The effect of acceptance and commitment therapy on the psychological flexibility and inflexibility of undergraduate students: A systematic review and three-level meta-analysis

Ti Hsu[a], **Jenna L. Adamowicz**[a], **Emily B. K. Thomas**[a]

[a]Department of Psychological and Brain Sciences, University of Iowa University of Iowa, G60 Psychological and Brain Sciences Building, Iowa City, Iowa 52242-1407

Abstract

Rising rates of mental health problems in undergraduate students is a critical public health issue. There is evidence supporting the efficacy of acceptance and commitment therapy (ACT) in decreasing psychological symptoms in undergraduates, which is thought to be facilitated through increases in psychological flexibility (PF) and decreases in psychological inflexibility (PIF). However, little is known about the effect of ACT on these processes in undergraduates. We conducted a systematic review and three-level meta-analysis examining this effect in 20 studies, which provided 56 effect sizes. A combined sample of 1,750 undergraduates yielded a small-to-medium overall effect ($g = .38$, $SE = .09$, $p < .001$, 95% CI: [0.20, 0.56]). This effect did not depend on control group type, intervention modality, number of sessions, the questionnaire used, whether PF or PIF was measured, or participant age. However, there was a significant mean effect only in studies with a specific clinical target, but not in those without one. Furthermore, the higher the percentage of female participants, the lower the reported effect size. Results suggested that ACT may increase PF and decrease PIF in undergraduates and highlighted various conceptual and measurement issues. Study protocol and materials were preregistered (https://osf.io/un6ce/).

Keywords

Acceptance and Commitment Therapy; Psychological Flexibility; Undergraduate Students; University Students; Multilevel Meta-Analysis

Introduction

Undergraduate Student Mental Health

The experience of pursuing an undergraduate degree is associated with heightened psychological distress (Pedrelli et al., 2015). The transition to college alone is associated with the onset of anxiety and depression, as well as heightened physiological and psychological stress (Rutter & Sroufe). In fact, out of 55,156 undergraduate students assessed during the 2020-2021 academic year in the U.S., 65.50% reported problems with anxiety, and 47.20% with depression (Center for Collegiate Mental Health, 2021). Additionally, among undergraduate students from 26 U.S. campuses who screened positive for depression, 76% reported at least one co-occurring disorder, such as generalized anxiety disorder (40%) and non-suicidal self-injury (37%; Eisenberg et al., 2013). Such mental health issues adversely affect students' quality of life, physical health, and academic performance (Petruzzello & Box, 2020).

It is important to acknowledge that pursuing an undergraduate degree is not a universal experience. For example, as of 2021, only 61.80 percent of high school graduates (ages 16-24) in the United States were enrolled in colleges or universities (U.S. Census Bureau, 2022). In terms of educational attainment in the United States, 39.60% of individuals 25-29 years of age have obtained a bachelor's degree or higher (National Clearinghouse Research Center, 2022). At the global level, 24% of all individuals who complete secondary education move on to pursue tertiary education within 5 years, even with large differences between countries (Roser & Ospina, 2013). Moreover, although undergraduate students are often characterized as individuals between the ages of 18-22 who are living on campus, attending classes full time, and are finally dependent on their parents (i.e., "traditional" students; Pelletier, 2010), there is a non-negligible proportion of undergraduate students who are "nontraditional," that is, students who have delayed their enrollment in college by a year or more after high school, are enrolled in classes part-time, are working full-time while enrolled, and are financially responsible for themselves or their families (Radford et al., 2015).

This variability does not belie the fact that pursuing an undergraduate degree is a stressful experience, especially if we consider research exploring differences in mental health and coping behaviors between undergraduate students and their non-university attending peers. Past research has shown that, compared to their non-college attending peers, undergraduates may experience higher levels of psychological distress (Cvetkovski et al., 2012) and report greater use of alcohol (Slutske et al., 2004) and other substances (Ford & Pomykacz, 2016). There is also evidence that undergraduate students may be more likely to employ maladaptive coping strategies (e.g., drinking to cope) when experiencing depression than their non-college attending peers (Kenney et al., 2018).

Furthermore, there are also contextual factors which highlight areas of growing need for undergraduate students. One example is the COVID-19 pandemic, which has contributed to increases in psychological distress and internalizing symptoms in undergraduates (Frazier et al., 2021; Son et al., 2020). Even prior to COVID-19, major barriers have interfered with undergraduate student access and uptake of mental healthcare. Results from a survey

completed by 13,984 students in eight countries provided descriptions of such barriers, including a preference to handle the problem alone, wanting to talk with friends or relatives instead, and being too embarrassed to seek help. These attitudinal barriers were rated as more important than structural barriers such as cost, transportation, and scheduling (Ebert et al, 2019). Students are not aware of how best to access mental health care, and report that internalizing symptoms are simply a part of the college experience (Eisenberg et al., 2007). Those who do identify a need for services are often skeptical about the efficacy of care and perceive available care to be inconvenient (Eisenberg et al., 2011). In sum, previous research suggests that undergraduate students may experience higher levels of psychological distress and negative coping than their non-college attending peers, and that these difficulties may be exacerbated by stress brought on by the COVID-19 pandemic as well as significant attitudinal barriers to accessing mental health care.

Efficacy of Acceptance and Commitment Therapy for Undergraduate Students

Given the critical state of undergraduate student mental health, it is crucial to identify and disseminate interventions that address these pressing needs. Acceptance and commitment therapy (ACT; Hayes et al., 1999) is a cognitive behavioral therapy that shows promise. A meta-analysis of 20 meta-analyses indicated that ACT is efficacious in addressing depression, substance use, chronic pain, and transdiagnostic clusters (Gloster et al., 2020). However, few systematic reviews have focused on ACT for undergraduate students, and none have focused on the effects of ACT on PF/PIF among undergraduate students. In fact, there are only two meta-analyses which include ACT as one out of many possible interventions addressing stress (Amanvermez et al., 2022) and depressive symptoms (Ma et al., 2021) in undergraduate students. Both meta-analyses focused only on online self-guided interventions in an individual format. Amanvermez and colleagues (2022) investigated the effects of stress management interventions for undergraduate students, and categorized such programs according to the theoretical orientation that the content was based upon, naming the following groups: 1) cognitive behavioral therapy (CBT); 2) third-wave therapies (e.g., interventions based on acceptance or mindfulness techniques, ACT, and mindfulness-based cognitive therapy); 3) skills training; and 4) mind-body interventions. Although there was a small and statistically significant effect of these online stress management interventions for perceived stress and internalizing symptoms overall, theoretical orientation did not significantly moderate this effect.

Ma and colleagues (2021) reviewed studies with online self-help interventions targeting depressive symptoms in undergraduates. Of the 19 primary studies they included, ten studies employed CBT, eight third-wave CBTs, and one involved a physical activity intervention. The third-wave interventions included five ACT studies and three studies that were based on mindfulness interventions. Findings revealed a small effect favoring intervention groups over controls overall, but moderation analyses comparing ACT-based interventions and CBT did not yield a significant effect. Therefore, although both of these meta-analyses include ACT interventions for college students, they do not focus exclusively on such interventions, and therefore are limited in what they can communicate specifically about the effects of ACT on undergraduates. They are also limited in that they only examined online self-guided interventions in an individual format, when ACT interventions for undergraduate students

exist in a variety of modalities and formats, including in person (Yadavaia et al., 2014), online (Gregoire et al., 2022), and through bibliotherapy (Muto et al., 2011), as well as in group (Fang et al., 2022) and individual formats.

Only one meta-analysis to date has exclusively examined the impact of ACT on undergraduate students (Howell & Passmore, 2019), but it focused on well-being, a construct that is related to but distinct from PF/PIF, as an outcome. Results indicated that, when compared to controls, ACT had a small, significant pooled effect on well-being ($d = 0.29$). The small number of meta-analyses or systematic reviews centered on ACT for undergraduate students is surprising, given the significant number of published clinical trials that support ACT's efficacy in decreasing psychological symptoms and promoting well-being in this group. Evidence suggests that compared to waitlist or active controls, students who receive ACT report increases in quality of life and well-being (Krafft et al., 2019), and decreases in symptoms of anxiety, depression, and stress (Larsson et al., 2022), as well as COVID-19 related distress (Copeland et al., 2021). The fact that ACT influences this diverse array of outcomes is often attributed to the therapy's transdiagnostic or process-based approach. That is, the claim that ACT is designed to strengthen certain psychological skills, and symptom reduction is the by-product of such learning (Ong et al., 2020). This philosophy is one of the defining hallmarks of ACT, compared to other cognitive behavioral therapies (CBTs) which emphasize symptom reduction as the primary metric of treatment efficacy (Linardon et al., 2017). Given evidence for ACT's efficacy across a wide range of disorders in undergraduate students, the question that follows is whether there is evidence that ACT affects the theorized psychological processes it was designed to target in this population.

Psychological Flexibility Processes and Undergraduate Students

The psychological flexibility (PF) or hexaflex model (Hayes et al., 1996) presents the processes thought to account for ACT's efficacy (Hayes et al., 2019). PF is composed of six processes: *acceptance, cognitive defusion, contact with the present moment, self-as-context, values, and committed action*. The model also includes the six processes of psychological *in*flexibility (PIF), the conceptual inverse of PF: *experiential avoidance, cognitive fusion, lack of contact with the present moment, self-as-content, lack of contact with values, and inaction*. A meta-analysis of ACT mediation studies suggested that improvements in mental health outcomes, quality of life, and well-being are all mediated by changes in PF processes (Stockton et al., 2019). However, few studies have provided evidence that ACT affects outcomes through PF in undergraduate students (e.g., Zhao et al., 2022), and the magnitude and direction of the effect across studies remains unexamined.

Examining PF in undergraduate students is important for several reasons. First, compared to older adults, undergraduate students may have stronger cognitive skills and less crystallized attitudes (Sears, 1986). These relatively higher levels of openness may make students more amenable to working on skills such as *self-as-context* and *acceptance*, which require perspective-taking and willingness to contact unpleasant thoughts, emotions, and physical sensations. Furthermore, a significant portion of undergraduates are emerging adults (ages 18-29), who are more flexible in adapting to multiple contexts compared to

adolescents (Kashdan & Rottenberg, 2010). Emerging adults also occupy the achieving stage of cognitive development (Schaie & Willis, 2000), in which focus is shifted from the acquisition to the application of knowledge, which may make undergraduates especially amenable to honing skills such as *committed action*, which involves behavior changes in service of values.

There is limited research, however, on whether the different PF processes proposed by the hexaflex model are differentially effective as treatment components, not just in undergraduates, but more generally speaking. One reason for this is the widespread use of the Acceptance and Action Questionnaire-II (AAQ-II; Bond et al., 2011) as a measure of PF. The AAQ-II, which contains 7 items, is a unidimensional measure, and does not assess all the components of the hexaflex model. Nevertheless, PF has been included as a secondary outcome in meta-analyses since 2017 (Association for Contextual Behavioral Science, ACBS, 2023). Small to large effects favoring ACT over controls have been observed in studies on chronic pain (Hughes et al., 2017; Trindade et al., 2021). Small effects have been found favoring self-help ACT and internet-based ACT over controls on PF (Thompson et al., 2021; French et al., 2017). Medium effects have been found favoring ACT over controls on PF in cancer patients (Zhao et al., 2021) as well as family caregivers (Han et al., 2020). Large effects have been found on studies on dysregulated eating (Di Sante et al., 2022). Some meta-analyses, however, have reported no significant effect of ACT on PF in healthcare professionals (Prudenzi et al., 2021) and patients with multiple sclerosis (Thompson et al., 2022). Of note, only two meta-analyses have focused on individual PF and PIF facets, finding small effects favoring ACT over controls on increasing *valued living* and *cognitive defusion* (Han & Kim, 2022), and reducing *cognitive fusion* (Han et al., 2020).

To summarize, the current meta-analytic evidence for the effects of ACT on PF and PIF suggests that there are small-to-medium effects of the intervention on overall PF, as well as some specific processes. However, these studies have focused on clinical and non-clinical adult populations, and no meta-analysis to date has directly focused on the effects of ACT on the PF or PIF of undergraduate students. Such a meta-analysis could provide an estimate of the overall effect and information regarding intervention and participant characteristics which may moderate this effect. Studies on ACT for undergraduates vary widely in terms of intervention characteristics and design. Treatment modalities range from in-person ACT workshops to ACT bibliotherapy, to internet-based ACT, and even to ACT mobile phone applications. Some implemented multi-session interventions whereas others employed single ACT sessions. These studies also cover diverse clinical indications, ranging from high depression (Zhao et al., 2022), to low self-compassion (Yadavaia et al., 2014). Examining which intervention and participant characteristics affect the overall effect could provide information regarding which features of interventions work best in which undergraduate students. Last, a meta-analysis would be able to examine how PF and PIF are assessed in this population, and whether the magnitude of the overall effect differs across measures.

Objectives of the Current Study

The current systematic review and meta-analysis examined the impact of ACT interventions on the PF and PIF of undergraduate students. Randomized controlled trials comparing ACT

to various control groups were evaluated to determine the effect of these interventions on the PF of undergraduate students at post-intervention and follow-up. Given that the aforementioned meta-analyses examining the effect of ACT on PF and PIF have identified small-to-medium effects at post-intervention and follow-up in clinical and non-clinical adult populations, we anticipated that we would also observe a similar small-to-medium significant effect favoring ACT over controls, since we did not expect that differences between intervention and control groups would necessarily be greater in undergraduates than in other samples of adults. We used a multilevel meta-analytic approach to estimate the overall effect, as well as conduct moderation analyses. We tested the effects of several study characteristics as well as participant characteristics as moderators of the overall effect size. Given that little is known about factors (e.g., type of controls, modality of sessions) which may influence the effect of ACT on undergraduate student PF and PIF, this is an especially important and novel aspect of the current study. Last, we tested the effect of publication bias in the overall effect size.

Method

We followed the Preferred Reporting of Items for Systematic Reviews and Meta-Analysis Protocols guidelines (Page et al., 2021; see Supplemental Table S1). Study protocol and materials were preregistered (https://osf.io/un6ce/).

Eligibility Criteria

Published and unpublished records were included if they: (a) were written in English, (b) included an ACT intervention with clearly defined content and duration, (c) employed a randomized controlled research design, (d) included undergraduate students (i.e., any student pursuing an undergraduate degree at a college or university), (e) included at least one validated self-report measure assessing PF/PIF or their facets, (f) had at least one pre-intervention, post-intervention, and/or follow-up measure of PF/PIF, or have an author agree to provide this information. Records were excluded if they: (a) did not include a non-ACT comparison group, (b) included experimental inductions of mood states, (c) included graduate or professional students.

Information sources and search strategy

The search strategy was developed by the first author and a librarian with expertise in systematic reviews. Extensive preliminary searching was conducted to identify databases and refine search strategies. Databases were chosen to include multiple disciplines, including psychology, education, and the health sciences. A systematic literature search was conducted in PubMed, APA PsycINFO (EBSCOhost), Scopus, Education Source (EBSCOhost) and ProQuest Dissertations and Theses Global. See Supplemental Table S2 for search strategies, which specified that records should include at least one term referring to undergraduate students and young/emerging adults, PF/PIF, and ACT in their titles, abstracts, and keywords. To address possible effects of publication bias, we included dissertations and theses in our search, and contacted authors about any unreported data.

The final search was conducted on 3/21/2023, and results were uploaded to EndNote (The EndNote Team, 2013) for deduplication. Results were next uploaded to Rayyan (Ouzzani et al., 2016) for initial screening of abstracts and titles. The full text of records that passed the initial screen were retrieved and assessed for eligibility. The initial abstract and title screening, as well as the screening of records that passed that first screen, was conducted by the first author, who consulted with the other authors if any questions or ambiguous cases arose. A backwards and forward citation search (Hinde & Spackman, 2014) was then conducted on eligible journal articles using Web of Science. We also hand-searched a list of ACT randomized controlled trials (Hayes, 2023), and ClinicalTrials.gov. Last, we contacted the ACBS and Association for Behavioral and Cognitive Therapies email listservs for any additional studies.

Selection of Outcome Variables

Measures of general PF/PIF and of PF/PIF facets were selected as outcomes. In anticipation of the widespread use of the AAQ-II, and its operationalization as both a measure of PIF (as validated) and reverse-scored as a measure of PF, coders recorded how the AAQ-II was scored by each study. However, for purposes of classification in the current study, we adhered to the AAQ-II's original designation as a measure of PIF. We only included instruments not originally designed to measure PF/PIF if authors provided conceptual links between the measured constructs and PF/PIF.

Selection of Moderator Variables

We examined both intervention and participant characteristics as moderators, including treatment indication (if participants met diagnostic criteria for psychological disorders), timepoint (post-intervention vs follow-up), length of follow-up, number of sessions, intervention format (group vs individual), intervention modality (online, in-person, bibliotherapy), control group type (active vs waitlist), instruments used to assess PF (AAQ-II vs not), and participant age, sex, and race.

Data Extraction

Two raters independently extracted and coded data from eligible studies. Raters attended preliminary training as well as weekly coding meetings and adhered to a standardized data extraction protocol. Overall percent agreement was 97.05%. The following variables were extracted from each study: study country; clinical indication, length of time to follow-up, intervention mode; intervention delivery; number of control groups; types of control groups; sample size; mean participant age; participant sex; participant race; intervention group size; control group size; number of PF/PIF outcomes; name(s) of PF/PIF outcomes; pre- and post-intervention and follow-up mean and standard deviation for PF /PIF outcome(s) for intervention and control group(s).

Study risk of bias assessment

Two coders used the Cochrane risk of bias tool (Higgins et al., 2011) to assess risk of bias across seven domains. Discrepancies were discussed until 100% agreement was reached. PF or PIF measures were also evaluated for quality. If studies did not report evidence of an

instrument's validity and reliability, then the effect size was excluded from analyses. Studies were also assessed on whether their protocols had been registered on ClinicalTrials.gov or the Open Science Framework.

Meta-Analytic Strategy

Effect Size Calculation—Hedges' *g* (Hedges & Olkin, 1985) values were calculated based on post-treatment and follow-up outcome differences between the intervention and control group as effect size measures. Outcomes were treated so that *increases* in general PF and PF facets were advantageous, whereas *decreases* in general PIF and PIF facets were advantageous. Effect sizes were calculated so that a positive effect size indicated an advantage (i.e., higher PF and lower PIF) of the intervention group at post-test or follow-up compared to controls.

The Three-Level Meta-Analytic Model—Almost half of all primary studies (45%) in this meta-analysis reported on more than one PF outcome, introducing the issue of dependency among effect sizes drawn from the same study. To address this, we employed a three-level meta-analytic model (Cheung, 2014; Van den Noortgate et al., 2013), which allows for the inclusion of all relevant effect sizes and addresses the issue of dependent effect sizes within studies by modeling three distinct sources of variation: variance between different studies at level 3 (i.e., between-study variance), variance among effect sizes derived from the same study at level 2 (i.e., within-study variance), and sampling variance of observed effect sizes at level 1 (i.e., sampling variance). Further, because level 2 accounts for sampling covariation, the multilevel approach does not require that the correlations between outcomes in primary studies be known. Last, the three-level meta-analytic model allows us to examine differences in outcomes within studies by estimating within-study heterogeneity, as well as differences between studies by estimating between-study heterogeneity. If there is indeed evidence of heterogeneity, moderator analyses can be conducted in the same framework with a three-level mixed effects model.

Software and Parameters—Referring to Assink and Wibbelink (2016), we fit three-level meta-analytic models in Rstudio (Rstudio Team, 2022) using the *metafor* package (Viechtbauer, 2010). The overall effect across primary studies was estimated using a three-level intercept only model, and potential moderators were examined by their addition as covariates. Restricted maximum likelihood was used to estimate model parameters, and Knapp and Hartung's (2003) method was used to test regression coefficients and confidence intervals. Continuous moderators were centered around their mean, and dummy variables were created for categorical variables (Tabachnik & Fidell, 2013).

Publication Bias Analysis

We tested for publication bias in multiple ways. First, we evaluated whether publication status moderated the overall effect size to assess the effect of publication status. Second, we constructed a funnel plot to examine the distribution of each effect size plotted against their standard error. Third, we used a multilevel extension of Egger's regression test (Rodgers & Pustejovsky, 2021) in the form of a multilevel random-effects model which includes the standard error of the effect sizes as a moderator to assess funnel plot asymmetry. The

intercept of this model deviating significantly from zero was indicative of the presence of the small-study effect.

Results

Study Selection

Figure 1 features a PRISMA 2020 Flow diagram (Page et al., 2021) delineating the study selection process. Our database search produced 1,621 studies, and 688 duplicates records were removed, resulting in 933 records subject to screening. The screening process excluded 873 studies, and 60 studies were assessed for eligibility, after which 41 studies were excluded. The most common reason for exclusion was the presence of graduate students in the sample (n = 11). See Supplemental Table S3 for elaboration and supplementary analyses including these studies. Our database search therefore yielded 19 eligible studies. Our backwards and forwards citation search identified an additional 1,749 records, 537 of which were duplicates. We next identified 12 studies from ACBS and ClinicalTrials.gov, which were also duplicates. 1,056 of the records of our citation search did not meet eligibility criteria, yielding a total of 12 studies. Eleven of these studies were duplicates of the studies identified through our database search. The ACBS and ABCT listservs yielded no additional eligible studies. These additional searches therefore yielded only one additional eligible study, resulting in a total of 20 primary studies (see Supplemental Table S4 for citations).

Study Characteristics

Please see Supplemental Table S5 for a summary of study characteristics. Of the 20 included studies, 15 (75%) were published journal articles, while five studies were master's theses or doctoral dissertations. Fifteen (75%) studies were conducted in North America (14 in the U.S., one in Canada), and five studies in other regions of the world (Iran, Finland, and China.) More than half of the studies (60%) used an intention-to-treat approach and included all randomized participants in statistical analyses. Most studies (70%) were published in or after 2016.

Participant Characteristics and Sample Size—Please also see Supplemental Table S5 for a summary of participant characteristics. The mean age across studies ranged from 18.73 to 30.90 (the latter mean was from a study conducted at an institution with a substantial proportion of nontraditional students). Weighted mean age across studies was 21.88 years ($SD = 3.03$). All studies reported on participant sex, and 3/20 (15%) reported on gender identity. On average, 65.63% of study samples were female. Of the studies (13/20; 65%) that reported on the racial/ethnic breakdown, 76.09% of participants across studies identified as White (12.59% Black/African American; 1.61% Native American/ Alaskan Native; 4.07% Asian; 0.63% Native Hawaiian and Pacific Islander; 2.02% Other/ Multiracial).

Intervention Characteristics—See Supplemental Table S6 for a summary of intervention characteristics. Nine studies (45%) did not require that participants meet diagnostic criteria for a psychological disorder. Of the 11 studies with specific treatment indications, three required that participants meet criteria for depression, two required that

participants meet criteria for social anxiety, one for anxiety regarding physical appearance, and the remaining studies for gambling pathology, impulsivity, negative body image, hoarding behaviors and low self-compassion and high distress. Sixteen (80%) studies had one control group, and four studies had two. Of these four studies, two included active control groups that were a variation of the ACT intervention, which were not considered in the current analysis. The other two studies employed active controls that were not variations of the ACT intervention applied, as well as a waitlist control in one study and a no-treatment control in another. Comparisons were made between the intervention group and both kinds of control groups for these two studies in our analyses.

Half of all studies employed active control groups, three studies (15%) had a no-treatment control, and nine studies (45%) employed a waitlist control, either alone or in addition to an active control. Only four studies (20%) employed single-session interventions. Fourteen studies (70%) had an individual rather than group intervention format. Eleven studies (55%) featured in-person interventions, five studies (25%) featured online interventions, one study (5%) featured a mobile app intervention, and three studies (15%) included bibliotherapy assisted by online modules. Nine studies (45%) assessed PF and PIF at follow-up, with a mean follow-up length of 1.91 months across studies. Only two studies had more than one follow-up timepoint.

PF and PIF Measurement—Please see Supplemental Table S7 for PF/PIF measures used in primary studies organized by hexaflex concepts, and Supplemental Table S8 for measure descriptions. Most studies (60%) assessed only PIF, six studies (30%) assessed PF and PIF, and two (10%) assessed only PF. The AAQ-II (Bond et al., 2011) was the most used measure across studies, with nine studies (45%) administering the AAQ-II and three studies (15%) administering the Chinese version. Most studies using the AAQ-II (11/13) operationalized it as a measure of PIF, in line with its original design, and two studies reverse-scored and used it as a measure of PF. Two studies used the Acceptance and Action Questionnaire (AAQ; Bond & Bunce, 2003; Hayes et al., 2004), the AAQ-II precursor. Three studies used "contextually sensitive" versions of the AAQ, such as the AAQ-Stigma (Levin et al., 2014), AAQ for Weight-Related Difficulties (Lillis & Hayes, 2008), and the Persian Acceptance and Action in Social Anxiety Questionnaire (Soltani et al., 2016).

Only two studies employed multidimensional measurements of PF and PIF, but one employed only the *values* and *lack of contact with values* subscales of the Multidimensional Psychological Flexibility Inventory (MPFI; Rolffs et al., 2016), and the other used only the general PF and PIF subscales of the MPFI-24 (Gregoire et al., 2020). Few studies used instruments designed to assess individual PF/PIF facets. Four studies used the Cognitive Fusion Questionnaire (Gillanders et al., 2014) and its Chinese translation (Zhang et al., 2014). The Valuing Questionnaire (Smout et al., 2014) was used by two studies, and the Valued Living Questionnaire (Wilson et al., 2010) by one.

Studies also employed measures not designed to assess PF as proxy measures of the construct. Dixon et al. (2016) used the Mindful Attention Awareness Scale (Brown & Ryan, 2003) to assess *contact with the present moment*. Levin et al. (2016) employed the acting with awareness subscale of the Five Facet Mindfulness Questionnaire (Baer et al.,

2006) to assess *contact with the present moment* and the nonreactivity subscale to assess *acceptance*. No study measured all facets of PF or PIF. Three facets of PF (*committed action, self-as-context, cognitive defusion*), and four of PIF (*lack of contact with the present moment, self-as-content, experiential avoidance, inaction*) were not assessed in any study.

Risk of Bias Assessment

Results for the risk of bias assessment are shown in Supplemental Figure S1. Less than half of all studies (45%) described the inclusion of a random component in the sequence generation process and received low risk of bias ratings for this domain. One study received a high risk of bias rating in this domain, due to the description of a non-random element in its randomization system. Most studies did not describe allocation concealment, resulting in "unclear" ratings. All but one study reported on planned outcomes, indicating low risk of reporting bias across studies. Most studies did not report on the masking of participants and personnel, or of the outcome assessment, resulting in unclear effects of performance and detecting bias. One study received a high risk of bias rating due to its acknowledgement that masking therapists is impossible in psychotherapy research. Every study but one received a "low" rating for attrition bias, indicating that most studies described methods of handling incomplete outcome data. In terms of other risk of bias, coders acknowledged that there may have been various sources of additional bias, but that there was insufficient information to assess whether risk of bias exists. Therefore, all studies received an "unclear" rating for risk of other bias. In terms of risk of bias for PF/PIF measures, all studies reported on the validity and reliability of measures. In terms of study preregistration, three studies (15%) reported being registered, with two on ClinicalTrials.gov, and one in the Iranian Registry of Clinical Trials.

Average Effect of ACT Interventions

In total, 20 studies provided 56 estimates of effect sizes at post-intervention and follow-up. A combined sample of 1,750 students yielded a significant overall effect size $g = 0.38$ ($SE = .09$, $p < .001$, 95% CI: [0.20, 0.56]), which falls in the small-to-medium range according to interpretation guidelines (small = 0.20, medium = 0.50, large = 0.80; Cohen 1992). Analyses revealed significant heterogeneity across studies ($\sigma^2_{level\ 3} = 0.12$, $p = .001$), but no significant variability between effect sizes extracted from the same study ($\sigma^2_{level\ 2} = 0.001$, $p = .92$). Only 0.86% of the variance among effect sizes was accounted for by the within-study level, 73.13% of the variance could be attributed to the between-study level, and the remaining 26.01% was attributable to random sampling error. Moderation analyses were conducted to examine this substantial between-study variability.

Moderators of ACT Interventions for Undergraduate Students

Table 1 presents the results of the moderation analysis. We first examined potential moderating effects of study characteristics. Results showed that there was significant variability in effect sizes based on whether participants were recruited based on elevated clinical symptoms ($F(1, 54) = 4.09$, $p = .048$). There was a positive significant effect of ACT on PF in studies that focused on specific clinical indications ($B_0 = 0.53$, $t = 4.71$, $p < .001$), but no significant effects on the mean effects of studies that had non-specific

clinical indications (B_0 =0.20, t = 1.65, p = .10). However, studies with non-specific clinical indications also had significantly lower effect sizes than those with significant clinical indications on average (B_1 =−0.33, t = −2.02, p = .048).

The overall effect of ACT on PF was not moderated by control group type ($F(1,54)$ = 2.76, p = .10), but ACT had significant mean effects in studies with waitlist control groups (B_0 =0.48, t = 4.66, p < .001), as well as those with active controls (B_0 =0.25, t = 2.25, p = .03). Next, we found that the overall effect did not depend on whether the intervention was delivered in a group or individual format ($F(1, 54)$ = 2.23, p = .14), although significant mean effects of ACT were observed in both group interventions (B_0 = 0.57, t = 3.66, p < .001) and individual ones (B_0 = 0.29, t = 2.80, p = .007). The overall effect was not moderated by whether the intervention was delivered in-person, online, or through bibliotherapy ($F(1,53)$ = 1.11, p = .34). There was a positive and significant mean effect of ACT for interventions delivered in person (B_0 =0.50, t = 4.00, p < .001), but not for those delivered online (B_0 =0.30, t = 1.95, p = .06) or through bibliotherapy (B_0 =0.16, t =0.72, p = .48). Number of sessions also did not moderate the overall effect ($F(1,54)$ = 1.16, p = .29), though there was a significant mean effect of ACT for interventions with multiple sessions (B_0 =0.43, t =4.26, p < .001), but not with single-session interventions (B_0 =0.04, t =0.16, p = .13).

Although choice of measure did not moderate the overall effect ($F(1,54)$ = 2.64, p = .11), ACT had significant mean effects for studies using AAQ-II and derivatives (B_0 =.33, t = 3.51, p <.001) and those that did not (B_0 =.46, t = 4.44, p <.001). Whether PF or PIF was assessed also did not moderate the overall effect ($F(1,54)$ = 1.26, p = .27), although there were significant mean effects for effect sizes derived from both PF (B_0 =0.38, t = 3.77, p <.001) and PIF measures (B_0 =0.30, t = 3.49, p <.001). Whether an effect size was from post-intervention or follow-up did not moderate the overall effect ($F(1,54)$ = 0.10, p = .75). However, there were significant mean effects at both post-intervention (B_0 =0.39, t =4.42, p <.001) and follow-up (B_0 =0.37, t = 3.65, p <.001). Finally, length of follow-up did not moderate the overall effect ($F(1,54)$ = 0.35, p = .56).

Next, we examined the potential moderating effects of participant characteristics on the overall effect. Results showed that the overall effect was not moderated by participant age ($F(1, 54)$ = 0.17, p = .69). However, the overall effect of ACT on PF was moderated by participant sex ($F(1, 54)$ = 5.87, p = .02), such that the higher the percentage of female participants, the lower the reported effect sizes in the primary studies (B_1 = −0.01, t = −2.48, p = .02). There were no differences in effect size based on the percentage of participants from any racial group.

Publication Bias Analysis

We first examined the possible moderating effect of publication status on the overall effect size. Analyses revealed that the overall effect was not moderated by whether the study was an unpublished dissertation or thesis (B_0 = .34, t = 1.91, p = .06), or whether it was a published study (B_0 =.39, t =3.70, p < .001; omnibus $F(1,54)$ = 0.05, p = .83), indicating the magnitude of effect sizes did not differ on the basis of whether it was drawn from a published or unpublished source. Supplemental Figure S2 features a funnel plot constructed

to examine the distribution of each effect size plotted against its precision (i.e., the inverse of the standard error). To account for the presence of multiple effect sizes per study, the plot was color coded so that effect sizes from the same study were the same color. This funnel plot appeared symmetric. The intercept of the multilevel extension of Egger's regression test did not deviate significantly from zero ($B_0 = -0.26$; $t = -0.93$ $p = .36$), indicating that the overall relationship between the sample sizes and their precision is symmetrical, and not indicative of the presence of the small-study effect.

Discussion

Overall Effects of ACT for Undergraduate Students

Analyses of the data from 20 studies of ACT for undergraduate students revealed a small-to-medium overall effect of ACT on PF ($g = 0.38$) at post-intervention and follow-up, favoring ACT over controls. The direction and magnitude of this effect is consistent with those from other meta-analyses on the effect of ACT on PF in general and clinical adult samples at post-intervention (Thompson et al., 2021; French et al., 2017). Thompson et al. (2017) also examined effects of iACT on PF at follow-up in a separate analysis, showing comparable effect sizes at each timepoint (pre-post $g = 0.32$; pre-follow-up $g = 0.36$). The three-level meta-analytic framework employed by our study allowed for effect sizes taken both at post-intervention and at follow-up to be included in the same model, and for a direct test of whether timepoint and length of follow-up affected the overall effect. Results indicated that the overall effect did not differ based on whether effect sizes were drawn from post-intervention or follow-up, nor on the length of the follow-up. Although previous research has reported an incubation effect of ACT on symptom-related outcomes such as smoking cessation and overall distress at follow-up (Gifford et al., 2004; Clarke et al., 2014), our finding is among the first to provide that this effect may also exist for PF across multiple studies.

It is also important to consider this overall effect in light of the one other meta-analysis to date which has examined the effects of ACT on undergraduate students. Howell & Passmore (2019) reported a small-to-medium ($d = 0.29$) effect of ACT on well-being. This effect falls in the same range as the estimated overall effect in this study. Previous research suggests that PF may mediate the effect of ACT on outcomes (Stockton et al., 2019), so one question to be explored by future meta-analyses is whether there is a significant mediating effect of PF between ACT and outcomes across studies, and the extent to which changes in PF contribute to symptom reduction. However, there is currently a dearth of mediational research on PF due to methodological challenges (Arch et al., 2022), and this is why this question was not explored by the current review.

The use of multilevel meta-analysis allowed us to examine variability in effect sizes at both the between- and within-study levels. Results suggested that there was no significant variability between effect sizes extracted from the same study. This finding is unsurprising given that 50% of studies featured only one measure of PF (i.e., the AAQ-II), and several studies including more than one measure featured the AAQ-II or a variant. Another explanation for this is that most studies had a small number of measures, with only one study using three measures, and the rest one or two.

Moderators of the Effectiveness of ACT for Undergraduate Students

Treatment indication emerged as a significant moderator of the overall effect, such that there was a positive significant effect of ACT on PF in studies that recruited participants who met certain cutoffs for psychological symptoms, but no significant effect for studies that did not require cutoffs. This moderation effect approached the threshold for statistical significance ($p = .048$) and should be interpreted with caution. However, given that primary studies reported screening for a range of different symptoms, ranging from low self-compassion to high depression, impulsivity, and hoarding behaviors, this finding seems to still suggest that PF skills are enhanced across a range of different symptom targets, speaking to the transdiagnostic nature of ACT. It is also possible that interventions with specific treatment indications recruited more participants with more severe symptoms, lower PF and higher PIF at baseline. These participants therefore may have had more room to improve compared to participants whose symptomology was not as severe. Whether studies featured an active or a waitlist control group did not affect the magnitude of the overall effect, and ACT interventions yielded significant effect sizes for waitlist and active controls. Of note, there was a wide variety of active control groups across primary studies, and the categorization of control group types into active or waitlist controls may have masked the effects of certain active control types. However, there were not enough studies with specific active controls for separate testing.

Results did not demonstrate a moderating effect of group versus individual formats, with ACT interventions yielding significant effect sizes for both formats. This result provides preliminary evidence that ACT delivered in a group format is no less effective than as in an individual format in promoting undergraduate student PF. Previous studies have demonstrated the efficacy of ACT groups, as well as delineated the advantages of delivering the therapy in this format (see Walser & Pistorello, 2004 for a review). Specifically, ACT's experiential exercises and metaphors lend themselves well to being presented in groups, where clients can receive support and validation, as well as alternative points of view from peers. Group members may also help each other understand abstract or difficult concepts. These elements may be especially powerful for undergraduate students, who may feel socially isolated and in need of interpersonal support, especially during the transition to college. Furthermore, the delivery of ACT in groups, rather than on an individual basis, may be a more cost- and resource-effective way of disseminating therapy resources on college or university campuses.

There was also no significant moderating effect of intervention delivery, but there was a significant mean effect of ACT interventions delivered in person, but not online or through bibliotherapy. Although previous research supports the efficacy of iACT compared to face-to-face ACT (e.g., Lappalainen et al., 2014). It is possible that internet-based modalities were not as engaging as in-person for undergraduate students, and were therefore less effective at promoting PF. Most of the online and bibliotherapy interventions included in primary studies featured limited interactions with interventionists, and participants completed treatment at their own self-guided pace. It is important to note that internet-based and self-guided ACT do offer participants more freedom regarding where and often when they receive the intervention and may potentially address some of the structural barriers

that prevent undergraduate students from accessing mental health care. However, it is also possible that students were more inconsistent with how often they engaged with online or bibliotherapy intervention materials compared to students who were required to attend intervention sessions in-person. It is possible that this inconsistency may have "diluted" the effects of the self-guided interventions, compared to in-person delivery. Future research should examine whether this variability affects the effectiveness of ACT interventions through direct comparisons between in-person and self-guided modalities.

Similarly, there was no significant moderating effect of the number of sessions, although multiple session interventions had a significant mean effect, but single-sessions interventions did not. This finding is curious, especially because there is some evidence that single-session ACT can be as short as 90 minutes and be no less effective than workshops that span three or six hours (Kroska et al., 2020). Single-session interventions also provide undergraduate students with an opportunity to be introduced to skills in a format that requires little time commitment. This may address some of the structural and attitudinal barriers that prevent undergraduates from seeking mental healthcare. At the same time, however, single session interventions may not provide enough time for students to become familiar with ACT skills and concepts for there to be a significant effect on their levels of PF. It is an important future direction to determine, on average, how many hours of intervention is required to result in lasting change in undergraduate student PF. Or even for research to explore change in PF over more extended periods of time. Overall, these results indicate that differences in efficacy (or lack thereof) based on treatment design is an important consideration when evaluating service delivery options for undergraduate students.

The PF measure used (i.e., whether the AAQ-II or a derivative was employed, or another measure of PF or PIF) did not moderate the overall effect. Again, it is important to note here that 15 of the 20 primary studies used the AAQ-II as either their sole or primary measure of PF/PIF, and there was an extremely limited range of other PF/PIF measures employed. The lack of a significant moderating effect is likely due to this lack of variability in the type of measures used. This is also a major limitation of the primary studies included in this meta-analysis and is discussed in detail later in this section. Next, whether primary studies assessed outcomes with PF or PIF measures also did not moderate the overall effect. This lack of a significant effect may also be attributed to the relatively small number of effect sizes drawn from measures of PF.

Next, participant race and age did not significantly moderate the overall effect. Given the relatively restricted age range of undergraduate student samples, and the fact that most studies consisted of primarily White samples from the U.S., this result is not unexpected. It is important to note, however, that several studies did not directly report on racial/ethnic breakdown, and instead described participant nationality and therefore were not included in analyses. Participant sex significantly moderated the overall effect, such that the higher the percentage of female participants in a sample, the lower the reported effect size. This was surprising, since there is little evidence suggesting that ACT improves PF differentially across sex. In terms of our publication bias analysis, we found no evidence of small study effects, or evidence of bias based on publication status. However, we were not able to test for selective reporting across studies, as existing tests (e.g., 3PSM models, p-curve analyses) do

not perform well in multilevel meta-analysis. Therefore, selective reporting cannot be ruled out, and the results of this review should be interpreted with caution.

Limitations of Evidence Included in Current Review

Our systematic review highlighted several notable issues regarding PF/PIF measurement. First was the anticipated, widespread use of the AAQ-II as the sole or primary measure of PF or PIF. More than half of the primary studies administered the AAQ-II. Across these studies, the AAQ-II was mostly scored so that higher scores indicated greater PIF, but sometime reverse-scored so that high scores indicated greater PF. This practice erroneously assumes that PF and PIF lie on a single dimension. Relatedly, as a unidimensional measure, the AAQ does not assess all components of PF and PIF, posing issues of content validity and incongruence with the PF model. However, the use of the AAQ-II persisted even in primary studies published after the introduction of multidimensional measures in the 2010s. In fact, only two primary studies used a multidimensional measure of PF/PIF, but neither assessed all facets of these constructs. Gregoire and colleagues (2022) employed the MPFI-24 (Gregoire et al., 2020) to only assess overall PF and PIF. Clark (2019) used the MPFI to only assess *connection with values* and *lack of connection with values*.

Studies published prior to this often-used measures of other constructs as proxies for PF/PIF components. For instance, measures originally designed to measure mindfulness, such as the MAAS and the FFMQ were used by some studies (i.e., Dixon et al., 2016; Levin et al., 2016) to measure facets of PF such as *contact with the present moment* and *acceptance*. Although these authors made theoretical arguments supporting the use of these measures, there is a lack of psychometric evidence suggesting that the measures originally designed to assess mindfulness indeed measure these PF facets. Another common strategy used by studies published prior to the development of multidimensional PF/PIF measures was employing multiple measures designed to measure individual facets (e.g., the CFQ, VLQ, or VQ), but these studies were not comprehensive in assessing all facets of the construct. These trends also highlight that there are PF/PIF facets that are not widely assessed in the literature, namely *committed action, self-as-context, cognitive defusion, lack of contact with the present moment, self-as-content, experiential avoidance, and inaction*. The absence of the assessment of these facets in current research results in a dearth of information regarding whether they are indeed being effectively targeted by intervention efforts. It is also crucial that future studies assess all PF/PIF facets to provide more theoretically consistent information on ACT's effect on its purported mechanisms of change. Multidimensional but brief assessments of PF, such as the CompACT-15 (Hsu et al., 2023), which has demonstrated good psychometric qualities and longitudinal measurement invariance, may assist in this effort. These issues regarding PF and PIF measurement are also important to consider when interpreting the results of the current meta-analysis, because they highlight content gaps that are not accounted for by our estimated overall effect due to a lack of multidimensional assessments of PF and PIF in the current literature.

As documented in our preregistered protocol, we had aimed to explore the extent to which ACT intervention studies focused on non-college-attending emerging adults, to address the fact that attending college or university is not a universal experience. In fact, only 39.60%

of individuals aged 25-29 in the U.S. have attained a bachelor's degree or higher (National Clearinghouse Research Center, 2023). We included search terms that broadened our search from undergraduates alone to their non-college attending, same-aged peers, but found no studies testing ACT exclusively in non-college-attending young adults. We therefore did not conduct exploratory moderation analysis to compare the effects of ACT on the PF of college-attending and non-college attending young adults.

Another limitation is that most primary studies were conducted in the U.S., and studies which reported on participant race/ethnicity had majority White samples. Of note, most studies conducted outside of the U.S. did not report on race/ethnicity, but instead on the nationality of participants, and there may be cultural differences in the conceptualization of race/ethnicity across countries. This distinctions between race, ethnicity, and nationality should be considered in future research. Study samples of our primary studies were also mostly female, with only three primary studies reporting on gender identity, and no studies reporting on both sex and gender identity. These patterns highlight a lack of racial as well as gender identity diversity in primary study samples. Exploring the efficacy of ACT for diverse groups is an important future direction.

It is insufficient to just acknowledge that findings of our meta-analysis may not generalize to the effect of ACT interventions on undergraduate students in other areas of the world. This is an important area of future research. There are other background variables that were not assessed by primary studies that could potentially provide important contextual information, including the type of institution attended, first-generation student status, and family socioeconomic status. In sum, there are challenges to the generalizability of current findings on the efficacy of ACT on undergraduate students, as well as more unassessed factors that also contribute to this trend. Future studies should include assessments of these background factors not only to promote generalizability, but also to explore whether there are differential effects based on these factors.

Another issue presented by the primary studies in the current meta-analysis is the variability in the duration of interventions. Although most studies which delivered in-person interventions reported on the length of the sessions in hours and the total number of sessions, other studies which featured an alternative delivery modality (e.g., websites, mobile apps) often reported the period during which participants had access to interventions, but not the length of active engagement with materials. Last, although registering clinical trials reduces selective reporting and publication bias (Viergever & Ghersi, 2011), only three primary studies were registered.

Study Limitation and Strengths

Limitations of the Review Processes—Meta-analyses are dependent on the availability, quality, and features of primary studies. We tested a range of methodological factors which may have contributed to the overall effect. However, the fact that few moderators significantly contributed to variability in PF indicates that there may be other unmeasured or unreported factors that could explain some of the heterogeneity between our primary studies. The current study also only included primary studies written in English. It is worth mentioning here again that this exclusion criteria may have resulted in the

omission of eligible studies because of publication language. It is also important to note that we excluded studies that included graduate students in addition to undergraduate students. Interestingly, our supplemental analyses which included studies that featured both undergraduate and graduate students did not indicate that the overall effect differed according to the percentage of graduate students in the sample. However, studies ranged widely in the number of graduate students included, and future research should directly assess whether ACT has different effects on PF and other outcomes in graduate students compared to undergraduates.

The results of the risk of bias assessment should be interpreted with several caveats. For instance, the masking of personnel to treatment conditions is impractical because clinicians must be aware of the intervention they are delivering. In fact, therapist allegiance has documented effects on outcomes in randomized controlled trials (Leykin & DeRubeis, 2009). Furthermore, because all primary studies utilized self-report questionnaires as outcome assessments, the masking of outcome assessors to participant allocation to groups also did not apply.

Strengths of the Current Study—The current meta-analysis provides the first assessment of the impact of ACT interventions on the PF and PIF of undergraduate students. We employed a multilevel meta-analytic framework, which allowed for the inclusion of multiple effect sizes per study and the examination of variability at the within- and between-study levels. Our systematic review broke down PF and PIF measures by concept, highlighting several measurement issues relevant to future measurement refinement efforts. Another strength of this study is the comprehensive nature of our search strategy. We adhered to PRISMA guidelines and worked with a librarian with expertise in systematic reviews. The search strategy was refined through extensive preliminary searching and was applied to databases from multiple academic disciplines. We also conducted a backwards and forwards citation search, which allows the searcher to capture parallel topic areas that may not be covered by keyword searching alone. We also hand-searched two online databases and emailed the two listservs for studies.

Conclusion

In sum, results provide support for the use of ACT in multiple forms in enhancing PF and decreasing PIF in undergraduate students, a population at high risk for mental health concerns but also lacking access to treatments. Findings also highlight that this effect was significant regardless of whether the effect was drawn from post-intervention or follow-up. Our review emphasized a critical need for future trials to use reliable and validated measures which reflect the multifaceted nature of PF and PIF. Although our findings underline some flaws in the current literature, they may also inform implementation of future ACT interventions that best serve undergraduate students.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

References

Amanvermez Y, Zhao R, Cuijpers P, de Wit LM, Ebert DD, Kessler RC, Bruffaerts R, Karyotaki E (2022). Effects of self-guided stress management interventions in college students: A systematic review and meta-analysis. Internet Interventions, 28, 100503. 10.1016/j.invent.2022.100503

Arch JJ, Fishbein JN, Finkelstein LB, & Luoma JB (2022). Acceptance and commitment therapy processes and mediation: Challenges and how to address them. Behavior Therapy, S0005789422000892. 10.1016/j.beth.2022.07.005

Assink M, & Wibbelink CJM (2016). Fitting three-level meta-analytic models in R: A step-by step tutorial. The Quantitative Methods for Psychology, 12(3), 154–174. 10.20982/tqmp.12.3.p154

Association for Contextual Behavioral Science (2023) Meta-analyses and systematic, scoping, or narrative reviews of the ACT evidence base. https://contextualscience.org/metaanalyses_and_systematic_scoping_or_narrative_reviews_of_the_act_evidence

Baer RA, Smith GT, Lykins E, Button D, Krietemeyer J, Sauer S, Walsh E, Duggan D, & Williams JMG (2008). Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. Assessment, 15(3), 329–342. 10.1177/1073191107313003

Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, Waltz T, & Zettle RD (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. Behavior Therapy, 42(4), 676–688. 10.1016/j.beth.2011.03.007

Brown KW, & Ryan RM (2003). The benefits of being present: Mindfulness and its role in psychological well-being. Journal of Personality and Social Psychology, 84(4), 822–848. 10.1037/0022-3514.84.4.822

Center for Collegiate Mental Health. (2022, January). 2021 Annual Report (Pub No. STA 22132)

Cheung MW-L (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. Psychological Methods, 19(2), 211–229. 10.1037/a0032968

Cohen J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159. 10.1037/00.3.3-2909.112.1.155

Conrad RC, Hahm H. Chris, Koire A, Pinder-Amaker S, & Liu CH (2021). College student mental health risks during the COVID-19 pandemic: Implications of campus relocation. Journal of Psychiatric Research, 136, 117–126. 10.1016/j.jpsychires.2021.01.054

Copeland WE, McGinnis E, Bai Y, Adams Z, Nardone H, Devadanam V, Rettew J, & Hudziak JJ (2021). Impact of COVID-19 pandemic on college student mental health and wellness. Journal of the American Academy of Child & Adolescent Psychiatry, 60(1), 134–141.e2. 10.1016/j.jaac.2020.08.466

Cvetkovski S, Reavley NJ, & Jorm AF (2012). The prevalence and correlates of psychological distress in Australian tertiary students compared to their community peers. The Australian and New Zealand Journal of Psychiatry, 46(5), 457–67. 10.1177/0004867411435290

Di Sante J, Akeson B, Gossack A, & Knäuper B (2022). Efficacy of ACT-based treatments for dysregulated eating behaviours: A systematic review and meta-analysis. Appetite, 171, 105929. 10.1016/j.appet.2022.105929

Ebert DD, Mortier P, Kaehlke F, Bruffaerts R, Baumeister H, Auerbach RP, Alonso J, Vilagut G, Martínez KU, Lochner C, Cuijpers P, Kuechler A, Green J, Hashing P, Lapsley C, Sampson NA, Kessler RC, & On behalf of the WHO World Mental Health—International College Student Initiative collaborators. (2019). Barriers of mental health treatment utilization among first-year college students: First cross-national results from the WHO World Mental Health International College Student Initiative. International Journal of Methods in Psychiatric Research, 28(2), e1782. 10.1002/mpr.1782

Eisenberg D, Golberstein E, & Gollust SE (2007). Help-seeking and access to mental health care in a university student population. Medical Care, 45(7), 594–601.

Eisenberg D, Hunt J, & Speer N (2013). Mental health in American colleges and universities: Variation across student subgroups and across campuses. The Journal of Nervous and Mental Disease, 201(1), 60–67. 10.1097/NMD.0b013e31827ab077

Eisenberg D, Hunt J, Speer N, & Zivin K (2011). Mental health service utilization among college students in the United States. Journal of Nervous and Mental Disease, 199(5), 301–308. 10.1097/NMD.0b013e3182175123

The EndNote Team. (2013). EndNote (EndNote X9). Clarivate.

Fang S, Ding D, Ji P, Huang M, & Hu K (2022). Cognitive defusion and psychological flexibility predict negative body image in the Chinese college students: Evidence from acceptance and commitment therapy. International Journal Of Environmental Research And Public Health, 19(24), 16519. 10.3390/ijerph192416519

Ford JA, & Pomykacz C (2016). Non-Medical use of prescription stimulants: A comparison of college students and their same-age peers who do not attend college. Journal of Psychoactive Drugs, 48(4), 253–260. 10.1080/02791072.2016.1213471

Frazier P, Liu Y, Asplund A, Meredith L, & Nguyen-Feng VN (2021). US college student mental health and COVID-19: Comparing pre-pandemic and pandemic timepoints. Journal of American College Health, 1–11. 10.1080/07448481.2021.1987247

French K, Golijani-Moghaddam N, & Schröder T (2017). What is the evidence for the efficacy of self-help acceptance and commitment therapy? A systematic review and meta-analysis. Journal of Contextual Behavioral Science, 6(4), 360–374. 10.1016/j.jcbs.2017.08.002

Gillanders DT, Bolderston H, Bond FW, Dempster M, Flaxman PE, Campbell L, Kerr S, Tansey L, Noel P, Ferenbach C, Masley S, Roach L, Lloyd J, May L, Clarke S, & Remington B (2014). The development and initial validation of the cognitive fusion questionnaire. Behavior Therapy, 45(1), 83–101. 10.1016/j.beth.2013.09.001

Gloster AT, Walder N, Levin ME, Twohig MP, & Karekla M (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. Journal of Contextual Behavioral Science, 18, 181–192. 10.1016/j.jcbs.2020.09.009

Grégoire S, Beaulieu F, Lachance L, Bouffard T, Vezeau C, & Perreault M (2022). An online peer support program to improve mental health among university students: A randomized controlled trial. Journal of American College Health. 10.1080/07448481.2022.2099224

Grégoire S, Gagnon J, Lachance L, Shankland R, Dionne F, Kotsou I, Monestès J-L, Rolffs JL, & Rogge RD (2020). Validation of the English and French versions of the multidimensional psychological flexibility inventory short form (MPFI-24). Journal of Contextual Behavioral Science, 18, 99–110. 10.1016/j.jcbs.2020.06.004

Han A, & Kim TH (2022). The effects of internet-based acceptance and commitment therapy on process measures: Systematic review and meta-analysis. Journal of Medical Internet Research, 24(8), e39182. 10.2196/39182 [PubMed: 36040783]

Han A, Yuen HK, Lee HY, & Zhou X (2020). Effects of acceptance and commitment therapy on process measures of family caregivers: A systematic review and meta-analysis. Journal of Contextual Behavioral Science, 18, 201–213. 10.1016/j.jcbs.2020.10.004

Hayes SC (2022). ACT Randomized Controlled Trials since 1986. https://contextualscience.org/act_randomized_controlled_trials_since_1986

Hayes SC, Strosahl KD, & Wilson KG (1999). Acceptance and commitment therapy: The process and practice of mindful change. Guilford Press.

Hayes SC, Strosahl K, Wilson KG, Bissett RT, Pistorello J, Toarmino D, Polusny MA, Dykstra TA, Batten SV, Bergan J, Stewart SH, Zvolensky MJ, Eifert GH, Bond FW, Forsyth JP, Karekla M, & McCurry SM (2004). Measuring experiential avoidance: A preliminary test of a working model. The Psychological Record, 54(4), 553–578. 10.1007/BF03395492

Hayes SC, Wilson KG, Gifford EV, Follette VM, & Strosahl K (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. Journal of Consulting and Clinical Psychology, 64(6), 1152–1168. 10.1037/0022-006X.64.6.1152 [PubMed: 8991302]

Hedges LV, & Olkin I (1985). Statistical methods for meta-analysis. San Diego, CA: Academic Press.

Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JAC, Cochrane Bias Methods Group, & Cochrane Statistical Methods Group. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ, 343(oct18 2), d5928–d5928. 10.1136/bmj.d5928 [PubMed: 22008217]

Hinde S, & Spackman E (2015). Bidirectional citation searching to completion: An exploration of literature searching methods. PharmacoEconomics, 33(1), 5–11. [PubMed: 25145803]

Howell AJ, & Passmore H-A (2019). Acceptance and commitment training (ACT) as a positive psychological intervention: A systematic review and initial meta-analysis regarding ACT's role in well-being promotion among university students. Journal of Happiness Studies, 20(6), 1995–2010. 10.1007/s10902-018-0Q27-7

Hsu T, Hoffman L, & Thomas EBK (2023). Confirmatory measurement modeling and longitudinal invariance of the CompACT-15: A short-form assessment of psychological flexibility. Psychological Assessment, 10.1037/pas0001214

Hughes LS, Clark J, Colclough JA, Dale E, & McMillan D (2017). Acceptance and commitment therapy (ACT) for chronic pain: A systematic review and meta-analyses. The Clinical Journal of Pain, 33(6), 552–568. 10.1097/AJP.0000000000000425 [PubMed: 27479642]

Kashdan TB, & Rottenberg J (2010). Psychological flexibility as a fundamental aspect of health. Clinical Psychology Review, 30(7), 865–878. 10.1016/j.cpr.2010.03.001 [PubMed: 21151705]

Kim H, Rackoff GN, Fitzsimmons-Craft EE, Shin KE, Zainal NH, Schwob JT, Eisenberg D, Wilfley DE, Taylor CB, & Newman MG (2022). College mental health before and during the COVID-19 pandemic: Results from a nationwide survey. Cognitive Therapy and Research, 46(1), 1–10. 10.1007/s10608-021-10241-5 [PubMed: 34177004]

Krafft J, Potts S, Schoendorff B, & Levin ME (2019). A Randomized Controlled Trial of Multiple Versions of an Acceptance and Commitment Therapy Matrix App for Well-Being. Behavior Modification, 43(2), 246–272. 10.1177/0145445517748561 [PubMed: 29262693]

Kroska EB, Roche AI, & O'Hara MW (2020). How much is enough in brief acceptance and commitment therapy? A randomized trial. Journal of Contextual Behavioral Science, 15, 235–244. 10.1016/j.jcbs.2020.01.009

Lappalainen P, Granlund A, Siltanen S, Ahonen S, Vitikainen M, Tolvanen A, & Lappalainen R (2014). ACT Internet-based vs face-to-face? A randomized controlled trial of two ways to deliver acceptance and commitment therapy for depressive symptoms: An 18-month follow-up. Behaviour Research and Therapy, 61, 43–54. 10.1016/j.brat.2014.07.006 [PubMed: 25127179]

Larsson A, Hartley S, & McHugh L (2022). A randomised controlled trial of brief web-based acceptance and commitment Therapy on the general mental health, depression, anxiety and stress of college Students. Journal of Contextual Behavioral Science, 24, 10–17. 10.1016/j.jcbs.2022.02.005

Levin ME, Haeger JA, Pierce BG, & Twohig MP (2017). Web-based acceptance and commitment therapy for mental health problems in college students: A randomized controlled trial. Behavior Modification, 41(1), 141–162. 10.1177/014S4455166S9645 [PubMed: 27440189]

Levin ME, Luoma JB, Lillis J, Hayes SC, & Vilardaga R (2014). The Acceptance and action questionnaire – stigma (AAQ-S): Developing a measure of psychological flexibility with stigmatizing thoughts. Journal of Contextual Behavioral Science, 3(1), 21–26. 10.1016/j.jcbs.2013.11.003 [PubMed: 25485230]

Lillis J, & Hayes SC (2008). Measuring avoidance and inflexibility in weight related problems. International Journal of Behavioral Consultation and Therapy, 4(4), 348–354. 10.1037/h0100865

Ma L, Huang C, Tao R, Cui Z, & Schluter P (2021). Meta-analytic review of online guided self help interventions for depressive symptoms among college students. Internet Interventions, 25, 100427. 10.1016/j.invent.2021.100427 [PubMed: 34401386]

McGuinness LA, & Higgins JPT (2021). Risk-of-bias visualization (robvis): An R package and shiny web app for visualizing risk-of-bias assessments. Research Synthesis Methods, 12(1), 55–61. 10.1002/jrsm.1411 [PubMed: 32336025]

Muto T, Hayes SC, & Jeffcoat T (2011). The effectiveness of acceptance and commitment therapy bibliotherapy for enhancing the psychological health of Japanese college students living abroad. Behavior Therapy, 42(2), 323–335. 10.1016/j.beth.2010.08.009 [PubMed: 21496516]

National Student Clearinghouse Research Center (2023, February 2). Current term enrollment estimates. National Student Clearinghouse Research Center. Retrieved February 7, 2023, from https://nscresearchcenter.org/current-term-enrollment-estimates/

Ouzzani M, Hammady H, Fedorowicz Z, & Elmagarmid A (2016). Rayyan—A web and mobile app for systematic reviews. Systematic Reviews, 5(1), 210. 10.1186/s13643-016-0384-4 [PubMed: 27919275]

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, … Moher D (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, n71. 10.1136/bmj.n71 [PubMed: 33782057]

Pelletier SG (2010). Success for adult students. Public Purpose, 12, 2–6.

Pedrelli P, Nyer M, Yeung A, Zulauf C, & Wilens T (2015). College students: Mental health problems and treatment considerations. Academic Psychiatry, 39(5), 503–511. 10.1007/s40596-014-0205-9 [PubMed: 25142250]

Petruzzello SJ, & Box AG (2020). The kids are alright—right? Physical activity and mental health in college students. Kinesiology Review, 9(4), 279–286. 10.1123/kr.2020-0039

Prudenzi A, Graham CD, Clancy F, Hill D, O'Driscoll R, Day F, & O'Connor DB (2021). Group-based acceptance and commitment therapy interventions for improving general distress and work-related distress in healthcare professionals: A systematic review and meta-analysis. Journal of Affective Disorders, 295, 192–202. 10.1016/j.jad.2021.07.084 [PubMed: 34479127]

Radford AW, Cominole M, & Skomsvold P (2015). Demographic and Enrollment Characteristics of Nontraditional Undergraduates: 2011-12. Web Tables. NCES 2015-025. National Center for Education Statistics.

Rodgers MA, & Pustejovsky JE (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. Psychological Methods, 26(2), 141–160.

Rolffs JL, Rogge RD, & Wilson KG (2018). Disentangling components of flexibility via the hexaflex model: Development and validation of the multidimensional psychological flexibility inventory (MPFI). Assessment, 25(4), 458–482. 10.1177/1073191116645905 [PubMed: 27152011]

RStudio Team (2022). RStudio: Integrated development environment for R. RStudio, PBC, Boston.

Schaie KW, & Willis SL (2000). A stage theory model of adult cognitive development revisited. In Rubinstein B, Moss M, & Kleban M (Eds.), The many dimensions of aging: Essays in honor of M. Powell Lawton (pp. 175–193). New York: Springer.

Sears DO (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. Journal of Personality and Social Psychology, 51(3), 515–530. 10.1037/0022-3514.51.3.515

Slutske WS, Hunt-Carter EE, Nabors-Oberg RE, Sher KJ, Bucholz KK, Madden PAF, Anokhin A, & Heath AC (2004). Do college students drink more than their non college-attending peers? Evidence from a population-based longitudinal female twin study. Journal of Abnormal Psychology, 113(4), 530–540. 10.1037/0021-843X.113.4.530 [PubMed: 15535786]

Smout M, Davies M, Burns N, & Christie A (2014). Development of the valuing questionnaire (VQ). Journal of Contextual Behavioral Science, 3(3), 164–172. 10.1016/j.jcbs.20l4.06.001

Soltani E, Bahrainian SA, Masjedi Arani A, Farhoudian A, & Gachkar L (2016). Psychometric properties of the Persian version of the social anxiety—acceptance and action questionnaire. Iranian Journal of Psychiatry and Behavioral Sciences, 10(2). 10.17795/ijpbs.3753

Stockton D, Kellett S, Berrios R, Sirois F, Wilkinson N, & Miles G (2019). Identifying the underlying mechanisms of change during acceptance and commitment therapy (ACT): A systematic review of contemporary mediation studies. Behavioural and Cognitive Psychotherapy, 47(3), 332–362. 10.1017/S1352465818000553 [PubMed: 30284528]

Thompson B, Moghaddam N, Evangelou N, Baufeldt A, & das Nair R (2022). Effectiveness of acceptance and commitment therapy for improving quality of life and mood in individuals with multiple sclerosis: A systematic review and meta-analysis. Multiple Sclerosis and Related Disorders, 63, 103862. 10.1016/j.msard.2022.103862 [PubMed: 35636265]

Thompson EM, Destree L, Albertella L, & Fontenelle LF (2021). Internet-based acceptance and commitment therapy: A transdiagnostic systematic review and meta-analysis for mental health outcomes. Behavior Therapy, 52(2), 492–507. 10.1016/j.beth.2020.07.002 [PubMed: 33622516]

Trindade IA, Guiomar R, Carvalho SA, Duarte J, Lapa T, Menezes P, Nogueira MR, Patrão B, Pinto-Gouveia J, & Castilho P (2021). Efficacy of online-based acceptance and commitment therapy for chronic pain: A systematic review and meta-analysis. The Journal of Pain, 22(11), 1328–1342. 10.1016/j.jpain.2021.04.003 [PubMed: 33892153]

Van den Noortgate W, López-López JA, Marín-Martínez F, & Sánchez-Meca J (2013). Three level meta-analysis of dependent effect sizes. Behavior Research Methods, 45(2), 576–594. 10.3758/s13428-012-0261-6 [PubMed: 23055166]

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

Viergever RF, & Ghersi D (2011). The quality of registration of clinical trials. PLoS ONE, 6(2), e14701. 10.1371/journal.pone.0014701 [PubMed: 21383991]

Wilson KG, Sandoz EK, Kitchens J, & Roberts M (2010). The valued living questionnaire: Defining and measuring valued action within a behavioral framework. The Psychological Record, 60(2), 249–272. 10.1007/BF03395706

Yadavaia JE, Hayes SC, & Vilardaga R (2014). Using acceptance and commitment therapy to increase self-compassion: A randomized controlled trial. Journal of Contextual Behavioral Science, 3(4), 248–257. 10.1016/j.jcbs.2014.09.002 [PubMed: 25506545]

Zhao C, Lai L, Zhang L, Cai Z, Ren Z, Shi C, Luo W, & Yan Y (2021). The effects of acceptance and commitment therapy on the psychological and physical outcomes among cancer patients: A meta-analysis with trial sequential analysis. Journal of Psychosomatic Research, 140, 110304. 10.1016/j.jpsychores.2020.110304 [PubMed: 33248396]

Zhao C, Ren Z, Jiang G, & Zhang L (2022). Mechanisms of change in an Internet-Based ACT study for depression in China. Journal of Contextual Behavioral Science, 24, 51–59.

Highlights

- We examined the effect of ACT on undergraduates' psychological flexibility and inflexibility.

- A three-level meta-analysis indicated that ACT had a small-to-medium effect on PF/PIF.

- Clinical indication and participant sex were significant moderators.

- Measurement issues limited knowledge regarding ACT's effect on PF/PIF components.

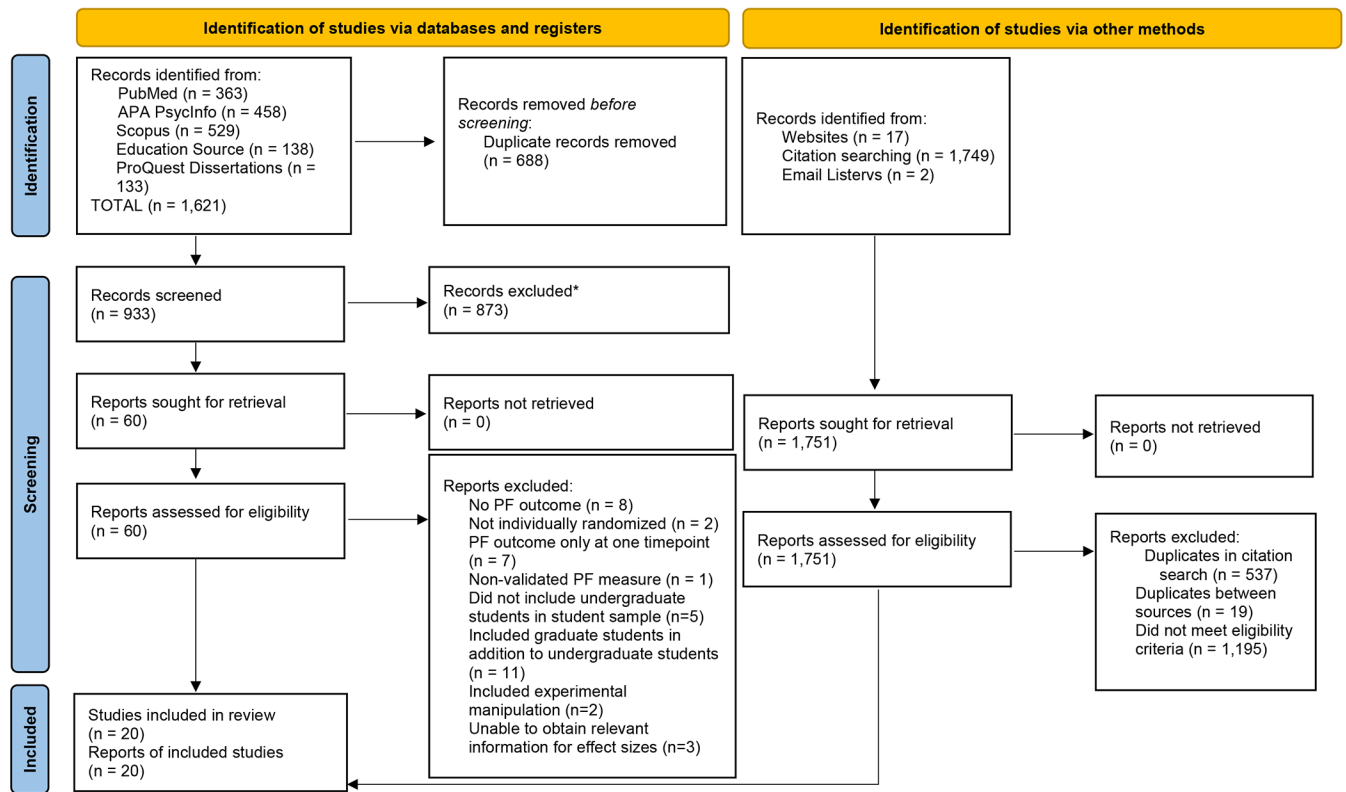- Results may aid efforts to refine the operationalization of PF/PIF.

Figure 1. Prisma 2020 Flow Diagram
*No automation tools were used during this process

Table 1.

Moderators of the effectiveness of ACT Interventions for Undergraduate Students

Moderator Variables	k	#ES	B_0/g	t_0	B_1	t_1	F(df1, df2)	p-value
Intervention Characteristics								
Treatment Focus							F(1, 54) = 4.09	.048
Non-Clinical (RC)	20	56	0.20	1.65				
Clinical	20	56	0.53	4.71[**]	−.33	−2.02[*]		
Control Group Type							F(1, 54) = 2.76	.10
Waitlist/No Treatment (RC)	20	56	0.48	4.66[**]				
Active Control	20	56	0.25	2.25[**]	.22	1.66		
Intervention Size							F(1, 54) = 2.23	.14
Individual (RC)	20	56	0.29	2.80[**]				
Group	20	56	0.57	3.66[**]	−.28	−1.49		
Delivery Modality							F(2, 53) = 1.11	.34
In-Person (RC)	20	56	0.50	4.00[**]				
Online	20	56	0.30	1.95	−.20	−.99		
Bibliotherapy	20	56	0.16	0.72	−.34	−1.36		
Number of Sessions							F(1, 54) = 2.42	.13
Single (RC)	20	56	0.04	0.16				
Multiple	20	56	0.43	4.26[**]	.39	1.55		
Outcome Measure							F(1, 54) = 3.02	.08
AAQ-II and Derivatives (RC)	20	56	0.33	3.51[**]				
Not AAQ-II and Derivatives	20	56	0.46	4.44[**]	.12	1.74		
Outcome Assessed							F(1, 54) = 1.08	.30
Psychological Flexibility (RC)	20	56	0.44	4.10[**]				
Psychological Inflexibility	20	56	0.36	3.99[**]	−.08	−1.04		
Timepoint							F(1, 48) = 2.61	.09
Post-Intervention	20	56	0.41	4.57[**]				
Follow-up	20	56	0.29	2.88[*]	−.12	−1.73		
Length of Follow-Up (in months)	19	9	0.28	2.11[*]	.05	.96	F(1, 17) = 0.93	.35
Participant Characteristics								
Participant Age	20	56	0.37	3.85[**]	−.01	−.41	F(1, 54) = 0.16	.69
Percentage Female	20	56	0.39	5.06[**]	−.01	−2.48[*]	F(1, 54) = 6.15	.02
Participant Race								
% White	13	39	0.33	2.73[**]	−.01	−1.07	F(1, 37) = 1.15	.29
% Black or African American	13	39	0.25	2.50[*]	.01	1.49	F(1, 37) = 2.22	.14
% Native American or Alaskan Native	13	39	0.27	2.35[*]	.002	0.05	F(1, 37) = .003	.96
% Asian	13	39	0.23	0.60	−.003	0.90	F(1, 37) = 0.02	.90
% Native Hawaiian or Pacific Islander	13	39	0.29	2.48[*]	−.06	−.52	F(1, 37) = 0.27	.61

Moderator Variables	k	#ES	B_0/g	t_0	B_1	t_1	F(df1, df2)	p-value
% Multiracial or Other	13	39	0.31	2.98[**]	−.05	−1.52	$F(1, 37) = 2.30$	.14

Note.

k = number of independent studies; #ES = number of effect sizes; B_0/g = intercept or mean effect size; t_0 = test of whether the mean effect size significantly deviates from zero; B_1 = estimated unstandardized regression coefficient; t_1 = test of difference in mean effect size with reference category; F(df1, df2) omnibus test of moderation; RC = Reference category.

For categorical moderators, each intercept represents the mean effect of a category, and each slope represents the difference in the mean effect between the category and reference category. Therefore, the slope of the reference category is not reported because it represents a comparison of the reference category with itself. The slope of a continuous moderator represents an increase or decrease in effect size with each unit increase in the variable.

[*]
$p < .05$;

[**]
$p < .01$.