EMPIRICAL RESEARCH

Non-Specific versus Targeted Approaches to Youth Mentoring: A Follow-up Meta-analysis

Kirsten M. Christensen[1] · Matthew A. Hagler[1] · Geert-Jan Stams[2] · Elizabeth B. Raposa[3] · Samantha Burton[1] · Jean E. Rhodes[1]

Abstract
Despite decades of increased research and funding, youth mentoring programs, overall, yield small effects on youth outcomes. As a result, there are growing calls for programs to utilize the mentoring relationship as context for intentional, targeted skills development, in which mentors employ targeted skills designed to match the presenting concerns of mentees. This targeted approach contrasts with the historically dominant, non-specific friendship model, which holds that a supportive relational bond—alone—promotes positive developmental change. The current study is a follow-up meta-analysis using a comprehensive dataset of all intergenerational, one-on-one mentoring program evaluations published between 1975 and 2018, investigating the comparative impact of targeted, skills-based versus non-specific, relational approaches to mentoring. Analyses of 48 mentoring studies of youth outcomes (average youth age of 12.25 years old) revealed the overall effect size of targeted programs to be more than double that of non-specific relational approaches, with significant moderator effects on academic, psychological, and social functioning. Findings suggest that youth mentoring programs can promote positive outcomes, particularly when mentors employ targeted approaches matched to the needs of their mentees.

Keywords Meta-analysis · Youth mentoring · Formal mentoring · Program evaluations · Skills-based · Relational

Introduction

Meta-analyses of the effectiveness of youth mentoring programs in improving youth outcomes have shown a remarkably consistent and relatively modest range of overall effects over time, despite steady investments in mentoring-related research designed to improve program impact (Fernandes-Alcantara 2018). DuBois and colleagues have published two comprehensive meta-analyses of youth mentoring programs, both of which showed relatively small effects of mentoring across outcomes (Hedge's g values ranging from 0.18 to 0.21; DuBois et al. 2002, 2011). Likewise, a 2019 meta-analysis of 70 intergenerational, one-on-one mentoring program evaluations once again showed that mentoring yielded a range of significant, but small overall effects (Hedges' $g = 0.21$; Raposa et al. 2019). In response to the lack of improvement in the effect sizes of youth mentoring over the past two decades, the authors called for "more rigorous adherence to evidence-based practices that target specific mechanisms underlying particular youth difficulties, rather than relying on a relatively low-intensity, non-specific approach with uneven adherence to practices that are research-informed" (Raposa et al. 2019, p. 438).

Nonetheless, the vast majority of mentoring programs provide non-targeted care, encouraging mentors to provide general friendship, support, and role modeling aimed at broad developmental goals (Garringer et al. 2017). This approach is based on the assumption that the mentor–youth relationship itself is the primary active ingredient of change. In particular, a close, supportive relationship with an adult is thought to provide youth with a "corrective experience" which, in turn is thought to lead to improvements in youth functioning across a broad range of developmental domains (Rhodes 2005). In fact, according to proponents of this non-specific, relationship-focused model of mentoring, targeted, skills-based approaches may actually be counter-indicated

✉ Jean E. Rhodes
    Jean.Rhodes@umb.edu

1 University of Massachusetts Boston, 100 William T. Morrissey
   Blvd, Boston, MA 02125, USA

2 University of Amsterdam, 1012 WX Amsterdam, Netherlands

3 Fordham University, 441 East Fordham Road, Bronx, NY 10458,
   USA

insofar as they may hamper relationship-building (Li and Julian 2012). Yet there is emerging evidence that more targeted, problem-specific approaches to mentoring may yield larger effects (e.g., Lyons et al. 2019). Programs that have taken a structured approach often do so in response to the fact that many youth who are referred to mentoring programs present with significant emotional, behavioral, or academic difficulties (Jarjoura et al. 2018). The current meta-analysis, which builds on the Raposa et al. (2019) study, contributes to this growing debate by examining whether non-specific approaches to youth mentoring show differential effects relative to more targeted, problem-focused approaches to mentoring.

The non-specific, relationship-focused approach to youth mentoring has been the dominant model of youth mentoring in the U.S. since mentoring programs first emerged in the early 1900s. In a recent study of thirty, nationally representative youth mentoring programs, nearly 2000 mentors were asked how they spent time with their mentees (Jarjoura et al. 2018). The responses reflected a non-specific approach to broad developmental goals. The most common response, "making time to have fun," was followed by activities such as discussing important people or personal issues, going to cultural or other special events, and engaging in creative activities (Jarjoura et al. 2018). Likewise, in a national survey of 1271 mentoring organizations affiliated with MENTOR: The National Mentoring Partnership (Garringer et al. 2017), the most common program goals, according to program directors, were the broad development of life skills (53.9%), general youth development (51.2%), and providing a caring adult relationship (44.6%). In contrast, comparatively fewer programs endorsed more specific and instrumental goals, such as supporting college access (17.9%), violence prevention (5.2%), STEM education (6.3%), or substance use prevention (2.9%). Finally, in an evaluation of Big Brothers Big Sisters (BBBS) school-based mentoring programs, only 11% of mentors endorsed having an instrumental goal (e.g., improving academic or school behavior), while 79% endorsed broad developmental goals (e.g., being a friend or helping the child feel good about him/herself; Raposa et al. 2016).

Researchers and theorists have long advocated this non-specific, relationship-focused approach as the most effective way to connect with one's mentee (Rhodes 2005). In an influential article, researchers Li and Julian (2012) compared the role of the mentor-mentee bond to that of fluoride in toothpaste, suggesting that other ingredients (e.g., color, taste) may add value, but they are not essential to the success of the match. These authors argued that the "enduring emotional attachment" is the only "active ingredient" in mentoring programs and that "scaled-up programs and policies serving children and youth often fall short of their potential impact when their designs or implementation drift

toward manipulating other 'inactive' ingredients (e.g., incentive, accountability, curricula) instead of directly promoting developmental relationships" (p. 157). Other researchers have reached similar conclusions, noting that "the common feature of successful interventions across all stages of the life cycle through adulthood is that they promote attachment and provide a secure base for exploration and learning for the child" (Heckman and Kautz 2013, p. 2). This type of bond is thought to compensate for a lack of the close bonds of a nuclear family and set youth on a more positive life course (Rhodes 2005).

One factor that may also help to explain the persistence of non-specific mentoring approaches relates to the wide range of presenting problems endorsed by the families of youth referred to large programs like BBBS. Many of these larger programs work with youth at diverse developmental stages (i.e., early childhood through late adolescence), from diverse backgrounds (e.g., foster care, unaccompanied refugees, previously incarcerated youth), who are often dealing with a wide variety of problems (e.g., depression and anxiety, peer rejection, substance use). For these programs, a generic, friendship-based model is a scalable approach that can be delivered to all youth, irrespective of their particular issues. Such programs rarely have access to the wide range of targeted, empirically supported interventions that would be needed to address their mentees' needs. Moreover, when they do specialize, they can only hit the mark with a subset of mentees. It is thus completely understandable that non-specific programs have defaulted to the common denominator, lighter-touch friendship models that can essentially be delivered to all youth, irrespective of their particular issues.

Yet, non-specific, relationship-based approaches alone may not adequately address the substantial emotional, behavioral, or academic difficulties that mentees face. Compared to national samples, youth who are referred to mentoring programs are often at significantly higher risk for a variety of difficult life circumstances (e.g., poverty, parental substance abuse), as well as behavioral and mental health issues like depression, anxiety, aggression, and attention difficulties (Jarjoura et al. 2018). Moreover, recent evidence suggests that families from marginalized communities often perceive mentoring programs as an alternative to professional healthcare services. For example, one study found that Black caregivers were twice as likely as White caregivers to turn to mentoring programs to address their children's externalizing behaviors such as aggression, hyperactivity, or conduct problems (Vázquez and Villodas 2018). Given the diverse barriers to accessing high-quality professional mental health services within marginalized communities, including language differences, discrimination, and financial costs, mentoring is sometimes seen as a less stigmatizing and more culturally congruent approach to

supporting youth with mental health needs (Cook et al. 2013; Vázquez and Villodas 2018). In general, prevention programs that work with youth who experience more acute behavioral, emotional, and academic problems tend to show stronger positive effects because, like a pendulum, there is more room for an upswing (Tanner-Smith et al. 2018). Despite this, non-specific mentoring programs have tended to yield relatively weaker effects than other interventions that serve youth at comparable levels of risk. It may be the case that, in the absence of specialized training and clear goals, many mentors in programs with a non-specific focus feel overwhelmed by their mentees' difficulties (Spencer 2007). In fact, higher rates of mentee stress exposure at home and behavioral difficulties (e.g., delinquency, academic problems) are often key predictors of poor mentor–youth relationship quality (Raposa et al. 2016). Moreover, this mismatch between parent and youth needs and mentor training can lead to high proportions of early match closures and youth disappointment, particularly for youth with more severe behavioral and emotional difficulties (Bernstein et al. 2009; Kupersmidt et al. 2017).

Thus, although efforts to develop strong mentor–youth relationships are certainly necessary, it may be equally important to provide mentors with training in targeted approaches that can more directly address the needs of mentees. In fact, several recent evaluations of more targeted approaches to youth mentoring suggest their effectiveness. For example, Sowers et al. (2017) evaluated the efficacy of a STEM-focused mentoring program for students with disabilities. Mentees were matched with a mentor in the STEM field, and mentoring activities and discussions specifically revolved around education and career development in STEM. Findings revealed moderate to large effects for a range of outcomes such as STEM-related knowledge, engagement, confidence, and career planning confidence. Other research suggests that, when mentoring activities are calibrated and targeted to specific challenges, youth see even stronger positive effects in academic (Wellington-Johnson 1997), psychological (Jent and Niec 2006), and social (King et al. 2002) outcomes.

In contrasting targeted and non-specific approaches to youth mentoring, it is helpful to look to an analogous, and perhaps more developed, discussion in psychotherapy. For decades, researchers and clinicians have debated the "active ingredient" of psychotherapy, with one side arguing for that the relationship between therapist and client as the primary vehicle of change, and the other arguing for the importance of applying evidence-based interventions that have demonstrated empirical evidence of effectiveness. This debate has been a topic of extensive research comparing "usual care" approaches, in which therapists "use their clinical judgment as they [see] fit, not constrained by evidence-based interventions or manuals," to evidence-based interventions, in which therapists apply empirically support therapy protocols (Weisz et al. 2005, p. 631). If the therapeutic relationship is the primary vehicle of change, one would expect "usual care" models to be as effective, if not more effective, compared to evidence-based approaches.

Yet, decades of studies have consistently shown that evidence-based practice shows superior effects compared to usual care. In a multilevel meta-analysis of 52 randomized controlled trials (RCT's), Weisz et al. (2013) found that evidence-based psychotherapy was significantly more effective than usual care, with an average difference in effect size of 0.29. The same research group replicated this finding in a more recent meta-analysis, finding an average difference in effect size of 0.30 between evidence-based practice and usual care (Weisz et al. 2017). The authors note that usual care seemed to be more rigorous and effective than other types of control conditions, such as waitlists, psychotherapy placebos, and case management, and that usual care likely has some active treatment ingredients, such as a therapeutic relationship and consistent contact. However, the consistent difference in effect size indicates that evidence-based interventions add value and effectiveness to treatment, and that techniques that are empirically and theoretically matched to presenting problems are also essential active ingredients. Applied to youth mentoring, these findings suggest that providing mentors with training and supervision in techniques that target youth's presenting difficulties, rather than relying on intuition and relationships alone, may enhance effectiveness. In practice, this would involve specifically targeting youth with particular presenting problems and/or assessing and identifying difficulties at the start of the program, and then applying interventions that theoretically, and (if known) empirically matched to these problems that can be delivered in the context of the mentoring relationship. As discussed above, although non-specific approaches have been far more common, some targeted mentoring programs have been developed and evaluated, presenting the possibility to conduct a meta-analytic comparison between targeted and non-specific approaches.

The Current Study

Although no studies to date have directly compared these two models of mentoring, the relative benefits of targeted and non-specific models can begin to be explored within the context of meta-analysis. The current study builds on a recent meta-analysis of 70 intergenerational, one-on-one mentoring program evaluations, representing more than 25,000 youth, that were conducted from 1975 through 2017 (Raposa et al. 2019). One potentially important moderator that this meta-

analysis did not code or analyze involves whether the program took a non-specific, relationship-focused approach versus a more targeted, problem-specific approach. To address this gap in the literature, the current meta-analysis sought to examine the relative impact of these two distinct approaches to formal mentoring. The present study built upon the analyses of Raposa et al. (2019) in two ways: (1) an additional literature search was conducted to find articles published since the completion of this previous meta-analysis (up through September 2018); and (2) a rigorous coding procedure was applied to these studies in order to specify whether the evaluated program was non-specific and purely relationship-focused versus targeted and problem-specific.

Using rigorous inclusion criteria established by Raposa et al. (2019), analyses included all relevant outcome studies of intergenerational, one-on-one formal youth mentoring programs (i.e., offering the support of a caring relationship with a non-parental adult) written in English between 1975 and 2018. A multilevel meta-analytic approach was used to estimate the overall effect size of youth mentoring programs and assess whether the magnitude of these effects was moderated by the approach to mentoring (i.e., non-specific model versus targeted approach).

It was hypothesized that the overall effect size would be larger for targeted mentoring programs compared to programs taking a non-specific approach. Moreover, in light of the research regarding stronger effects for programs serving higher-risk youth, youth risk status was examined as a moderator. Finally, analyses explored whether mentoring approach (targeted versus non-specific) would significantly moderate five most common outcome domains in mentoring program evaluations (i.e., academic functioning, psychological symptoms, social functioning, health problems, cognitive functioning).
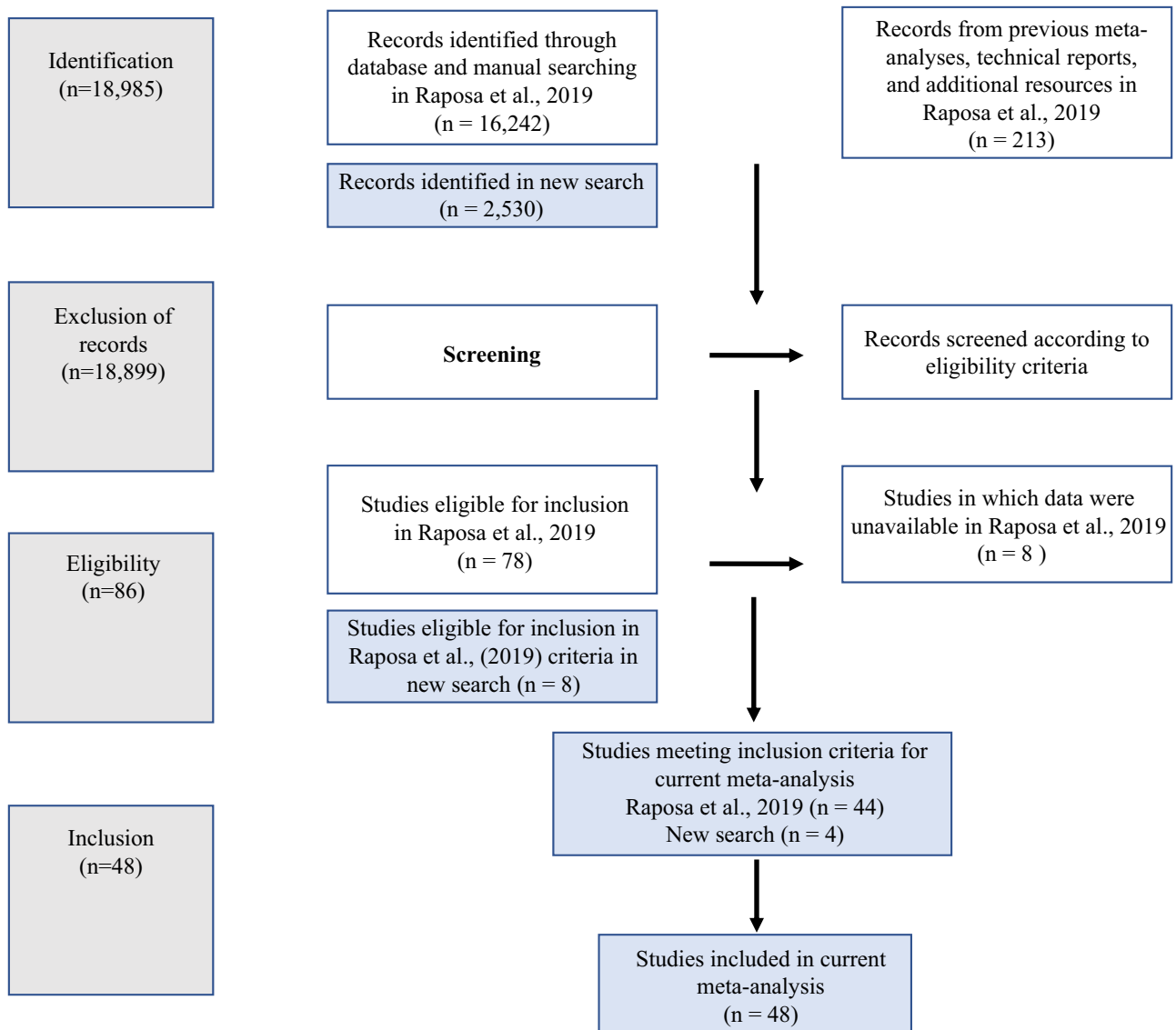
Method

Study Selection

The current analyses drew on an updated set of the original studies used in Raposa et al. (2019). As reported in Raposa et al. (2019), these studies were found through a comprehensive literature search of three online databases (i.e., PsychINFO, ERIC, ProQuest), identifying all evaluations of formal mentoring programs (i.e., peer reviewed articles, unpublished dissertations, technical reports) published prior to September 2017 using the following search terms and combinations of terms: Youth mentoring, Mentor + program, Mentor + evaluation, Mentor + intervention, Mentor + outcomes, Mentor + effects, Mentor + comparison, Big Brother, Big Sister, Protégé + mentor, Apprentice + mentor, Nonparental adult + mentor, Mentor mentee

relationship, Mentor + randomized control trial, Mentor + RCT, and Mentor + experimental. Prior meta-analyses and qualitative reviews were manually searched to identify additional studies. Websites of multiple national formal mentoring programs, agencies that fund or evaluate mentoring research, and the Office of Juvenile Justice and Delinquency Prevention were searched for additional articles and reports. Additional research recommendations were requested from experts in the field of youth mentoring, including those involved with previous mentoring meta-analyses as well as those with ample experience conducting youth mentoring research. This search process resulted in 16,455 potential articles, dissertations, and reports. Duplicate studies were eliminated prior to evaluation for inclusion and the most comprehensive studies were utilized in cases where multiple evaluations included the same sample.

Original inclusion criteria from the Raposa et al. (2019) study were used and included the following: (1) A formal mentoring program with mentoring defined as a non-parental adult or older youth serving in a non-professional helping role for a specific youth to promote positive outcomes through the relationship. More professional relationships such as coaches or tutors did not meet inclusion criteria. (2) An evaluation with a comparison group, such as a randomized control trial and/or quasi-experimental design. Original study exclusion criteria included: (1) same-age peer mentoring, (2) exclusively group mentoring, (3) adult mentees (i.e., mentees older than 18 years), (4) inadequate treatment versus comparison group distinction (e.g., both groups received mentoring or the treatment group included a considerable amount of youth who did not receive mentoring), (5) evaluations in which mentoring was adjunctive, or not the primary or secondary intervention, (6) measured outcomes could not be categorized into one of the following broad categories: psychological, social, academic, health, or cognitive, (7) the mentoring program could not be categorized into either a non-specific or targeted approach, (8) inadequate data to compute an effect size, and the author was unable to be reached to provide necessary information or did not respond to requests for information within a specified timeframe, and (9) the study was written in a language other than English.

To update Raposa et al. (2019) database, an additional literature search was conducted to find articles published between September 2017 and September 2018 using the same databases and search terms. This resulted in a total of 2530 articles being screened, of which eight additional studies were identified according to Raposa et al.'s original criteria. Of these eight new studies, only four contained adequate information to be coded as "targeted" or "non-specific model." Thus, the final analytic sample for the current meta-analysis included 48 studies and 535 effect sizes (see Fig. 1 for an overview of study selection).

Fig. 1 Flowchart of the study selection process

Study Coding Procedures

Procedures for coding youth outcomes were identical to the methods described in detail in Raposa et al. (2019). Of particular relevance for this follow-up analysis, five raters with advanced statistical experience and who were trained by a meta-analysis expert double-coded all selected studies. Raters met weekly with each other, and with meta-analysis and mentoring experts, to resolve ambiguities and discrepancies. Five broad categories of youth outcomes were coded in each study, including relevant statistical information (e.g., sample size, means, standard deviations, frequencies) to calculate an effect size for each outcome. These categories, established *a priori* to align with theories of

developmental psychopathology and positive youth development, included academic functioning, psychological symptoms, social functioning, health problems, and cognitive functioning. The original meta-analysis also coded for several mentee characteristics. In addition to being coded for youth age, gender, and race/ethnicity, each study was coded for the type of youth population for which the mentoring intervention was designed. Studies were coded as to whether the program was designed to serve a "general" youth population or targeted one of five "at-risk" sample types: racial or ethnic minority youth, youth with single parents, youth from low socioeconomic status families, youth in foster care, or youth with multiple risk factors such as these. Given that Raposa et al. (2019) did not find a

moderator effect for these five different risk groups, they were collapsed into one "at-risk" category in the current analyses.

Coding of Non-Specific versus Targeted Approaches

For this study, an additional coding procedure was implemented to identify whether each study involved a program that took a non-specific approach versus a targeted, problem-specific approach to mentoring (see Table 1). The senior author, who has over 30 years of experience in the mentoring research field, as well as practical experience as a match support specialist in a mentoring program, developed the initial coding scheme in consultation with other mentoring experts and based on the body of the research reviewed above. Studies were coded as tests of the non-specific model if they involved a program in which mentors were trained to act as a caring friend, primarily engaging in non-specific recreational activities with their mentee, rather than targeting a particular youth challenge or need. For example, a study of a Swedish mentoring program aiming to prevent substance use in youth had the program objective of "[establishing] a safe and supportive relationship, by which the youth is assumed to benefit in social, emotional, and academic development, and as a consequence, be less prone to use alcohol and drugs" (Bodin and Liefman 2011, p. 441).

In contrast, articles were coded as testing a targeted approach if the mentoring intervention both (a) targeted a specific youth population (e.g., trauma exposed youth) or challenge (e.g., depression, academic difficulties) *and* (b) implemented a mentoring intervention specifically designed to match the needs of the target population or challenge. Although it was not necessary for programs to use a manualized curriculum to be considered targeted, one example of this type of program included the Achievement Mentoring Program (Holt 2008), a manualized cognitive-behavioral mentoring program that seeks to enhance school-related cognitions and behaviors in students at risk of not graduating from high school.

Rigorous coding standards were implemented to reduce ambiguity and achieve clarity and precision about the nature of the mentoring program. For example, articles with the following features were excluded from analyses: (1) articles that did not provide sufficient information about the mentoring program; (2) non-specific mentoring programs that occasionally exposed their mentees to group-based, more targeted workshops; and (3) mentoring programs that aimed to target specific problems, but were "mismatched" in terms of mentee challenges and actual mentoring activities (e.g., targeting youth with academic challenges but the mentor is a "friend" who engages in recreational, rather than educational activities).

Each article was coded independently (i.e., double-coded) by two graduate research assistants who were involved in the development of the coding scheme and received training from the senior author regarding the coding scheme. These graduate research assistants then convened to discuss discrepancies until reaching agreement. In the event that an agreement was not reached, a consensus meeting was held with both graduate research assistants and the senior author to review the article and determine the final code. Inter-rater reliability tests demonstrated that the overall agreement across these two coders was 90%, representing almost perfect agreement ($\kappa = 0.85$; Cohen, 1960, as cited in McHugh 2012).

Effect Size Calculation and Data Analysis

Effect sizes were calculated by obtaining the standardized mean difference between the mentoring intervention group and the control group for each outcome, where a positive standardized mean value indicates higher performance for the mentoring intervention groups. To adjust for small sample sizes as well as differences in sample size across studies, standardized mean difference values were then transformed into Hedge's g (Hedges and Olkin 1985). Study authors were contacted for more information in cases where necessary data (i.e., sample sizes, means, standard deviations, other values) was not reported.

Given that many studies assessed more than one youth outcome, multiple effect sizes were calculated for a majority of the included studies. To account for interdependency of these effect sizes, a three-level meta-analysis approach was implemented (Van den Noortgate et al. 2014). Three-level meta-analysis approaches are more rigorous than traditional meta-analysis methods as they allow for all effect sizes obtained from the same study (i.e., dependent effect sizes) to be included in analysis. In addition to this major advantage, three-level approaches to meta-analysis account for between- and within-study variability, increase statistical power, and allow for more moderator analyses than traditional approaches.

Three-level meta-analysis allows for modeling of three sources of variance, including (1) the sampling variance of the observed effect sizes, (2) the variance between effect sizes obtained from the same study, and (3) the variance between studies. The first level of variance (sampling variance of observed effect sizes) was estimated using a previously established formula (Cheung 2014). To determine whether there was significant variance present at the second (within-study) and/or third (between-study) levels, log-likelihood-ratio-tests compared the full-model deviance relative to the deviance of models excluding one of the variance parameters (Assink and Wibbelink 2016). Significant variance at level 2 or level 3 suggests there is

Table 1 Article codes of mentoring approach

Author	Year	Code Non-specific (1) Targeted (2) Excluded (X)
Fo & O'Donnell	1975	X
Dicken, Bryson, & Kass	1977	1
Flaherty	1985	X
Davis	1988	X
Galvin	1989	1
Rowland	1991	1
McPartland & Murray Nettles	1991	X
Slicker & Palmer	1993	1
Newton	1994	X
Jones	1994	2
Tomlin	1994	2
Blakely, Menon, Jones, Pruitt, Ash, & Johnson	1995	2
Turner & Scherman	1996	1
Wellington Johnson	1997	2
Abbott, Meredith, Self-Kelly & Davis	1997	1
Grossman & Tierney	1998	1
Lee & Crammond	1999	1
Houston	1999	X
Gordon	2000	X
Blechman, Maurice, Buecker, & Helberg	2000	X
Aseltine, Dupre, & Lamlein	2000	X
Lee	2001	2
Barron-McKeagney, Woody, & D'Souza	2001	1
Karcher, Davis, & Powell	2002	2
Starks	2002	X
King, Vidourek, Davis, & McClellan	2002	2
Keating, Tomishima, Foster, & Alessandri	2002	X
Schobitz	2003	X
Holloway	2004	X
Saintonage, Achille, & Lachance	2005	1
Portwood, Ayers, Kinnison, Waris, & Wise	2005	X
Karcher	2005	2
Jent & Niec	2006	2
Anderson	2006	X
Bauldry	2006	X
Drake & Bernoski	2006	2
Moore	2006	X
Schmidt, McVaugh, & Jacobi	2007	X
Jarjoura	2007	2
Holt	2007	2

Table 1 (continued)

Author	Year	Code Non-specific (1) Targeted (2) Excluded (X)
De Wit et al.	2007	1
DeSocio et al.	2007	2
Herrera et al.	2008	1
Karcher	2008	1
De Blank	2009	X
Brown et al.	2009	X
Johnson	2009	X
Bernstein et al.	2009	X
Converse & Lignugaris	2009	X
Clarke	2009	2
Child, Rossler, & Nichol	2009	X
Ellege, Cavell, Ogle, & Newgent	2010	1
Taussig & Culhane	2010	2
Gibbs-Roseboro	2010	X
ICF International (OJJDP Amachi Texas)	2011	1
Dolan et al.	2011	1
Herrera, Grossman, Kauh, & McMaken	2011	1
Bodin & Leifman	2011	1
Kuperminc et al.	2011	2
McQuillin, Smith, & Strait	2011	2
Komosa-Hawkins	2012	2
McQuillin	2012	2
Herrera, DuBois, & Grossman	2013	1
Simoes & Alarcao	2013	2
Cantu	2013	X
Henneberger et al.	2013	2
Morris	2014	1
Sue	2014	1
Gray	2015	X
McQuillin	2015	2
Eddy et al.	2015	1
Weiler et al.	2015	1
Iver et al.	2017	X
Markle et al.	2017	X
Rodriguez-Planas	2017	X
Sowers et al.	2017	2
Lau et al.	2017	X
Angus & Hughes	2017	X
Heppen et al.	2017	2
Sharpe et al.	2017	2

heterogeneity in effect size distribution, indicating that the effect sizes cannot be considered estimates of a common effect size. In such cases, moderator analyses assessing variables such as sample characteristics, program

Table 2 Descriptives non-specific versus targeted mentoring approaches

	Type of mentoring					
	Non-specific		Targeted		Total	
	k	M (SD)	k	M (SD)	k	M (SD)
Sample size	23	371.70 (497.32)	25	170.44 (167.02)	48	266.88 (374.62)
Age	21	11.70 (2.35)	23	12.76 (2.07)	44	12.25 (2.25)
Percent male	20	58.61 (19.38)	22	51.33 (23.53)	42	54.80 (21.71)
Ethnicity						
Percent White[a]	14	45.54 (33.26)	19	22.01 (25.77)	33	32.00 (31.02)
Percent Black[b]	13	22.39 (26.54)	20	50.89 (29.46)	33	39.66 (31.30)
Percent Hispanic	12	24.17 (23.70)	18	23.38 (26.95)	30	23.70 (25.28)
Percent Asian	10	1.10 (1.85)	16	7.60 (24.76)	26	5.10 (19.49)
Percent Hawaiian	9	0.11 (0.33)	12	0.00 (0.00)	21	0.05 (0.22)
Percent Indian	9	0.83 (1.70)	16	0.80 (2.31)	25	0.81 (2.07)
Percent at risk	23	82.61 (38.76)	25	72.00 (45.83)	48	77.08 (42.47)

Overrepresentation of White youth in non-specific mentoring programs, and an overrepresentation of African American/Black youth in targeted mentoring programs. t-test White: $t(31) = 2.292$, $p = 0.029$; t-test Black: $t(31) = -2.821$, $p = 0.008$

k number of studies, M mean, SD standard deviation

[a]Difference at $p < 0.05$

[b]Difference at $p < 0.01$

Table 3 Overall effect of mentoring programs on youth outcomes

Outcome	k	#ES	Mean g	95% CI	t	$\sigma^2_{level\ 2}$	$\sigma^2_{level\ 3}$	% Var. Level 1	% Var. Level 2	% Var. Level 3
Youth outcomes	48	535	0.19	0.13, 0.24	6.410***	0.020***	0.031***	22.01	30.30	47.69

Youth outcomes academic and vocational, social-emotional, physical health, psychosocial problems, k number of studies, *#ES* number of effect sizes, *mean ES* mean effect size (g), *CI* confidence interval, $\sigma^2_{level\ 2}$ variance between effect sizes extracted from the same study, $\sigma^2_{level\ 3}$ variance between studies, *% Var.* percentage of variance distributed

*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$

characteristics, and/or study methodology characteristics can help explain heterogeneity in effect sizes.

All analyses were conducted in R (version 3.2.0) using the *metafor* package. Given the diversity in study samples, settings, interventions, and outcomes, a multilevel random effects model was estimated (Assink and Wibbelink 2016). All model parameters were estimated using the restricted maximum likelihood estimate. The Knapp and Hartung (2003) method was utilized to test individual regression coefficients of the meta-analytic models and to calculate corresponding confidence intervals (Assink and Wibbelink 2016). Dichotomous dummy variables were created for categorical variables (Tabachnik and Fidell 2013). Program type (targeted versus non-specific approach) was examined as a moderator for overall effect size of youth mentoring programs in addition to effect sizes for each outcome domain. Youth risk status was examined as an additional moderator of overall effect size to determine whether programs taking targeted versus non-specific mentoring approaches have different effects in at-risk and general

youth populations. In order to increase statistical power to detect a significant moderator effect, and to prevent chance capitalization due to multiple testing, we collapsed five types of risk into one "at-risk" group.

Results

Youth sociodemographic information is presented in Table 2. Table 3 shows that the average effect size across all 48 studies and 535 effect sizes was $g = 0.19$ ($p < 0.001$; 95% CI: 0.13–0.24), which is a small effect and not significantly different from the average effect size found by Raposa and colleagues ($g = 0.21$; 95% CI: 0.14–0.28; 2019), given the overlapping 95% confidence intervals of the different meta-analyses. There was significant heterogeneity within studies (σ^2 level 2 = 0.02, $p < 0.001$) and between studies (σ^2 level 3 = 0.03, $p < 0.001$). A total of 30% of the variance among effect sizes was accounted for by the within-study level, and 48% by the between-study

Table 4 Targeted versus non-specific approach to mentoring

Moderator variable	k	#ES	B_0/g	t_0	B_1	t_1	$F(\mathrm{df_1, df_2})$
Approach							$F(1, 533) = 6.447^*$
Targeted (RC)	25	228	0.25	6.576[***]			
Non-specific	23	307	0.11	2.902[**]	–0.14	–2.539[*]	
Risk status							$F(1, 533) = 0.430$
No risk (RC)	11	107	0.15	2.492[*]			
At risk	37	428	0.20	5.902[***]	0.05	0.656	
Targeted * Risk							$F(3, 531) = 2.406^+$
Non-specific + No risk (RC)	4	59	0.07	0.812			
Non-specific + At risk	19	248	0.12	2.801[**]	0.05	0.502	
Targeted + No risk	7	48	0.20	2.678[**]	0.13	1.100	
Targeted + At risk	18	180	0.27	5.993[***]	0.20	1.973[*]	
Outcome Sub-Categories							
Academic							$F(1, 180) = 5.040^*$
Targeted (RC)	21	112	0.28	5.142[***]			
Non-specific	15	70	0.09	1.433	–0.19	–2.245[*]	
Psychological symptoms							$F(1, 133) = 5.196^*$
Targeted (RC)	12	47	0.22	4.800[***]			
Non-specific	18	88	0.07	1.750[+]	–0.15	–2.280[*]	
Health							$F(1, 25) = 1.379$
Targeted (RC)	4	9	0.24	2.560[*]			
Non-specific	7	18	0.11	1.850[+]	–0.13	–1.174	
Cognition							$F(1, 61) = 0.108$
Targeted (RC)	13	28	0.11	2.239[*]			
Non-specific	13	35	0.13	3.631[***]	0.02	0.329	
Social functioning							$F(1, 126) = 4.731^*$
Targeted (RC)	15	32	0.24	4.194[***]			
Non-specific	16	96	0.09	2.471[*]	–0.15	–2.175[*]	

k number of independent studies, #ES number of effect sizes, B_0 intercept/mean effect size (g), t_0 difference in mean g with zero, B_1 estimated regression coefficient, t_1 difference in mean g with reference category, F (df_1, df_2) omnibus test, (RC) reference category

$^+p < 0.10$; $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

level, while random sampling error accounted for 22% of the variance.

Table 4 presents the results of the moderator analyses. The omnibus F-test for targeted versus non-specific programs proved to be significant ($p = 0.011$); the overall effect size for targeted, problem-focused programs ($g = 0.25$) was more than twice as large as that of non-specific programs ($g = 0.11$). In terms of clinical significance (Kraemer and Kupfer 2006), these statistics indicate that targeted, problem-specific programs, on average, tended to yield a 14% improvement in youth outcomes relative to no mentoring, while non-specific mentoring models predicted only a 6% increase in outcomes, relative to no mentoring at all.

Subsequent moderator analyses tested whether programs targeting a particular high-risk population of youth (e.g., youth in foster care, youth from families with low socio-economic status) showed stronger effects when the curriculum was problem-specific, relative to programs treating a more general population of youth. Omnibus F-test results revealed a trend approaching statistical significance ($p = 0.067$). In general, the effect of the targeted approach to mentoring was larger in high risk samples ($g = 0.27$) and smaller, or sometimes not statistically different from zero in general population samples ($g = 0.07$).

Finally, analyses of the five outcome sub-domains revealed that the moderator effect for the targeted versus non-specific approach to mentoring was significant for academic functioning, (which includes outcomes such as academic performance and school engagement), psychological symptoms (which includes mental health related outcomes such as self-regulation and internalizing symptoms) and social functioning (which includes outcomes such as social skills and social support), but not for health problems and cognitive functioning.

Discussion

Although previous meta-analyses have examined the impact of mentoring programs on youth development, no studies to date have directly compared the influence of distinct models of mentoring in promoting youth outcomes. The purpose of this study was to compare the effectiveness of "targeted" mentoring programs (i.e., those that explicitly matched specific interventions to mentees' presenting problems) to the effectiveness of "friendship-based" mentoring programs (i.e., those exclusively focused on relationship-building and non-specific recreational activities). The average effect sizes of targeted, problem-focused and non-specific programs were compared. Further, analyses examined whether the respective effect sizes of targeted and non-specific programs differed as a function of outcome domains (i.e., academic, psychological, social, cognitive, physical health) and mentees' risk status. These analyses were conducted by drawing upon the most comprehensive, rigorous meta-analytic dataset on youth mentoring accumulated to date (Raposa et al. 2019), while adding new studies published after this dataset was originally compiled.

The meta-analysis yielded an overall effect size of 0.19, an impact not significantly different from that found in Raposa and colleagues' meta-analysis (0.21) and one that corresponds with a conventionally "small" effect (Cohen 1988). However, when type of program was examined, targeted and more problem-specific programs had an average effect size of 0.25, which was more than double the average effect size non-specific programs ($g = 0.11$). These findings are in line with recent calls from mentoring researchers for stronger alignment with theoretical and evidentiary standards of prevention science (e.g., Cavell and Elledge 2015). Such standards typically require a close association between structured interventions and identified target problems in youth.

Across outcome subcategories (i.e., academic, psychological, cognitive, social, physical health) the effects of targeted programs ranged from 0.11 (cognitive) to 0.28 (academic), while the effects of non-specific programs ranged from 0.07 (psychological) to 0.13 (cognitive). Targeted programs were significantly more effective than non-specific programs in improving academic, psychological, and social functioning, but there was no significant difference in terms of physical health and cognitive functioning. This is perhaps not surprising, given that mentoring interventions are primarily psychosocial in nature. Indeed, most mentoring interventions coded as "targeted" specifically recruited youth with particular academic, psychological, or social difficulties, and equipped mentors with training and treatment approaches that directly addressed these respective issues (e.g., DeSocio et al. 2007; Karcher et al. 2002; King et al. 2002). In contrast, the majority of programs did not tend to target cognitive and physical health outcomes.

Additionally, it is notable that relatively few studies—targeted or non-specific—reported on physical health outcomes at all, perhaps resulting in insufficient statistical power to detect a significant difference between program models in this domain.

Research suggests that significantly elevated rates of mentee stress exposure, as well as pre-existing emotional and behavioral difficulties, might be one reason why mentors struggle to successfully implement non-specific models of mentoring (Raposa et al. 2016; Spencer 2007). To examine the effects of targeted versus non-specific mentoring as a function of mentees' presenting risk, a second set of moderator analyses was conducted. Although applying a targeted approach had a marginally greater benefit relative to non-specific approaches when samples involved higher-risk youth (0.27 versus 0.12) as opposed to a general population of youth (0.20 versus 0.07), results indicated that youth risk status did not statistically significantly moderate the impact of mentoring program approach ($p = 0.067$). However, these analyses used indicators of risk status that were primarily demographic (e.g., racial and ethnic minority status, low SES, single parent household) and, by nature of the meta-analytic approach, were at the program level rather than the participant level. These findings likely do not capture more acute problems often targeted by mentoring models (e.g., academic difficulties or delinquency). Future studies should include a more detailed assessment of youth risk at both individual and contextual levels.

It is important to note that relationship-building activities and training are not antithetical to targeted mentoring approaches. Indeed, many of the programs that were coded as "targeted" provide training to mentors aimed at developing sufficiently strong mentor–mentee bonds as the context and catalyst for the targeted intervention. Even the strongest proponents of more targeted approaches (e.g., Cavell and Elledge 2015) do not eschew the fact that mentoring is, at its core, a relational intervention and that overly prescriptive, rigid approaches could threaten relationship quality and mentees' persistence in the intervention. This tension runs parallel to long-held debates in psychotherapy research about the impact of structured, evidence-based therapies relative to the non-specific or "common factors" of therapist warmth, empathy, and support provided across therapeutic modalities (e.g., Weisz et al. 2017). Meta-analyses of child and adolescent psychotherapy have consistently found that, across treatment modalities, therapist-youth working alliance has a moderate effect size on youth outcomes, even in the context of structured treatments like cognitive-behavioral therapy, and that certain relational variables (e.g., counselor empathy, genuineness, and warmth; counselor direct influence skills on youth; youth willingness to participate) significantly boost outcomes (Karver et al. 2018; Karver et al. 2006).

Training in these universal characteristics of effective helping relationships can ensure a strong foundation for targeted skills development and remediation. However, the twofold difference in effect sizes between targeted and non-specific mentoring interventions in the current study suggests that relationship-building may be a necessary, but not sufficient, basis for helping mentees achieve change.

Study Strengths, Limitations, and Future Research Directions

The following limitations should be acknowledged when interpreting the study findings. First, evaluations of more targeted, problem-specific programs benefit from the fact that they measure a more restricted range of youth outcomes that are more directly aligned with the treatment goals. In non-specific programs, differential patterns of improvement can be essentially washed out when widely diverse outcomes are calculated into overall averages (Rose 2016). In addition, many other variables (e.g., sample, program size and length, implementation or evaluation quality) could account for different effects across these types of programs. It was not possible to estimate a multivariate model including all possible confounders of the targeted non-targeted effect because of missing data in many of these possible confounders; however, many of these variables (including many sociodemographic variables) did not emerge as significant in the meta-analysis by Raposa et al. (2019) in univariate analyses. Nonetheless, it will be important for future research to include randomized trials that control for these variables and directly compare the effects of non-specific to targeted program models.

It is also important to note that there was significant heterogeneity in outcomes both within and across studies, and a number of the targeted mentoring models examined in these analyses produced relatively small effects. Thus, although the meta-analytic data support targeted programs more generally, there are certainly instances where this is not the case. Future program evaluations should consider and measure the wide variety of mentor, youth, and program characteristics that may account for differences in program effectiveness (even within the more targeted, problem-specific category). It is also notable that programs were coded as targeted if they matched specific interventions to youth's presenting problems, irrespective of the evidence base, rigor, and fidelity of their interventions (and only included studies that specifically delineated program practices such that they could be clearly categorized into one of two mentoring models were included). Again, it is important for mentoring programs to articulate the theories and evidence base from which they develop their interventions as well as the fidelity with which these interventions are carried out, and for future meta-analyses to examine these factors as moderators of program effectiveness.

The analyses only included studies that were written and published in English, and all but two (i.e., Simões and Alarcão 2014; Bodin and Liefman 2011) were conducted in North America. As a result, the results are not necessarily generalizable to mentoring programs operating outside of North America. Further, although the Raposa et al. (2019) meta-analysis from which the subset of studies was drawn did not find strong evidence of publication bias, it remains likely that this sample of studies is not representative of all mentoring programs being implemented and evaluated.

Despite these limitations, there are a number of strengths to the current study. This meta-analysis explored a critical, previously untested moderator of youth mentoring effectiveness. Previous studies (e.g., Raposa et al. 2019) have examined some program characteristics as predictors of program effectiveness (e.g., structured vs. unstructured vs. semi-structured, whether or not programs use a curriculum), but found no impact on outcomes. It is possible that these previous findings were null because these categorizations do not take into account whether structures and curricula are actually matched with youth's presenting problems.

Another strength of the current study is the use of rigorous statistical methodologies, particularly multi-level meta-analyses to account for variability in effect sizes both within and across studies. The analyses suggested that a substantial portion of the variance among effect sizes was accounted for within studies (30%). That is, even within the same program evaluation, there were significant differences in the size of observed effect sizes for different outcome measures or constructs. This finding suggests that the multi-level approach to meta-analysis should be considered as a particular strength of this study, given that a traditional meta-analysis only accounts for between-study heterogeneity among effect sizes. Future studies in this area should use a multi-level approach to more accurately explore the specific outcome-level (e.g., self-report versus teacher-report) and study-level (e.g., demographic background of the youth in the program) constructs that might be accounting for all sources of variability in effect sizes.

Implications for Practice

The findings from the current study suggest several recommendations for future research and practice. Currently, there is inconsistency among mentoring programs in adherence to empirically supported program practices (e.g., recruitment and training strategies) recommended by organizations such as the National Mentoring Resource Center or MENTOR's Elements of Effective Practice for Mentoring (Garringer et al. 2015). Implementation of evidence-based practices has the potential to increase mentoring match lengths (Kupersmidt et al. 2017), thus increasing the likelihood of positive youth outcomes (Grossman et al.

2012; Herrera et al. 2011). Moreover, mentoring programs and researchers can partner to shift from non-specific approaches and activities to targeted and precise interventions that have stronger theoretical and empirical foundations (McQuillin and Lyons 2016; Weiler et al. 2017). Still, it is important to acknowledge that even the largest effects of targeted programs found in the analyses (i.e., 0.28, on academic outcomes) are still considered small by standard conventions, suggesting the need for ongoing innovation, research, and implementation of evidence-based practices to increase impact for youth participants.

Of course, it will be important to introduce and implement changes to mentoring practices in ways that are compatible with existing organizational structures. In particular, it may not be realistic for large mentoring programs that serve a broad array of youth, like Big Brothers Big Sisters, to develop interventions and trainings that target every potential presenting problem. Rather, large programs can partner with other organizations and embed their volunteers into programs that already offer specialized, evidence-based interventions and deploy mentors to play supporting roles. Likewise, the burgeoning field of mental health apps and other technology-delivered interventions have the potential to provide large, non-specific programs with a growing array of targeted, evidence-based interventions. In this "blended" model, youth learn and practice targeted skills through youth-friendly apps, with mentors reinforcing this engagement. In both embedded and blended approaches, mentors encourage youth to complete lessons and activities and provide contexts for practicing new skills (Mohr et al. 2011).

Conclusion

In recent decades, mentoring programs have become increasingly prevalent interventions to support positive developmental trajectories for youth experiencing a range of emotional, behavioral, and academic difficulties. The dominant "non-specific" model of youth mentoring programs assumes that a supportive relational bond between volunteers and youth can promote positive developmental changes and prevent a range of negative outcomes. However, findings from years of systematic reviews and meta-analyses of mentoring programs yielding small effects on youth outcomes have led to growing calls for mentors to employ targeted skills designed to match the presenting concerns of mentees. Despite these recommendations, no studies to date have directly compared non-specific versus targeted approaches to mentoring. Results from the current meta-analysis of 48 mentoring outcome studies suggest that non-specific, relationship-based models of mentoring may be less effective than more targeted, problem-specific approaches in addressing the needs and challenges of

today's youth. Moreover, moderator analyses revealed that targeted approaches may be particularly effective for promoting important developmental outcomes such as youths' academic performance, psychological wellbeing, and social functioning. Yet, as the field corrects for an overemphasis on non-specific approaches and moves in more targeted, evidence-based directions, it should resist veering too far from the relational component that sets the field apart from pure tutoring or other skills-based training. Programs should strive to find equilibrium between relational bonds and the delivery of more targeted and specific approaches to mentoring. When this balance is achieved, the mentoring relationship may be poised to better address the particular needs and challenges of today's youth.

Compliance with Ethical Standards

References

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

Bernstein, L., Rappaport, C. D., Olsho, L., Hunt, D., & Levin, M. (2009). Impact evaluation of the US Department of Education's

Student Mentoring Program. Final Report. NCEE 2009-4047. *National Center for Education Evaluation and Regional Assistance*.

Bodin, M., & Leifman, H. (2011). A randomized effectiveness trial of an adult-to-youth mentoring program in Sweden. *Addiction Research & Theory*, 19(5), 438–447.

Cavell, T. A., & Elledge, L. C. (2015). Mentoring and prevention science. In M. J. Karcher & D. L. DuBois (Eds.), *Handbook of youth mentoring* (2nd edn., pp. 29–43). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412996907.n3.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological Methods*, 19, 211–229.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. https://doi.org/10.1177/001316446002000104.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd edn., Hillsdale: Erlbaum.

Cook, B. L., Barry, C. L., & Busch, S. H. (2013). Racial/ethnic disparity trends in children's mental health care access and expenditures from 2002 to 2007. *Health Services Research*, 48(1), 129–149. https://doi.org/10.1111/j.1475-6773.2012.01439.x.

DeSocio, J., VanCura, M., Nelson, L. A., Hewitt, G., Kitzman, H., & Cole, R. (2007). Engaging truant adolescents: results from a multifaceted intervention pilot. *Preventing School Failure: Alternative Education for Children and Youth*, 51(3), 3–9.

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: a meta-analytic review. *American Journal of Community Psychology*, 30(2), 157–197. https://doi.org/10.1023/A:1014628810714.

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest*, 12(2), 57–91. https://doi.org/10.1177/1529100611414806.

Fernandes-Alcantara, A. L. (2018). *Vulnerable youth: federal mentoring programs and issues*. Washington, D.C: Congressional Research Services.

Garringer, M., Kupersmidt, J. B., Rhodes, J. E., Stelter, R. L., & Tai, T. (2015). *Elements of effective practices for mentoring*. Boston, MA: MENTOR: The National Mentoring Partnership.

Garringer, M., McQuillin, S., & McDaniel, H. (2017). Examining youth mentoring services across America: findings from the 2016 National Mentoring Program survey. Boston, MA: MENTOR: The National Mentoring Partnership.

Grossman, J. B., Chan, C. S., Schwartz, S. E. O., & Rhodes, J. E. (2012). The test of time in school-based mentoring: the role of relationship duration and re-matching on academic outcomes. *American Journal of Community Psychology*, 49(1–2), 43–54. https://doi.org/10.1007/s10464-011-9435-0.

Heckman, J. J., & Kautz, T. (2013). Fostering and measuring skills: interventions that improve character and cognition (No. w19656). National Bureau of Economic Research.

Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York: Academic Press.

Herrera, C., Grossman, J. B., Kauh, T. J., & McMaken, J. (2011). Mentoring in schools: an impact study of Big Brothers Big Sisters school-based mentoring. *Child Development*, 82(1), 346–361. https://doi.org/10.1111/j.1467-8624.2010.01559.x.

Holt, L. J., Bry, B. H., & Johnson, V. L. (2008). Enhancing school engagement in at-risk, urban minority adolescents through a school-based, adult mentoring intervention. *Child & Family Behavior Therapy*, 30(4), 297–318.

Jarjoura, G. R., Tanyu, M., Forbush, J., Herrera, C., & Keller, T. E. (2018). *Evaluation of the mentoring enhancement demonstration program: Technical report*. Washington, D.C.: American Institute for Research.

Jent, J. F., & Niec, L. N. (2006). Mentoring youth with psychiatric disorders: the impact on child and parent functioning. *Child & Family Behavior Therapy*, 28(3), 43–58.

Johnson, A. W. (1997). Mentoring at-risk youth: a research review and evaluation of the impacts of the Sponsor-A-Scholar program on student performance. Dissertations available from ProQuest. AAI9727243. https://repository.upenn.edu/dissertations/AAI9727243.

Karcher, M. J., Davis, C., & Powell, B. (2002). The effects of developmental mentoring on connectedness and academic achievement. *The School Community Journal*, 12(2), 35–50.

Karver, M. S., De Nadai, A. S., Monahan, M., & Shirk, S. R. (2018). Meta-analysis of the prospective relation between alliance and outcome in child and adolescent psychotherapy. *Psychotherapy*, 55(4), 341 https://doi.org/10.1037/pst0000176.

Karver, M. S., Handelsman, J. B., Fields, S., & Bickman, L. (2006). Meta-analysis of therapeutic relationship variables in youth and family therapy: the evidence for different relationship variables in the child and adolescent treatment outcome literature. *Clinical Psychology Review*, 26(1), 50–65. https://doi.org/10.1016/j.cpr.2005.09.001.

King, K. A., Vidourek, R. A., Davis, B., & McClellan, W. (2002). Increasing self-esteem and school connectedness through a multidimensional mentoring program. *Journal of School Health*, 72(7), 294–299.

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17), 2693–2710.

Kraemer, H. C., & Kupfer, D. J. (2006). Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry*, 59(11), 990–996.

Kupersmidt, J. B., Stump, K. N., Stelter, R. L., & Rhodes, J. E. (2017). Predictors of premature match closure in youth mentoring relationships. *American Journal of Community Psychology*, 59(1–2), 25–35. https://doi.org/10.1002/ajcp.12124.

Li, J., & Julian, M. M. (2012). Developmental relationships as the active ingredient: a unifying working hypothesis of "what works" across intervention settings. *American Journal of Orthopsychiatry*, 82(2), 157–166. https://doi.org/10.1111/j.1939-0025.2012.01151.x.

Lyons, M. D., McQuillin, S. D., & Henderson, L. J. (2019). Finding the sweet spot: investigating the effects of relationship closeness and instrumental activities in school-based mentoring. *American Journal of Community Psychology*, 63(1–2), 88–98.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276–282.

McQuillin, S. D., & Lyons, M. D. (2016). Brief instrumental school-based mentoring for middle school students: theory and impact. *Advances in School Mental Health Promotion*, 9(2), 73–89. https://doi.org/10.1080/1754730x.2016.1148620.

Mohr, D., Cuijpers, P., & Lehman, K. (2011). Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research*, 13(1), e30.

Raposa, E. B., Rhodes, J. E., & Herrera, C. (2016). The impact of youth risk on mentoring relationship quality: do mentor characteristics matter? *American Journal of Community Psychology*, 57, 320–329. https://doi.org/10.1002/ajcp.12057.

Raposa, E. B., Rhodes, J. E., Stams, G. J. J. M., Card, N., Burton, S., Schwartz, S. E. O., … Hussain, S. B. (2019). The effects of youth mentoring programs: a meta-analysis of outcome studies. *Journal of Youth and Adolescence*. https://doi.org/10.1007/s10964-019-00982-8.

Rhodes, J. E. (2005). A theoretical model of youth mentoring. In D. L. DuBois & M. J. Karcher (Eds.), Handbook of youth mentoring (pp. 30–43). Thousand Oakes: Sage Press.

Rose, T. (2016). *The end of average: how we succeed in a world that values sameness.* New York, NY: Harper One.

Simões, F., & Alarcão, M. (2014). Promoting well-being in school-based mentoring through basic psychological needs support: does it really count? *Journal of Happiness Studies*, 15(2), 407–424.

Spencer, R. (2007). "It's not what I expected": a qualitative study of youth mentoring relationship failures. *Journal of Adolescent Research*, 22(4), 331–354.

Sowers, J. A., Powers, L., Schmidt, J., Keller, T. E., Turner, A., Salazar, A., & Swank, P. R. (2017). A randomized trial of a science, technology, engineering, and mathematics mentoring program. *Career Development and Transition for Exceptional Individuals*, 40(4), 196–204.

Tabachnik, B. G., & Fidell, L. S. (2013). *Using multivariate statistics.* 6th edn. Boston: Allyn and Bacon.

Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: a review of meta-analyses. *Prevention Science*, 19(8), 1091–1101. https://doi.org/10.1007/s11121-018-0942-1.

Vázquez, A. L., & Villodas, M. T. (2018). Racial/ethnic differences in caregivers' perceptions of the need for and utilization of adolescent psychological counseling and support services. *Cultural Diversity and Ethnic Minority Psychology*. https://doi.org/10.1037/cdp0000255.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2014). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, 47, 1274–1294. https://doi.org/10.3758/s13428-014-0527-2.

Weiler, L. M., Chesmore, A. A., Pryce, J., Haddock, S. A., & Rhodes, T. (2017). Mentor response to youth academic support–seeking behavior: does attunement matter? *Youth & Society*, 1–22. https://doi.org/10.1177/0044118×17697235.

Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual clinical care: a multilevel meta-analysis. *JAMA Psychiatry*, 70(7), 750–761.

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., & Weersing, V. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. *American Psychologist*, 72(2), 79.

Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, 60(6), 628.

Kirsten M. Christensen is a doctoral candidate in clinical psychology at the University of Massachusetts Boston. Her major research interests include youth-adult partnerships in mentoring and community-based youth program contexts.

Matthew A. Hagler is a doctoral candidate in clinical psychology at the University of Massachusetts Boston. His research examines socioeconomic and demographic factors that impact the distribution, quality, and nature of mentoring relationships.

Geert-Jan Stams is a Professor of Forensic Child and Youth Care Sciences at the University of Amsterdam. His major research interests include meta-analysis and youth delinquency.

Elizabeth B. Raposa is an Assistant Professor in the Department of Psychology at Fordham University. Her research investigates how early life stressors influence risk for mental health problems across development, and how close relationships with parents, peers, or others mitigate the negative impact of stressors on youth.

Samantha Burton is a doctoral candidate in clinical psychology at the University of Massachusetts, Boston. Her major research interests include intergenerational and peer mentoring relationships, including their impact on youth outcomes.

Jean E. Rhodes is the Frank L. Boyden Professor of Psychology and the Director of the Center for Evidence-Based Mentoring at the University of Massachusetts Boston. She has devoted her career to understanding the role of both natural and assigned mentoring relationships in the lives of marginalized youth.