

Journal of Counseling Psychology

The Efficacy of Empathy Training: A Meta-Analysis of Randomized Controlled Trials

Emily Teding van Berkhout and John M. Malouff

Online First Publication, July 20, 2015. <http://dx.doi.org/10.1037/cou0000093>

CITATION

van Berkhout, E. T., & Malouff, J. M. (2015, July 20). The Efficacy of Empathy Training: A Meta-Analysis of Randomized Controlled Trials. *Journal of Counseling Psychology*. Advance online publication. <http://dx.doi.org/10.1037/cou0000093>

The Efficacy of Empathy Training: A Meta-Analysis of Randomized Controlled Trials

Emily Teding van Berkhout and John M. Malouff
University of New England, Australia

High levels of empathy are associated with healthy relationships and prosocial behavior; in health professionals, high levels of empathy are associated with better therapeutic outcomes. To determine whether empathy can be taught, researchers have evaluated empathy training programs. After excluding 1 outlier study that showed a very large effect with few participants, the meta-analysis included 18 randomized controlled trials of empathy training with a total of 1,018 participants. The findings suggest that empathy training programs are effective overall, with a medium effect ($g = 0.63$), adjusted to 0.51 after trim-and-fill evaluation for estimated publication bias. Moderator analyses indicated that 4 factors were statistically significantly associated with higher effect sizes: (a) training health professionals and university students rather than other types of individuals, (b) compensating trainees for their participation, (c) using empathy measures that focus exclusively on assessing understanding of the emotions of others, feeling those emotions, or commenting accurately on the emotions, and (d) using objective measures rather than self-report measures. Number of hours of training and time between preintervention assessment and postintervention assessment were not statistically significantly associated with effect size, with 6 months the longest time period for assessment. The findings indicate that (a) empathy training tends to be effective and (b) experimental research is warranted on the impact of different types of trainees, training conditions, and types of assessment.

Keywords: efficacy, empathy, meta-analysis, training

This meta-analysis examined the effects of empathy training. Over the past century, a variety of disciplines from the arts to neuroscience have shown interest in the topic of empathy (Tudor, 2011). Studies have demonstrated that psychotherapist empathy is an important feature of successful treatment by psychologists (Watson, Steckley, & McMullen, 2014), social workers (Gerdes & Segal, 2009), and substance abuse counselors (Moyers & Miller, 2013). Studies have also found that empathy is associated with better patient outcomes for physicians (Hojat et al., 2011) and increased patient adherence to treatments (Vermeire, Hearnshaw, Van Royen, & Denekens, 2001). In nonprofessionals, research results have indicated that high levels of empathy are associated with enhanced personal relationships (Long, Angera, Carter, Nakamoto, & Kalso, 1999) and prosocial behavior (Telle & Pfister, 2012).

Conversely, studies have shown that a lack of empathy is associated with negative outcomes, including aggressive behavior such as bullying and sexual offending (Ang & Goh, 2010; Jolliffe & Farrington, 2006; Salmon, 2003). Additionally, research findings have suggested that disorders such as autism may be associated with neurological impairments in the empathy systems

(Shamay-Tsoory, 2011), and that deficits in empathy are linked to psychopathy (Ali, Amorim, & Chamorro-Premuzic, 2009; Blair, 2008).

The research findings on empathy suggest that it would be worthwhile to increase empathy in individuals. Many studies have examined whether that is possible, but, as yet, no meta-analysis of the effects of empathy training has been published.

Defining Empathy

The concept of empathy gained popularity in the 1950s after Carl Rogers emphasized it as essential for successful psychotherapy (Gladstein, 1983). However, the definition of empathy remains contested. This controversy arises in part from the debate whether empathy is predominantly a cognitive or an affective process.

Cognitive and Affective Empathy

Cognitive approaches to empathy suggest that it is an intellectual ability enabling an individual to view the world from another person's perspective (Duan & Hill, 1996; Gladstein, 1983). Some scholars adopted cognitive views of empathy as an ability to sense or understand the experience, feelings, or mental state of another person while remaining an objective observer (Hogan, 1969; Kohut, 1959; Rogers, 1992). In this cognitive view, empathy is predominantly an intellectual, perspective-taking process. Other scholars (e.g., Gladstein, 1983; Hoffman, 2000; Mehrabian & Epstein, 1972) embraced an affective approach to empathy, wherein empathy involves having a matching or corresponding emotional reaction to the emotions of another individual.

Emily Teding van Berkhout and John M. Malouff, Department of Psychology, University of New England, Australia.

Correspondence concerning this article should be addressed to John M. Malouff, Department of Psychology, University of New England Psychology, Armidale NSW 2351, Australia. E-mail: jmalouff@une.edu.au or emily.t.v.b@gmail.com

Multidimensional Views of Empathy

Cognitive and affective approaches to empathy have been criticized, with advocates of both streams suggesting that one component of empathy occurs as a result of the other and is therefore secondary (Clark, 2007; Kerem, Fishman, & Josselson, 2001). A multidimensional approach to empathy embraces a more nuanced understanding in which empathy includes multiple processes occurring in isolation or together. Davis's (1983) suggestion that empathy involves reacting to the observed experience of another person, either cognitively or affectively, is a common multidimensional conceptualization of empathy. Davis created the Interpersonal Reactivity Index (IRI; Davis, 1983) and proposed four subscales measuring these reactions—two assessing the cognitive dimensions of empathy (Perspective Taking and Fantasy) and two assessing the affective dimensions (Empathic Concern and Personal Distress). Other common multidimensional understandings of empathy include social work models. Gerdes, Lietz, and Segal (2011), for example, suggested that behaviors performed in response to the experience of the components of empathy are important in conceptualizing empathy.

Empathy Defined for This Meta-Analysis

For this meta-analysis, we defined *empathy* as understanding the emotions another person is feeling, feeling the same emotions another person is feeling, or commenting accurately on the emotions another person is feeling.

Reviews of the Effects of Empathy Training

Researchers have attempted to teach individuals the meaning of empathy, to recognize emotions in others, to take the perspective of others, and to show empathy in various social situations. Researchers have examined the efficacy of empathy training programs developed for use with individuals from a variety of backgrounds. Butters's (2010) unpublished meta-analysis of training programs in client populations found an overall large effect of training for the 24 studies examined. Butters's analysis was limited by the inclusion of poor quality studies with no control groups, resulting in potential internal validity issues wherein the training outcome may have been a result of methodological factors rather than training. A further limitation of this meta-analysis is that it did not consider training outside client populations.

Lam, Kolomitro, and Alamparambil (2011) also examined empathy training in various populations in a review of 29 articles, including qualitative papers, randomized controlled trials (RCTs), and quasi-experimental studies. The review concluded that training programs were generally successful as participants learned about the concept of empathy, but that the impact of training on the outward demonstration of empathy was lacking. The authors suggested that limitations within the studies, for example, incongruences between empathy taught and empathy measured, may have accounted for this lack of success.

Many empathy training program studies have attempted to increase empathy levels in physicians, therapists, medical students, and nurses. An unpublished meta-analysis of empathy training in helping professions (Dexter, 2012) found an overall large effect of empathy training. However this meta-analysis was subject to sig-

nificant publication bias and included studies with various methodological limitations, such as nonrandomization and lack of comparison groups. A systematic review of empathy training for undergraduate medical students concluded that, for 15 of the 18 studies included in the review, training showed significant increases in empathy levels (Batt-Rawden, Chisolm, Anton, & Flickinger, 2013). Similar to the review of Lam et al. (2011), however, the results were limited in that a review by its nature does not test the statistical significance of those increases in empathy or include a systematic search for moderators of effect size.

Brunero, Lamont, and Coates (2010) conducted a review of 17 empathy training program studies in nursing populations. Eleven of the studies identified significant improvements in empathy, while six did not—and two of these reported negative results. Again, this study was limited in that the overall effect of the training programs examined cannot be determined. Further, the review included predominantly nonrandomized trials and studies without control groups.

In addition to empathy training programs for health workers, researchers have evaluated training programs for couples (e.g., Rogge, Cobb, Lawrence, Johnson, & Bradbury, 2013) and for children and adolescents. A meta-analysis of programs designed to reduce prejudice among children found that interventions involving empathy training produced stronger effects than those without (Beelmann & Heinemann, 2014). While not a direct examination of empathy training, the results of this meta-analysis suggested potential for empathy levels in children to be enhanced. A final popular area for empathy training is psychiatric patients and criminal offenders (e.g., Lomis & Baker, 1985; Pecukonis, 1990).

Potential Variables Moderating Empathy Training

Type of Trainee

Empathy begins to develop at an early age, but the brain regions used for these skills may not fully develop until late adolescence (Choudhury, Blakemore, & Charman, 2006). It is possible that training empathy has different effects on young individuals and on adults. Learning capacity, as indicated by educational status of the trainees, might also affect outcomes. In the present meta-analysis, the types of trainees fit best into these categories: university students, health professionals, patients, other adults, teens, and children.

Type of Empathy Trained

Three types of empathy could be targeted in training: cognitive, affective, and behavioral. Some types of empathy may be more conducive to training than others. For example, Pecukonis (1990) found that training increased the affective empathy of adolescent females but had no impact on cognitive empathy. In the current meta-analysis, the type of empathy trained fit best into these categories: cognitive and affective; cognitive and behavioral; and cognitive, affective, and behavioral.

Teaching Methods

Empathy training programs employ a number of methods. In a review, Lam et al. (2011) found that the most common methods

included experiential training (instructors provide “experiences” such as games and role-play), didactic (lecture based), skills training (lectures, demonstrations and practice), and mixed methods. Many of these correspond to methods found in behavioral skills training, which includes modeling, instructions, rehearsal, and feedback (Mittenberger, 2015). For the present meta-analysis, we divided training methods into those that included all four components of behavioral skills training and those that did not include all four components.

Intervention Length

There is currently no consensus regarding the optimum length of time for an empathy training course or whether differences in duration or intervals impact training outcomes. Some reviews of interpersonal skills training, including empathy training, have suggested a training length of 1–3 days (Berkhof, van Rijssen, Schellart, Anema, & van der Beek, 2011), and others, such as Butters (2010), failed to find an association between training time and effect size, concluding that as little as 1 hr may be effective.

Compensation

Providing a financial incentive for research participation might create measurement bias in which participants respond with socially or trainer-desired answers (see Head, 2009). Further, compensation may motivate extra learning effort or keep in training individuals who would ordinarily drop out and therefore not benefit from the training (see Ripley, 2006). In the present meta-analysis, compensation fit into three categories: monetary compensation, partial academic credit for a university course, and no compensation.

Type of Control Group

Control groups can be active in the sense of providing participants with a special activity as part of the study or they can be inactive in the sense of waiting to enter training or involving training as usual. Researchers generally consider active control conditions a better test of an intervention because they may be able to control for placebo effects (Boot, Simons, Stothart, & Stutts, 2013). In the current meta-analysis, we categorized control groups as either active or waiting list/training as usual.

Scope of Empathy Measured

Different developers of empathy measures have used more or less expansive content for their measures. We view empathy as understanding, feeling, or commenting accurately on another’s emotions. We included test measures such as the Carkhuff (1969) Empathy Rating Scale, in which a person responds either orally or in writing to a statement and raters then evaluate the level of empathy shown. We also included observational measures where patients, for instance, rate a health professional on level of empathy shown in actual interactions. Finally, we included self-report measures of empathy. We coded those measures as narrow measures of empathy. Other “empathy” measures include some items that ask about having sympathy (feeling sorry for individuals who have had bad experiences), being nice, feeling distressed at the plight of others, or other matters not central to empathy. For

instance, the La Monica (1981) empathy measure asks about soft-heartedness and other constructs, and the total IRI (Davis, 1983) includes items asking about personal distress and sympathy. The Empathy Quotient (Baron-Cohen & Wheelwright, 2004) includes as an item “I really enjoy caring for other people.” We included these measures, coded as broad measures of empathy, because they have at least some evidence of validity as a measure of empathy and some empathy researchers consider them to be good enough measures to evaluate the effects of empathy training. In the present meta-analysis, we categorized empathy measure scope as narrow, broad, or mixed.

Whether Outcome Measures Are Self-Report or Objective

Distinguishing between objective and self-report measures may impact the results of empathy training. Butters’s (2010) unpublished meta-analysis found that the effect sizes of each study were smaller when self-report measures were used. Similarly, in an unpublished dissertation, Reed (1996) reported no significant changes in empathy using self-report, but found higher levels of behavioral empathy when using objective measures. In the present meta-analysis, we categorized objective measures as those that involve tests or ratings by independent others, such as patients. Two other categories used were self-report/other and mixed objective and self-report/other.

Duration of Effects

It could be that the effects of empathy training are brief. On the other hand, if increased empathy is reinforced or otherwise becomes habitual, the benefits of empathy training might endure. One way to evaluate how long effects last is to evaluate over what time periods empathy training programs assess empathy after the baseline assessment.

Present Meta-Analysis

The aim of the present study was to use meta-analysis of RCTs to determine the overall effect of empathy training programs. A further aim was to evaluate possible moderators of effect size. In setting these aims, we hoped to minimize the limitations of prior reviews and meta-analyses relating to the effects of empathy training by (a) including all high-quality (randomized controlled) trials up to the current time, (b) calculating a conservative meta-analytic effect size across all studies, and (c) statistically examining possible moderators of effect size across the studies.

The main hypothesis was that empathy training would be efficacious. The remaining hypotheses were that the following factors would be associated with effect size: the type of empathy (cognitive, affective, or behavioral) targeted in training, the type of measure (self-report or objective), whether the population was specific types of adults or children/adolescents, whether a training program included all four behavioral skills training principles (modeling, instruction, practice, and feedback), whether participants received compensation, the total number of training hours, and the amount of time between pre- and posttesting.

Method

Literature Search

We searched the electronic databases PsycINFO, ProQuest, SAGE, Google Scholar, and ProQuest Dissertations and Theses to find English-language published or unpublished RCTs examining empathy training. We completed the search in July 2014 using the search terms *empathy*, *train* or *teach* or *course* or *intervention*, *random*, and *control group*. We also searched the reference lists of these articles for additional studies. Figure 1 provides an overview of the search process.

Inclusion Criteria

We used the following criteria to select studies for inclusion in the meta-analysis:

1. The study must be an RCT with an empathy training group and a control group that did not receive empathy training.
2. The study must use appropriate randomization methods. If the randomization procedure was compromised, we excluded the study.
3. Empathy must be a main target of training.
4. Empathy must be measured.
5. The study must provide empirical information on the outcome of training—either an effect size or data that can be used to calculate an effect size.

Coding Process

We coded information on study results (e.g., effect size data, direction of effect), descriptive data (e.g., the specific empathy

measures used), and information about the potential moderators mentioned above. One of us coded the articles and the other checked the coding. We occasionally changed category definitions to conceptualize better potential moderators. We assessed interrater agreement with three moderators. For these three, the level of agreement for initial coding was: trainee type, 18 of 19 (95%); control group type, 15 of 18 (84%); and scope of outcome measures per study, 19 of 19 (100%). We resolved all disagreements by discussion and used consensus coding for the meta-analysis.

Meta-Analytic Procedures

Following the example of Lambert and Alhassoon (2015), we calculated study effect sizes using the relatively conservative Hedges's g , which corrects a bias of Cohen's d while presenting the mean group difference over the pooled standard deviation. We based the study effect sizes on the data provided in the reports, such as pre and post means and standard deviations of groups. When a study had two training conditions with different levels of comprehensiveness, we used the results for the more comprehensive training. When a study had multiple relevant outcome measures, we used as the study effect size the mean of the relevant effect sizes for the study. We used a random-effects model because we viewed the studies as a sample of possible studies, and we wanted to generalize beyond the sample. For assessing homogeneity, we used the Q statistic and I^2 . We assessed potential publication bias using the funnel plot technique and Orwin's fail-safe N .

Results

The literature search resulted in 19 studies eligible for inclusion in the meta-analysis. Table 1 provides a descriptive summary of the included studies. We use "significant" in this section to mean statistically significant at $p \leq .05$, unless otherwise indicated.

Overall Effect of Training

With all 19 studies, the overall effect size for the random effects model was statistically significant, $g = 0.73$, $p < .001$, 95% confidence interval (CI) [0.45, 1.02]. However, the studies included one extreme outlier, with an effect size of g equal to 5.83 (Crabb, Moracco, & Bender, 1983), more than twice as high as that for any other study. The Q value for heterogeneity for all 19 studies was 59; without the outlier, the Q was much lower at 38. According to Hedges and Olkin (1985), when a meta-analysis has a much lower Q without a study, that study is an outlier that does not fit in the sample of studies. Lipsey and Wilson (2001) suggested that, because outlier effect sizes can significantly distort results, researchers should remove outliers. We decided to present the overall effect size with and without the outlier. Without the outlier study, for 18 studies with a total of 1028 participants, the overall effect size was g equal to 0.63, $p < .001$, 95% CI [0.39, 0.87].

The Q value of 38 for all the studies except the outlier indicated significant heterogeneity, with p equal to .003, and I^2 equal to 55, indicating that 55% of variability was caused by true heterogeneity rather than error.

We used all studies except the outlier to assess the risk of publication bias. Figure 2 shows the funnel plot of effect sizes. The

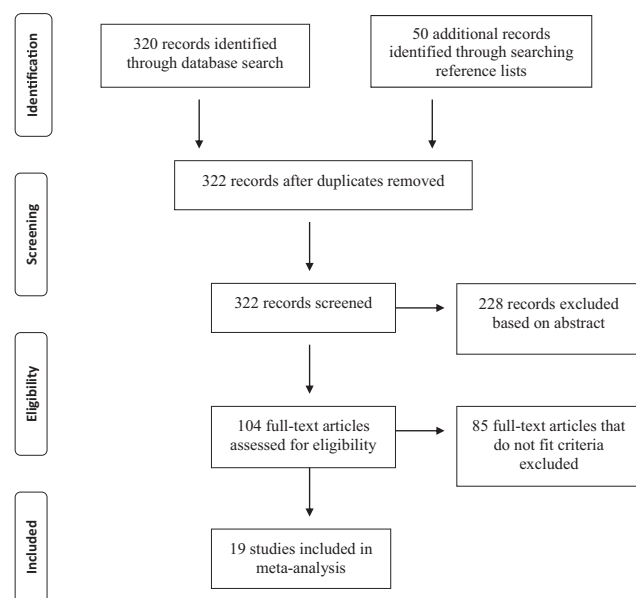


Figure 1. Flow of articles.

Table 1
 Summary of Studies Included in the Meta-Analysis

Author (Year)	Trainee type	N		Type of empathy targeted	Outcome measure	Outcome type	Outcome scope ^b	Compensation	Active control	Behavioral skills training components	Hours trained	Months pre-post	Hedges's g
		Training	Control										
Archer & Kagan (1973)	University students	4	4	A, B, C	ASS	Objective	Narrow	No	No	Yes	24	1	1.875
Begeer et al. (2011)	Children with autism	19	17	A, C	IECA	Self-report	Mixed	No	No	No	24	4	0.118
Bunn & Terpstra (2009)	Medical students	100	50	A, C	JSPE	Self-report	Mixed	No	No	No	0.66	0	0.364
Crabb et al. (1983) ^a	Religious lay-persons	7	8	B, C	Carkhuff	Objective	Narrow	No	No	No	8	1.25	5.83
Császár (2012)	Student teachers	36	34	A, C	IRI	Objective	Mixed	No	No	No	NA	1.5	0.756
Dadds et al. (2012)	Children w. behavioral problems	66	68	A, C	GEM	Objective, other	Narrow	No	No	Yes	6	6	0.062
Daniels et al. (1998)	Student nurses	24	29	B, C	Carkhuff	Objective	Mixed	No	No	Yes	25	NA	0.877
Handmaker et al. (1999)	Health practitioners	15	15	B, C	RS	Objective	Narrow	Yes	Yes	No	0.33	0	1.219
Higgins (1990)	Medical students	7	6	B, C	BLRI	Objective	Narrow	No	No	Yes	12	1	1.832
Hodge et al. (1978)	University students	24	12	B, C	Carkhuff	Objective	Narrow	Yes	Yes	Yes	NA	0	1.5
Jacobs (1977)	High school students	20	10	A, B, C	BLRI	Self-report	Broad	No	Yes	Yes	7.5	0.06	0.824
Lomis & Baker (1985)	Forensic inpatients	8	8	B, C	Carkhuff, HES, RS, FC	Self-report, objective	Mixed	Yes	Yes	Yes	7.5	0.033	0.51
Newman (1993)	Senior citizens	9	10	A, C	IRI	Self-report	Broad	No	Yes	No	15	0.16	0.212
O'Neill & McMillan (2012)	Head injury patients	12	12	A, C	EQ	Self-report	Broad	No	Yes	No	0.5	0	0.137
Pecukonis (1990)	Aggressive teen females	10	10	A, B, C	HES	Self-report	Broad	No	No	Yes	6	2.3	0.563
Perry (1975)	Clergymen	11	11	B, C	Carkhuff	Objective	Narrow	No	Yes	No	NA	0	0.255
Reed (1996)	Medical students	12	11	A, B, C	AES	Objective	Mixed	No	Yes	Yes	30	2.5	1.069
Riess et al. (2012)	Physicians	54	45	B, C	BEES, NE, EFD, CRE	Self-report, objective	Mixed	No	No	No	3	NA	0.454
Uhlemann et al. (1976)	University students	20	10	A, C	Carkhuff	Objective	Narrow	Yes	Yes	No	0.33	0	1.499

Note. A = affective; AES = Accurate Empathy Scale; ASS = Affect Sensitivity Scale; B = behavioral; BEES = Balanced Emotional Empathy Scale; BLRI = Barrett-Lennard Relationship Inventory; C = cognitive; Carkhuff = Carkhuff Empathy Rating Scale; CRE = Consultation and Relational Empathy Measure; EFD = Ekman Facial Decoding test; EQ = Empathy Quotient; FC = Feelings Checklist; GEM = Griffith Empathy Measure; HES = Hogan Empathy Scale; IECA = Index of Empathy for Children and Adolescents; IRI = Interpersonal Reactivity Index; JSPE = Jefferson Scale of Physician Empathy; NA = information not available in study article; NE = Neurobiology of Empathy; RS = Study Developed Scale.

^a Study is an outlier. ^b Outcome scope: narrow = measures understanding, feeling, or commenting accurately on emotions of others; broad = measures empathy and other constructs such as being nice; mixed = both narrow and broad measures used in the study.

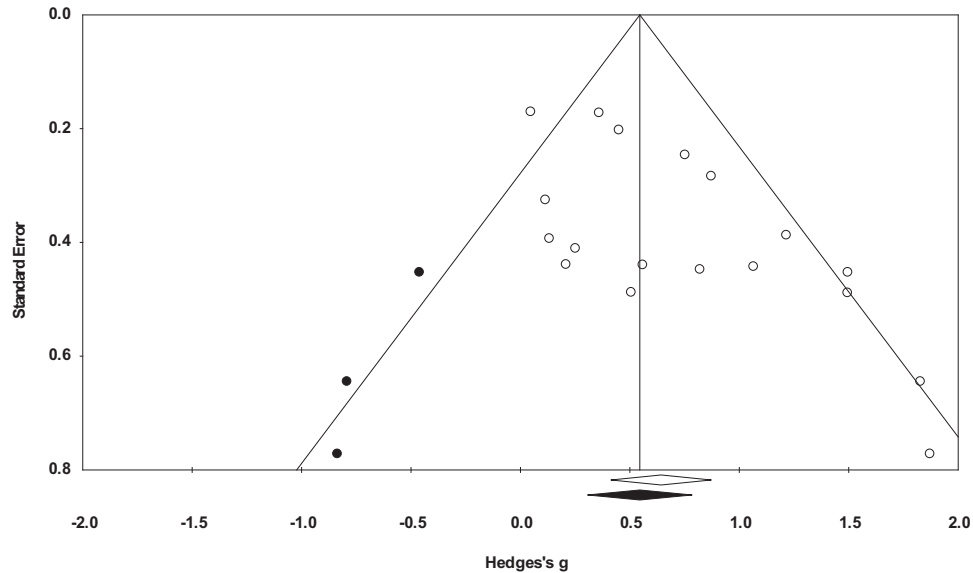


Figure 2. Funnel plot of effect sizes and standard errors. Open circles represent included study values. Solid circles represent imputed studies.

plot indicates a small amount of asymmetry, suggesting possible bias. Duval and Tweedie's trim-and-fill method imputed three missing studies into the plot. The adjusted effect size was g equal to 0.51, 95% CI [0.25, 0.77]. Orwin's fail-safe N of 43 ($k = 17$, trivial $g = 0.15$, $g = 0$ in missing studies) indicated that 43 studies of zero effect would be needed to reduce the mean effect to the trivial effect size of g equal to 0.15.

Moderator Analyses

To reduce the likelihood of distorted results in moderator analyses, we excluded the outlier study from the moderator analyses. Table 2 shows the results of the analyses, with the main outcome Q values comparing different levels of a variable. Four variables showed significant moderation: trainee type, compensation, scope of empathy measures, and whether the empathy measures were objective.

To further evaluate objective measures versus self-report, we determined the median effect size for objective and self-report measures in the two studies that used both types. For Lomis and Baker (1985), the median g was 0.74 for objective and 0.30 for self-report; for Riess, Kelley, Bailey, Dunn, and Phillips (2012), it was 0.78 for objective and 0.20 for self-report.

Metaregression with method of moments analysis showed a nonsignificant association between hours of training and effect size, slope estimate = .00, 95% CI [-0.02, 0.03], $p = .72$. Time between pre- and posttraining assessments ranged from 0–6 months, with a median time of 1 month. Metaregression showed a nonsignificant trend in the direction of a negative association of time between pre- and postassessments with effect size, slope estimate = -0.12, 95% CI [-0.28, 0.01], $p = .06$.

Discussion

The main aim of this meta-analysis was to determine the overall effect of empathy training programs. The overall effect size of the

19 studies was significant, supporting the hypothesis that empathy training would be efficacious. The overall effect size (g) ranges from 0.51 to 0.73, depending on whether an outlier study was included and on whether the results were adjusted for statistically presumed missing studies. All overall effect sizes were in the range of a moderate effect of training.

Sometimes outliers can provide useful information. In the present meta-analysis, the outlier study (Crabb et al., 1983) was unusual in that it presented only posttraining assessment data. Further, its participants were evangelical lay-persons and the training groups were exposed to a "biblical theory of personality" in which morality and prosocial behavior (e.g., empathy) were emphasized, directly prior to training in empathy. Research has indicated that prosocial behavior becomes stronger when participants have been exposed to religious priming (Pichon, Boccato, & Saroglou, 2007). Thus, the potential priming effects and the religious zeal of participants at the time of testing may have led to the extreme training effects on the objective outcome measure.

Moderators of Effect Size

Type of trainees. Studies that involved training health professionals and university students showed significantly higher effect sizes than studies with youths or with other types of adults. Studies with other types of adults or with children or teenagers failed to show a significant overall effect. The four studies that examined children and teens may not be reflective of the general population because three of the studies included youths with behavioral difficulties, aggression, and autism. Nevertheless, this result is consistent with developmental and neurological research findings that have suggested individuals must reach a certain level of neurological maturity before they can adequately understand and display empathy (Choudhury et al., 2006; Decety & Lamm, 2006).

Scope of empathy measures. Studies using outcome measures that assessed empathy in the narrow sense of understanding

Table 2
Categorical Moderator Analysis

Moderator	<i>k</i>	<i>g</i>	95% CI		<i>p</i>	Homogeneity analysis			<i>I</i> ²	
			Lower	Upper		<i>Q</i>	<i>df</i>	<i>p</i>		
Trainee type, $Q(5) = 15.89, p = .01$										
University student	8	1.06	0.65	1.47	.001	18.10	7	.01	61	
Health professional	2	0.76	0.29	1.50	.04	3.04	1	.01	67	
Adult patient	2	0.28	-0.32	0.89	.36	0.35	1	.55	0	
Other adult	2	0.24	-0.35	0.82	.43	0.00	1	.94	0	
Teenager	2	0.38	-0.22	0.98	.22	0.33	1	.57	0	
Child	2	0.07	-0.22	0.37	.63	0.02	1	.88	0	
Type of empathy trained, $Q(2) = 4.87, p = .09$										
Cognitive and affective	7	0.36	0.09	0.64	.01	10.72	6	.10	44	
Cognitive and behavioral	7	0.91	0.48	1.34	<.001	14.23	6	.03	58	
Cognitive, affective, and behavioral	4	0.77	0.18	1.35	.01	4.38	3	.22	32	
Used four components of behavior skills training, $Q(1) = 1.80, p = .18$										
Yes	9	0.87	0.39	1.35	<.001	26.48	8	.01	26	
No	9	0.50	0.26	0.74	<.001	11.29	8	.19	11	
Compensation provided, $Q(2) = 9.20, p = .01$										
Academic credit	2	1.50	0.85	2.15	<.001	0	1	.99	0	
Money	2	0.93	0.24	1.61	.01	1.29	1	.26	22	
None	14	0.49	0.27	0.70	<.001	21.43	13	.07	39	
Type of control group, $Q(1) = 0.70, p = .40$										
Active	9	.75	0.33	1.16	<.001	17	8	.03	52	
Waiting list or treatment as usual	9	.53	0.25	0.81	<.001	18	8	.02	56	
Scope of empathy measure used, $Q(2) = 6.40, p = .04$										
Narrow	8	0.94	0.43	1.45	<.001	30	7	.01	77	
Broad	5	0.23	-0.12	0.58	.20	1	4	.95	0	
Mixed or other	5	0.67	0.42	0.92	<.001	3	4	.61	0	
Self-report vs. objective measure used, $Q(2) = 15.69, p < .001$										
Self-report	6	0.30	0.06	0.54	.02	1.02	5	.96	0	
Objective	7	1.09	0.78	1.40	<.001	4.92	6	.55	0	
Both or other rating	5	0.53	0.12	0.95	.001	10.98	4	.03	64	

Note. CI = confidence interval; *df* = degrees of freedom.

the emotions of another, feeling those emotions, or commenting on the emotions had significantly higher effect sizes than studies using outcome measures that, on their face, assessed empathy plus other constructs such as being warm-hearted or nice. This finding makes sense because training in the studies focused on increasing empathy in the narrow sense.

Whether outcome measures were objective. Studies using objective measures, including written tests of ability to determine another person's emotions and ratings of empathic behavior by patients, showed significantly higher effect sizes than those using self-report, with studies that used both types of measures or a different type of measure intermediate. Further, the two studies that included both objective and self-report measures showed a pattern of the objective measures having over twice the mean effect size of the self-report measures, supporting the view that empathy training may lead to greater effects on objective measures than on self-report measures.

Training compensation. Prior reviews and meta-analytic research have not assessed compensation as a potential moderator. The association found in the present meta-analysis between compensation and higher effect sizes might be linked to compensation motivating participants to participate fully in training. That sort of motivation would not typically occur in nonresearch empathy training, except with regard to planned training activities for students.

Use of behavioral skills training principles. Empathy training studies involving all four components of behavioral skills training (instruction, modeling, practice, and feedback) had higher, but not significantly higher, effect sizes than other studies. The trend found is consistent with the meta-analytic conclusion of Hill and Lent (2006) that helping-skills training programs using multiple behavioral skills training methods produce significantly higher effect sizes than programs that do not.

Type of empathy trained. Studies that targeted cognitive and behavioral, or cognitive, affective, and behavioral, empathy showed higher, but not significantly higher, effect sizes than studies targeting cognitive and affective empathy. It could be that behavioral training provides a valuable aspect to training. Interestingly, all included studies targeted at least cognitive empathy. This may be because cognitive empathy is considered to involve processes that can be consciously acquired, whereas affective empathy is considered to be more autonomic, and behavioral empathy is considered to occur in response to the affective or cognitive empathy process (Batt-Rawden et al., 2013; Elliott, Bohart, Watson, & Greenberg, 2011; Jolliffe & Farrington, 2004).

Type of control group. There was no significant difference in empathy effects related to type of control group used in a study. That might be because it is hard to make participants think they are receiving empathy training without actually providing empathy training.

Training length. The metaregression found no significant evidence that number of training hours was associated with effect size. It is possible that these results are a consequence of population differences rather than the unimportance of training dose. This analysis involved populations ranging from children with disabilities to physicians, and it is possible that the time needed to produce empathy changes in each group varies. For example, Berkhof et al. (2011) have suggested that between 1 and 3 days of training is effective for physicians, while Gresham, Sugai, and Horner (2001) have suggested that over 30 hr of training is insufficient for children with disabilities. Thus, although the current results suggest that the amount of training has no effect on outcome, the nonsignificant result should be interpreted with caution.

Time from pre- to postassessment. The metaregression showed a nonsignificant trend of a negative association between longer periods between baseline assessment and final assessment and lower effect size. With these ambiguous results and a maximum time of 6 months from baseline to final assessment in the studies, it is hard to judge whether the effects of empathy training endure for long enough to make providing the training worthwhile.

Limitations of the Meta-Analysis

This meta-analysis addressed some of the limitations of previous reviews by employing stringent inclusion criteria requiring studies to be RCTs, by statistically combining study effect sizes, and by taking a conservative approach in using Hedges's g for effect size. Despite use of a comprehensive search strategy, including unpublished studies, it is possible that we did not discover all past RCTs of empathy training. We attempted to address this possibility by testing for publication bias and adjusting effect-size estimates accordingly.

Additionally there is the potential in meta-analysis for different moderators to overlap with or be related to each other, and therefore have a confounding impact on the moderator analyses, which, in the best of circumstances, are quasi-analytic and therefore cannot support causal conclusions (Lipsey, 2003).

With only 18 studies included in the moderator analyses, these analyses had enough power to identify as statistically significant only large differences in effect size. Nonsignificant trends might or might not also provide clues about possible moderators of effect size.

Conclusion and Future Directions

This meta-analysis found that empathy training programs tend to be effective in increasing empathy levels. The present overall results suggest that it could be worthwhile to train individuals in empathy and to evaluate, at least informally, the effects. The moderator results suggest that the training might work best with health professionals and university students, who are compensated for their time, with training in cognitive and behavioral empathy, using objective outcome measures that assess empathy in the sense of understanding, feeling, and commenting accurately on the emotions of others. However, much is still unclear about the effects of empathy training, including the mechanisms involved in positive effects. Future studies might examine experimentally (a) whether empathy training increases empathy in trainees other than univer-

sity students and health professionals, (b) to what extent training benefits endure after the end of training, (c) whether teaching a mix of cognitive and behavioral empathy leads to better effects than training not targeting behavioral empathy, (d) whether employing the four components of behavioral skills training leads to better effects than using only some of the components, and (e) whether compensating participants leads to better effects than not compensating them.

References

- References marked with an asterisk were included in the meta-analysis.
- Ali, F., Amorim, I. S., & Chamorro-Premuzic, T. (2009). Empathy deficits and trait emotional intelligence in psychopathy and Machiavellianism. *Personality and Individual Differences, 47*, 758–762. <http://dx.doi.org/10.1016/j.paid.2009.06.016>
- Ang, R. P., & Goh, D. H. (2010). Cyberbullying among adolescents: The role of affective and cognitive empathy, and gender. *Child Psychiatry and Human Development, 41*, 387–397. <http://dx.doi.org/10.1007/s10578-010-0176-3>
- *Archer, J., & Kagan, N. (1973). Teaching interpersonal relationship skills on campus: A pyramid approach. *Journal of Counseling Psychology, 20*, 535–540. <http://dx.doi.org/10.1037/h0035098>
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*, 163–175.
- Batt-Rawden, S. A., Chisolm, M. S., Anton, B., & Flickinger, T. E. (2013). Teaching empathy to medical students: An updated, systematic review. *Academic Medicine, 88*, 1171–1177. <http://dx.doi.org/10.1097/ACM.0b013e318299f3e3>
- Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of Applied Developmental Psychology, 35*, 10–24. <http://dx.doi.org/10.1016/j.appdev.2013.11.002>
- *Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., & Boer, F. (2011). Theory of mind training in children with autism: A randomized controlled trial. *Journal of Autism and Developmental Disorders, 41*, 997–1006. <http://dx.doi.org/10.1007/s10803-010-1121-9>
- Berkhof, M., van Rijssen, H. J., Schellart, A. J. M., Anema, J. R., & van der Beek, A. J. (2011). Effective training strategies for teaching communication skills to physicians: An overview of systematic reviews. *Patient Education and Counseling, 84*, 152–162. <http://dx.doi.org/10.1016/j.pec.2010.06.010>
- Blair, R. J. R. (2008). Fine cuts of empathy and the amygdala: Dissociable deficits in psychopathy and autism. *The Quarterly Journal of Experimental Psychology, 61*, 157–170. <http://dx.doi.org/10.1080/17470210701508855>
- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445–454. <http://dx.doi.org/10.1177/1745691613491271>
- Brunero, S., Lamont, S., & Coates, M. (2010). A review of empathy education in nursing. *Nursing Inquiry, 17*, 65–74. <http://dx.doi.org/10.1111/j.1440-1800.2009.00482.x>
- *Bunn, W., & Terpstra, J. (2009). Cultivating empathy for the mentally ill using simulated auditory hallucinations. *Academic Psychiatry, 33*, 457–460. <http://dx.doi.org/10.1176/appi.ap.33.6.457>
- Butters, R. P. (2010). *A meta-analysis of empathy training programs for client populations* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3398227).
- Carkhuff, R. R. (1969). *Human and helping relations* (Vols. 1 & 2). New York, NY: Holt, Rinehart and Winston.

- Choudhury, S., Blakemore, S. J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience, 1*, 165–174. <http://dx.doi.org/10.1093/scan/nsi024>
- Clark, M. A. (2007). *Cognitive and affective empathy: Exploring the differential effects of empathy components on work-family conflict and emotional labor* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1446883).
- *Crabb, W. T., Moracco, J. C., & Bender, R. C. (1983). A comparative study of empathy training with programmed instruction for lay helpers. *Journal of Counseling Psychology, 30*, 221–226. <http://dx.doi.org/10.1037/0022-0167.30.2.221>
- *Császár, I. (2012). *The effect of loving kindness meditation and student teachers stress and empathy* (Doctoral dissertation, Louisiana State University). Retrieved from <http://etd.lsu.edu/docs/available/etd-06272012-080110/unrestricted/csaszardiss.pdf>
- *Dadds, M. R., Cauchi, A. J., Wimalaweera, S., Hawes, D. J., & Brennan, J. (2012). Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Research, 199*, 201–207. <http://dx.doi.org/10.1016/j.psychres.2012.04.033>
- *Daniels, T. G., Denny, A., & Andrews, D. (1988). Using microcounseling to teach RN nursing students skills of therapeutic communication. *The Journal of Nursing Education, 27*, 246–252.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113–126. <http://dx.doi.org/10.1037/0022-3514.44.1.113>
- Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The Scientific World Journal, 6*, 1146–1163. <http://dx.doi.org/10.1100/tsw.2006.221>
- Dexter, V. J. (2012). *Research synthesis with meta-analysis of empathy training studies in helping professions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3510491).
- Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology, 43*, 261–274. <http://dx.doi.org/10.1037/0022-0167.43.3.261>
- Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy, 48*, 43–49. <http://dx.doi.org/10.1037/a0022187>
- Gerdes, K. E., Lietz, C. A., & Segal, E. A. (2011). Measuring empathy in the 21st century: Development of an empathy index rooted in social cognitive neuroscience and social justice. *Social Work Research, 35*, 83–94. <http://dx.doi.org/10.1093/swr/35.2.83>
- Gerdes, K. E., & Segal, E. A. (2009). A social work model of empathy. *Advances in Social Work, 10*, 114–127. Retrieved from <http://journals.iupui.edu/index.php/advancesinsocialwork/article/view/Article/235>
- Gladstein, G. A. (1983). Understanding empathy: Integrating counseling, developmental, and social psychology perspectives. *Journal of Counseling Psychology, 30*, 467–482. <http://dx.doi.org/10.1037/0022-0167.30.4.467>
- Gresham, F. M., Sugai, G., & Horner, R. H. (2001). Interpreting outcomes of social skills training for students with high-incidence disabilities. *Exceptional Children, 67*, 331–344.
- *Handmaker, N. S., Hester, R. K., & Delaney, H. D. (1999). Videotaped training in alcohol counseling for obstetric care practitioners: A randomized controlled trial. *Obstetrics and Gynecology, 93*, 213–218. [http://dx.doi.org/10.1016/S0029-7844\(98\)00377-9](http://dx.doi.org/10.1016/S0029-7844(98)00377-9)
- Head, E. (2009). The ethics and implications of paying participants in qualitative research. *International Journal of Social Research Methodology: Theory & Practice, 12*, 335–344. <http://dx.doi.org/10.1080/13645570802246724>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic.
- Higgins, H. M. (1990). *Empathy training and stress: Their role in medical students' responses to emotional patients* (Doctoral dissertation). Retrieved from <https://circle.ubc.ca/handle/2429/30612>
- Hill, C. E., & Lent, R. W. (2006). A narrative and meta-analytic review of helping skills training: Time to revive a dormant area of inquiry. *Psychotherapy: Theory, Research, Practice, Training, 43*, 154–172. <http://dx.doi.org/10.1037/0033-3204.43.2.154>
- *Hodge, E. A., Payne, P. A., & Wheeler, D. D. (1978). Approaches to empathy training: Programmed methods versus individual supervision and professional versus peer supervisors. *Journal of Counseling Psychology, 25*, 449–453. <http://dx.doi.org/10.1037/0022-0167.25.5.449>
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511805851>
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307–316. <http://dx.doi.org/10.1037/h0027580>
- Hojat, M., Louis, D. Z., Markham, F. W., Wender, R., Rabinowitz, C., & Gonnella, J. S. (2011). Physicians' empathy and clinical outcomes for diabetic patients. *Academic Medicine, 86*, 359–364. <http://dx.doi.org/10.1097/ACM.0b013e3182086fe1>
- *Jacobs, D. (1977). *Role playing role reversal as a method for training adolescents in empathy* (Master's thesis). Retrieved from <http://www.ruor.uottawa.ca/handle/10393/8194>
- Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior, 9*, 441–476. <http://dx.doi.org/10.1016/j.avb.2003.03.001>
- Jolliffe, D., & Farrington, D. P. (2006). Examining the relationship between low empathy and bullying. *Aggressive Behavior, 32*, 540–550. <http://dx.doi.org/10.1002/ab.20154>
- Kerem, E., Fishman, N., & Josselson, R. (2001). The experience of empathy in everyday relationships: Cognitive and affective elements. *Journal of Social and Personal Relationships, 18*, 709–729. <http://dx.doi.org/10.1177/0265407501185008>
- Kohut, H. (1959). Introspection, empathy, and psychoanalysis: An examination of the relationship between mode of observation and theory. *Journal of the American Psychoanalytic Association, 7*, 459–483. <http://dx.doi.org/10.1177/000306515900700304>
- La Monica, E. L. (1981). Construct validity of an empathy instrument. *Research in Nursing & Health, 4*, 389–400. <http://dx.doi.org/10.1002/nur.4770040406>
- Lam, T. C. M., Kolomiro, K., & Alamparambil, F. C. (2011). Empathy training: Methods, evaluation practices, and validity. *Journal of Multidisciplinary Evaluation, 7*, 162–200. Retrieved from http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/314
- Lambert, J. E., & Alhassoon, O. M. (2015). Trauma-focused therapy for refugees: Meta-analytic findings. *Journal of Counseling Psychology, 62*, 28–37. <http://dx.doi.org/10.1037/cou0000048>
- Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad and ugly. *Annals of the American Academy of Political and Social Science, 587*, 69–81. <http://dx.doi.org/10.1177/0002716202250791>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage.
- *Lomis, M. J., & Baker, L. L. (1985). Microtraining of forensic psychiatric patients for empathic counseling skills. *Journal of Counseling Psychology, 32*, 84–93. <http://dx.doi.org/10.1037/0022-0167.32.1.84>
- Long, E. C. J., Angera, J. J., Carter, S. J. C., Nakamoto, M., & Kalso, M. (1999). Understanding the one you love: A longitudinal assessment of an empathy training program for couples in romantic relationships. *Family Relations: An Interdisciplinary Journal of Applied Family Studies, 48*, 235–242. <http://dx.doi.org/10.2307/585632>

- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*, 525–543. <http://dx.doi.org/10.1111/j.1467-6494.1972.tb00078.x>
- Mittenberger, R. G. (2015). *Behavior modification: Principles and procedures* (6th ed.). Boston, MA: Cengage Learning.
- Moyers, T. B., & Miller, W. R. (2013). Is low therapist empathy toxic? *Psychology of Addictive Behaviors, 27*, 878–884. <http://dx.doi.org/10.1037/a0030274>
- Newman, J. A. (1993). *Affective empathy training with senior citizens using Zazen (Zen) meditation* (Doctoral dissertation). Retrieved from <http://arizona.openrepository.com/arizona/handle/10150/186497>
- *O'Neill, M., & McMillan, T. M. (2012). Can deficits in empathy after head injury be improved by compassionate imagery? *Neuropsychological Rehabilitation, 22*, 836–851. <http://dx.doi.org/10.1080/09602011.2012.691886>
- *Pecukonis, E. V. (1990). A cognitive/affective empathy training program as a function of ego development in aggressive adolescent females. *Adolescence, 25*, 59–76.
- *Perry, M. A. (1975). Modeling and instructions in training for counselor empathy. *Journal of Counseling Psychology, 22*, 173–179. <http://dx.doi.org/10.1037/h0076713>
- Pichon, I., Boccato, G., & Saroglou, V. (2007). Nonconscious influences of religion on prosociality: A priming study. *European Journal of Social Psychology, 37*, 1032–1045. <http://dx.doi.org/10.1002/ejsp.416>
- *Reed, D. L. (1996). *Effects of a combined didactic and experiential death education/empathy training program on death anxiety and empathic ability of medical students* (Doctoral dissertation). Retrieved from <http://escholarshare.drake.edu/handle/2092/425>
- *Riess, H., Kelley, J. M., Bailey, R. W., Dunn, E. J., & Phillips, M. (2012). Empathy training for resident physicians: A randomized controlled trial of a neuroscience-informed curriculum. *Journal of General Internal Medicine, 27*, 1280–1286. <http://dx.doi.org/10.1007/s11606-012-2063-z>
- Ripley, E. B. (2006). A review of paying research participants: It's time to move beyond the ethical debate. *Journal of Empirical Research on Human Research Ethics, 1*, 9–20. <http://dx.doi.org/10.1525/jer.2006.1.4.9>
- Rogers, C. R. (1992). The necessary and sufficient conditions of therapeutic personality change. 1957. *Journal of Consulting and Clinical Psychology, 60*, 827–832. <http://dx.doi.org/10.1037/0022-006X.60.6.827>
- Rogge, R. D., Cobb, R. J., Lawrence, E., Johnson, M. D., & Bradbury, T. N. (2013). Is skills training necessary for the primary prevention of marital distress and dissolution? A 3-year experimental study of three interventions. *Journal of Consulting and Clinical Psychology, 81*, 949–961. <http://dx.doi.org/10.1037/a0034209>
- Salmon, S. (2003). Teaching empathy: The PEACE curriculum. *Reclaiming Children and Youth, 12*, 167–173. Retrieved from https://reclaimingjournal.com/sites/default/files/journal-article-pdfs/12_3_Salmon.pdf
- Shamay-Tsoory, S. G. (2011). The neural bases for empathy. *The Neuroscientist, 17*, 18–24. <http://dx.doi.org/10.1177/1073858410379268>
- Telle, N., & Pfister, H. (2012). Not only the miserable receive help: Empathy promotes prosocial behavior toward the happy. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 31*, 393–413. <http://dx.doi.org/10.1007/s12144-012-9157-y>
- Tudor, K. (2011). Understanding empathy. *Transactional Analysis Journal, 41*, 39–57. <http://dx.doi.org/10.1177/036215371104100107>
- *Uhlemann, M. R., Lea, G. W., & Stone, G. L. (1976). Effect of instructions and modeling on trainees low interpersonal-communication skills. *Journal of Counseling Psychology, 23*, 509–513. <http://dx.doi.org/10.1037/0022-0167.23.6.509>
- Vermeire, E., Hearnshaw, H., Van Royen, P., & Denekens, J. (2001). Patient adherence to treatment: Three decades of research. A comprehensive review. *Journal of Clinical Pharmacy and Therapeutics, 26*, 331–342. <http://dx.doi.org/10.1046/j.1365-2710.2001.00363.x>
- Watson, J. C., Steckley, P. L., & McMullen, E. J. (2014). The role of empathy in promoting change. *Psychotherapy Research, 24*, 286–298. <http://dx.doi.org/10.1080/10503307.2013.802823>

Received January 4, 2015

Revision received April 29, 2015

Accepted April 30, 2015 ■