Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates
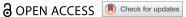
Teaching cognitive and affective empathy in medicine: a systematic review and meta-analysis of randomized controlled trials

Hazel L. Ngo, Nina Sokolovic and Jennifer M. Jenkins

Department of Applied Psychology and Human Development, University of Toronto, Toronto, Canada

ABSTRACT

Purpose: Empathy can be divided into cognitive empathy (CE) and affective empathy (AE). CE is defined as the accurate understanding and appropriate response to others' thoughts whereas AE is defined as the accurate understanding and appropriate response to others' emotions. The overall purpose of this systematic review and meta-analysis is to assess the effectiveness of empathy interventions in physicians and physicians-in-training in increasing CE and AE. Specifically, we are interested in examining whether specific teaching methods and intervention designs may contribute to greater empathy intervention effectiveness for CE and AE outcomes.

Method: Studies searched included randomized controlled trials conducted between 1971 to 2022 examining empathy interventions for medical students and physicians. Thirty-six studies, consisting of 3,833 participants, met the inclusion criteria. Data were analysed using random-effects pairwise meta-analysis.

Results: Empathy interventions have moderate effect sizes on both CE [$d = .50$ (95% CI = .30, .70)] and AE [$d = .46$ (95% CI = .30, .62)]. Heterogeneity of effects was evident for both analyses. The effectiveness of interventions on AE was moderated by measurement type. Intervention effectiveness was not significantly moderated by other intervention and study characteristics for either type of empathy.

Conclusions: There is evidence of key differences, and similarities, in how CE and AE is taught to medical students and professionals; however, the small number and high level of heterogeneity in studies makes this difficult to confirm. Research in this field will benefit from investigators standardizing teaching and research methods across studies.

PRACTICE POINTS

- Empathy interventions are moderately effective at increasing cognitive and affective empathy.
- Affective empathy appears more sensitive to variations in training methods and measurement than cognitive empathy.
- Short interventions, focused on didactic teaching and observation, that avoid public rehearsal, may be important to try.
- Future empathy intervention studies would benefit from stronger standardization and measurement of cognitive and affective empathy.

Introduction

The importance of physician empathy in physician-patient interactions is well established [1,2] Empathy contributes to better patient health outcomes [2,3], buffers against physician burnout [4,5], and is associated with greater clinical competence [6]. Given empathy's benefits, there is great interest in how to increase it in medical practice [7]. Previous meta-analyses demonstrate that empathy can be taught across various healthcare professionals, with interventions yielding moderate effect sizes [8–10]. There are, however, different types of empathy which are associated with different functions in clinical care,

and it is unclear how best to support the development of these separate types.

Empathy has been divided into two related types: *cognitive empathy* and *affective empathy* [7,11]. Cognitive empathy (CE) is the accurate understanding and appropriate response to others' thoughts whereas affective empathy (AE) is the accurate understanding and appropriate response to others' emotions. A majority of meta-analyses have shown that interventions are effective in increasing cognitive empathy to a moderate degree [8–10], while a smaller number have reported modest effect sizes on the training on emotional skills, including AE [12].

Nonetheless, several studies demonstrate the importance of CE [6,13,14] and AE [15,16] in medicine, as both have been positively associated with greater clinical competence in physicians. This makes sense, as core competencies, such as *patient care* and *interpersonal and communication skills*, depend on both CE and AE [3,17]. No studies, however, have yet examined how exactly these types of empathy are best trained in clinical settings. There are important distinctions between CE and AE which warrant separate study and perhaps separate training approaches.

First, while these types of empathy are positively correlated [18,19], CE and AE activate different neurobiological mechanisms and involve distinct brain regions [2,18,20–23]. Second, in medical practice, CE and AE can be differentially important for different tasks and interactions [7]. For example, CE is activated when the physician is curious about the patient's perspective or adjusting medical jargon to a patient to ensure understanding [24]. AE may be emphasized when facilitating emotion recognition of the patient, developing rapport, and providing reassurance as needed. Both types of skills are key to patient interactions and may be more or less emphasized depending on the context. Third, there is evidence of differential trajectories of CE and AE in medical students over time, reflecting the need to not treat empathy as a single entity, but as a complex, multifaceted one [25]. Therefore, knowing what type of empathy skills to target in a specific physicians' practice is an important consideration when designing a targeted intervention [26].

Despite the many approaches to teaching empathy in medicine, there are no meta-analyses to date which have examined whether CE and AE should be trained differently. There may be several reasons for this. One is that the measurement of empathy has not been consistent [27], making data aggregation challenging. Another is that there may be ambivalence towards the affective dimension of empathy within the medical field [7,24,28]. For instance, Sulzer et al. (2016) reported that 85% of empathy definitions in medical contexts predominantly involved, cognitive components of empathy, whereas affective components were the least prevalent. Indeed, medicine has traditionally focused on increasing cognitive empathy, despite the importance of AE in patient care [29,30]. The lack of emphasis on physician emotion in medicine has consequently translated to extremely limited knowledge on the effective methods of increasing AE in medicine and whether or not that differs from increasing CE.

Empathy interventions generally use the following five teaching methods: didactic (learning via presentations and written information), rehearsal (practicing empathy skills and/or role-playing as patient), observational (watching enactments pertaining to the

skill and/or exposure to patient's perspective), reflection (discussing the skill with others or reflecting by oneself), and feedback (receiving personalized feedback from peers or instructor) [31,32]. For example, one meta-analysis found that reading fiction had a small but significant impact on CE [33], whereas another meta-analysis found being immersed in another's perspective via virtual reality may improve AE, but not CE [34].

Beyond teaching methods, different intervention characteristics (e.g., intervention format, facilitator type, and duration), and study characteristics (i.e., measurement type, trainee education level, region, and overall risk of bias judgement) may also be associated with greater effectiveness for CE versus AE. These moderators have been investigated in meta-analyses for various adult [35,36], healthcare [37,38], and medical practitioner [9,39] samples, however, none have examined how these characteristics may separately impact the training of CE and AE. This study represents the first systematic analysis to go beyond treating empathy as a singular, general construct, instead adopting a multidimensional perspective. By doing so, it aims to identify the most effective strategies for enhancing distinct empathy components, offering a more nuanced understanding of empathy training. Thus, the purpose of this systematic review and meta-analysis is to evaluate the effect of empathy interventions in physicians and physicians-in-training, specifically examining which teaching methods and intervention designs may significantly impact CE and AE outcomes, respectively.

Method

Search strategy

A protocol for this study was developed and registered with the International Prospective Register of Systematic Reviews (Protocol number: CRD42018100100). The current study was part of a larger search to assess effective methods for teaching adults empathy across different populations. Other papers that derived from this search, but do not overlap, include a meta-analysis for mental health practitioners [38] and parents [36]. The current paper focuses on training individuals who are currently in, or who have completed, physician training. The search strategy was developed in consultation with librarians using the Peer Review of Electronic Search Strategies checklist. The search was conducted in PsycINFO, Medline, CINAHL, Social Work Abstracts, ERIC, ABI/INFORM, and the Cochrane Central Register of Controlled Trials from the inception of the respective databases to 9 October 2022. Reference lists of both included studies and relevant review articles retrieved in the search were also

scanned for additional studies. Grey and non-peer reviewed literature were included (e.g., dissertations, reports). For full search strategy, see Section 1 of the Supplemental Information. For full references of included studies, see Section 2.

Inclusion and exclusion criteria

Studies that were included met the following criteria: (1) randomized trial of a behavioural training targeting an empathy-related skill, (2) sample is physicians (i.e., professionals) or physicians-in-training (i.e., currently in undergraduate or postgraduate medical school or residency), (3) reports on at least one quantitative measure of cognitive and/or affective empathy, and (4) the study is available in English (due to coding capacities). Studies were excluded if there were no outcomes for AE and/or CE (see Data Extraction section below for more details). To control for non-independent sampling between studies, we screened studies to see if any shared the same sample. If there was overlap, but the studies reported different outcomes within the same category (i.e., CE or AE), we followed the common practice of creating the mean of the two effects $(k = 2)$ [40]. For studies in which summary data could not be converted into an effect size, we contacted the author if an e-mail address was available. If the author did not respond or were unable to be contacted (i.e., the provided e-mail no longer worked), the study was excluded.

Data extraction

All coders were advanced undergraduate and graduate student research assistants who were trained by the first author (HN) on study selection and data extraction. Assistants had to obtain an interrater reliability of at least 80% (percent agreement) on a sample of articles to pass the training. Independent pairs of coders screened abstracts and full-texts, and assessed for study quality [41] (see section below for details). Study outcomes (i.e., CE and AE) were determined before data extraction (see Table 1 for examples of items classified as CE or AE, respectively). Coders extracted outcomes and moderators using a predefined data extraction sheet (for full definitions of the study variables, see Section 3 of the Supplemental Information). All data were double coded with discrepancies resolved by

a consensus process involving the first author. This ensured consistency across the final dataset. No recurring or unresolved disagreements arose that would suggest systematic misinterpretation from the coders.

Outcomes were extracted for each type of empathy if available. In order for a measure to be extracted, at least 50% of the measure had to have items pertaining to either CE and AE (Table 1). If the scale did not meet this criteria, we extracted at the subscale level, and if the measure's subscales did not meet the criteria, we extracted at the item level. When multiple measures of empathy were reported for the same outcome (e.g., two measures of CE), we selected in the following order: objective measures (i.e., ability tests and/or observational measures), other-reported, first-person measures, because objective measures of empathy have been shown to be more predictive of client outcomes [42] and less biased [43] than clinician self-reports. If multiple measures were of the same methodological quality (e.g., multiple *objective* measures of CE), a pooled effect size was calculated across these measures to obtain one effect size per comparison without artificially deflating the standard error [44]. See Section 4 of the Supplemental Information for details on extracted measures.

Cochrane quality assessment

As part of standard practice, we used the Cochrane risk-of-bias tool [41] to assess study quality. This tool is focused on rating studies on methodological features across categories such as randomization, allocation, blinding of participants, and outcome measurement. Based on these criteria, a study given a rating of *high, low*, or *unclear* risk for each of the seven Cochrane categories (see Section 4 of the Supplemental Information, Figures 4.1 and 4.2). Overall risk of bias judgement was determined in the following manner. Studies were labelled 'low risk' if four or more of the seven Cochrane categories were determined to be *low risk*. Studies were labelled 'high risk' if four or more of the seven Cochrane categories were determined to be *high* or *unclear risk*. The assessment of whether a study is ultimately judged as 'high risk' or 'low risk' reflects the likelihood that bias may influence the results (i.e., the effects of an intervention may be underestimated or overestimated).

Table 1. Cognitive and affective empathy example skill classification.

Cognitive Empathy Skills Items	Affective Empathy Skills Items
Clarified questions as needed	Recognized patient's emotions
Checked for patient understanding	Appropriately provided reassurance
Accurately summarized what patient said	Attuned to the patient's body language
Avoided using technical language	Accurately reflected the patient's feelings
Actively listened to the patient's perspective	Made the patient feel comfortable

Statistical analysis

Comprehensive Meta-Analysis (CMA; Biostat, New Jersey, United States of America) software was used to calculate a standardized mean difference (SMD; Cohen's d) and standard error (SE) for each comparison. For studies in which two of the arms were functionally equivalent (i.e., the arms used the same teaching methods and format of delivery, such as lecture and written materials that are both didactic), the mean effect size was calculated for those two intervention arms relative to the other arm(s) in the study ($k = 2$). For studies with more than one functionally non-equivalent intervention (i.e., different combinations of teaching methods), we only used the comparison between the study arm with the greatest number of teaching methods (i.e., most intensive) and the control group [36] ($k = 2$). Pairwise analyses were separately conducted for CE outcomes and AE outcomes. A random-effects model was chosen as study populations were assumed to differ [45]. Outliers were defined as those with an SMD three standard deviations above or below the mean. Statistical heterogeneity was assessed with I^2

the statistic and p value and explored using mixed-effects model meta-regressions using method of moments estimation [39]. Publication bias was assessed by inspecting funnel plots and running Egger's test [41]. Statistical significance was determined at $p < .05$ and marginal significance at $p < .10$.

Results

As outlined in the PRISMA diagram (Figure 1), 12,002 citations were screened and 1,407 of those articles were further assessed for eligibility. Thirty-six studies, with 3,833 participants met criteria for inclusion. Of the 36 studies extracted, 13 studies (36%) reported both CE and AE outcomes, 20 studies reported CE outcomes, and 29 studies reported AE outcomes. Studies which reported a 'high' overall risk of bias judgement are as follows: 50% of studies with CE outcomes and 48% of studies with AE outcomes. Studies' publication dates ranged from 1971 to 2022, though most (61%) were published between 2009 and 2019 inclusive. Only 58% of studies reported on participants' gender identities; of these studies, most
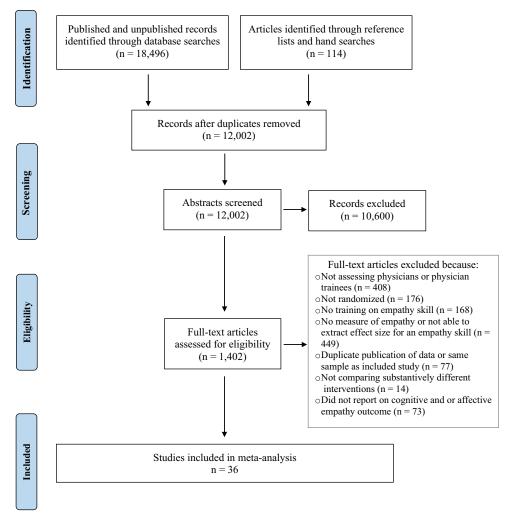

Figure 1. PRISMA flow diagram.

samples were approximately 50% female. Most studies did not report on the ethnicity of participants (75%); however, those that did report had primarily European American samples. In terms of education level, 61.1% represented medical students, 33.3% represented physicians, and 5.6% represented a mixed sample of medical students and physicians (see Section 4 of the Supplemental Information). No outliers were identified.

The CE analysis revealed an overall average effect size of $d = .50$ (95% CI = .30, .70). Comparably, the AE analysis revealed an overall average effect size of $d = .46$ (95% CI = .30, .62). These effect sizes are considered moderate [46]. Publication bias was not evident in the CE analysis but was found in the AE analysis. Further assessment of publication bias can be found in Section 5 of Supplemental Information. Since there was evidence of heterogeneity for both CE ($I^2 = 73.67$, $p < .001$) and AE ($I^2 = 66.19$, $p < .001$),

a series of meta-regressions were run to test whether this could be explained by individual teaching methods, intervention characteristics, or study characteristics.

The results of univariate meta-regressions are presented in Table 2. For CE outcomes, it was noted that interventions using the observation teaching method were marginally more effective ($d = .72$, 95% CI: .37, 1.06), on average, than those that did not use this method ($d = .32$, 95% CI: .14, .49). No differences were noted for other teaching methods (didactic, reflection, and feedback). Effects of intervention characteristics for CE outcomes did not vary by intervention format or duration. Other study characteristics, including measurement type, education level, region, or overall risk of bias judgement, were not found to be significant. There were not enough studies to examine effects of rehearsal teaching method and facilitator type.

Table 2. Pairwise random-effects meta-analyses.

	Cognitive Empathy						Affective Empathy					
	k	d	95% CI	Q	p	R^2	k	d	95% CI	Q	p	R^2
Teaching Methods												
Didactic												
Yes	15	.61	.35, .87				24	.54	.35, .74			
No	5	.29	.02, .56	2.01	.16	.00	5	.20	.03, .38	2.96	.08	.00
Rehearsal[c]												
Yes	–	–	–				23	.38	.21, .54			
No	–	–	–	–	–	–	6	.78	.32, 1.23	3.18	.07	.00
Reflection												
Yes	12	.62	.35, .90				17	.41	.20, .62			
No	8	.34	.05, .62	1.66	.20	.00	12	.53	.28, .77	.51	.48	.00
Observation												
Yes	10	.72	.37, 1.06				14	.65	.37, .94			
No	10	.32	.14, .49	3.06	.08	.04	15	.32	.14, .49	3.65	.06	.06
Feedback												
Yes	12	.53	.21, .85				19	.50	.31, .70			
No	8	.48	.22, .73	.01	.94	.00	10	.40	.15, .65	.48	.49	.04
Intervention Characteristics												
Format[a,b]												
Group	13	.61	.30, .93				18	.62	.36, .88			
Group + Individual	4	.34	−.06, .74	.85	.36	.00	5	.27	.03, .51			
Online/Independent	–	–	–				5	.24	.05, .42	3.58	.17	.00
Facilitator[a, c]												
Professional	–	–	–				20	.53	.33, .74			
Researcher	–	–	–				3	.73	.07, 1.38			
Other	–	–	–	–	–	–	5	.24	.05, .42	2.52	.28	.00
Number of Sessions[a,b]												
<5 Sessions	13	.62	.42, .82				21	.55	.36, .75			
5–10 Sessions	5	.32	−.20, .84	2.15	.14	.09	6	.19	−.15, .54	3.09	.08	.03
Study Characteristics												
Measurement Type[a]												
Objective	14	.60	.33, .98				20	.51	.32, .70			
Other-report	6	.41	.23, .73	.70	.40	.00	6	.19	.07, .31	4.65	.03	.22
Education Level[a,b]												
Professional	6	.79	.32, 1.27				10	.54	.24, .84			
Resident	–	–	–				5	.73	.04, 1.41			
Postgraduate	10	.46	.22, .70	1.34	1.66	.00	10	.40	.19, .61	1.07	.59	.00
Region[a,b]												
North America	8	.44	.28, .60				11	.53	.39, .76			
Europe	8	.47	−.01, .95				10	.33	.10, .56			
East Asia & Pacific	3	.79	.20, 1.40	1.40	.50	.00	5	.28	−.12, .67	1.84	.40	.00
Overall Risk of Bias Judgement												
High	10	.55	.21, .89				14	.52	.20, .84			
Low	10	.47	.23, .72	.04	.83	.00	15	.37	.22, .53	.22	.64	.00

[a]Pairwise analyses excluded categories when there were less than three studies in the AE analysis.
[b]Pairwise analyses excluded categories when there were less than three studies in the CE analysis.
[c]Pairwise analyses could not be run as only one category had more than three studies.

For AE outcomes, using didactic teaching methods were marginally more effective $(d = .54$, 95% CI: .35, .74), on average, than those that did not include them $(d = .20$, 95% CI: .03, .38). This was also true for observation teaching methods $(d = .65$, 95% CI: .37, .94 versus $d = .32$, 95% CI: .14, .49). In contrast, interventions using rehearsal teaching methods were marginally *less* effective $(d = .20$, 95% CI: .21, .54), on average, than those that did not include them $(d = .78$, 95% CI: .32, 1.23). No differences were found for reflection and feedback teaching methods.

In terms of intervention characteristics for AE outcomes, format and facilitator type were not related to intervention effects, but interventions with fewer than five sessions were marginally more effective $(d = .55$, 95% CI: .36, .75), on average, than those with more than five sessions $(d = .19$, 95% CI: −.15, .54). When examining the impact of study characteristics, a significant effect of measurement type was found. Larger intervention effects were found in studies using objective measures to measure AE $(d = .54$, CI: 95% CI: .34, .75) compared to those using other-report measures $(d = .19$, CI: 95% CI: .07, .31). No significant or marginal differences were found for other study characteristics, including education level, region, or overall risk of bias judgement. For individual study effect sizes, forest plot, and risk of bias judgements, see Supplemental Information Section 4 (Figures 4.1 and 4.2 for CE and AE outcomes, respectively).

Discussion

The goals of this paper were to explore the most effective teaching methods and intervention characteristics for increasing CE and AE in medical students and physicians. Despite the importance of using both CE and AE skills in patient interactions [7,30], it has been unclear how effective empathy interventions are at increasing CE and AE as separate outcomes. Overall, a moderate effect size was found for increasing both CE and AE. The effect sizes reported in this paper were similar in magnitude to meta-analyses that have not differentiated CE and AE [9,10]. Moderator effects were found to be small in magnitude and evident more for affective than cognitive empathy.

When examining whether effectiveness differed based on the use of specific teaching methods, results revealed no significant moderation for either CE or AE outcomes. There were, however, indications (based on marginally significant effects) that both types of empathy might benefit from observation and didactic teaching methods. The importance of didactic teaching has been shown for general empathy in a previous meta analysis [38]. Didactic teaching helps to build the surface knowledge essential for being able to acquire and enact the skill for practice [47]. Observational learning in single studies [48], and in previous meta-analysis [49], has also been shown to be an effective teaching tool for complex skills such as interpersonal skills.

Interestingly, the inclusion of rehearsal teaching methods in interventions tended to *decrease* AE, which is unexpected given the importance of skill practice in learning [50]. It is possible this pattern reflects performance anxiety: when physicians are asked to demonstrate skills in front of others as part of their training, this may negatively impact their skill learning [51]. Anxiety aroused in learning may particularly impact AE as there is a robust association between AE and emotion regulation [19]. However, this finding should be interpreted with caution, as it was at trend level and needs a larger number of studies to ensure accuracy.

In terms of intervention characteristics, there was a marginally significant effect which showed that interventions with *less than five* sessions tended to be more effective, on average, than interventions with *more than five* sessions for AE. This aligns with prior research which shows that focused interventions are more effective than long interventions when teaching empathy-related skills to adults [52]. This effect of duration was not evident when examining CE. Other intervention characteristics (i.e., format, facilitator type) were not found to be significant for CE nor AE.

Lastly, the only study characteristic that was associated with intervention effectiveness was measurement type for AE outcomes. Studies using objective measures had, on average, greater effect sizes than those using other-report measures. This could mean that objective measures are better measures of affective empathy (i.e., are more valid and reliable). Questionnaire measures have been shown to be influenced by various response biases which can inflate associations [53,54].

Several limitations and future directions should be considered. First, small cell sizes are evident for some moderators, resulting in low power for these analyses. Linked to this, several moderator analyses omitted certain category comparisons because of insufficient studies (i.e., online format, facilitator type, rehearsal teaching methods). This power issue indicates the need for more single intervention studies that differentiate CE and AE, as well as a future meta-analysis to inform training initiatives. Second, a vast majority of the AE and CE measures collected in the present study were at the subscale or item level, rather than scale level. Greater attention to the development of rigorous, standardized CE and AE measurement is an important next step.

Third, most descriptions of empathy interventions are lacking detail. Future studies should report the relative focus on AE and CE skills in the intervention, as well as more details relating to the intervention. For example, the variable 'number of sessions', was only partially able to capture the

exact 'dose' of the intervention. It is possible that one session could equate to 3 hours in one study and 30 minutes in another study. However, because a vast majority of studies did not report at this level of detail (i.e., time spent per activity), a more precise coding of this important construct was not possible. Fourth, while we were able to capture an aspect of cultural diversity in CE and AE by examining the region of the study, there were not enough studies to examine certain regions (i.e., Sub-Saharan Africa, Middle East and North Africa). It is a limitation that most studies were from North America and Europe, which mitigates against being able to examine cultural differences in the training of empathy more accurately [55]. There is a strong need for more studies from diverse cultural, ethnic, and linguistic backgrounds.

Despite its limitations, this study is the first to systematically evaluate the effectiveness of interventions when it comes to enhancing different facets of empathy. The results provide greater understanding of what constitutes effective empathy interventions in the medical context. When it comes to developing either CE and AE skills, the present study is aligned with other meta-analyses of empathy training, indicating certain aspects of training to be effective, such as didactic and observational training methods [36,49]. However, while previous meta-analyses on empathy interventions in healthcare [8] and medical students [9] showed that rehearsal training methods were effective, we found the opposite trend when it comes to developing AE skills. Additionally, unlike previous medical education research which has either treated empathy as a general concept [37,56] or focused solely on CE while overlooking affective empathy (AE) [10], the current meta-analysis shows that both CE and AE can be trained to a moderate degree. Marginal and significant effects of moderators suggest that affective empathy is more sensitive to variations in training methods and measurement than cognitive empathy.

In terms of educational practices, our research provides valuable information on what is more, or less, effective when it comes to creating interventions to increase CE and AE in physicians. The following recommendations for medical educators include: (1) incorporating didactic and observational methods into training, while noting rehearsal teaching methods may negatively impact AE; (2) all education levels are comparably trainable in both CE and AE empathy; (3) expensive intervention components, such as feedback teaching methods and the individual format, may not be necessary for effective empathy interventions; and (4) after five sessions of training, the length of intervention may be comparably effective, or even less effective, and (5) using objective measures of empathy may yield higher intervention effect sizes,

at least when measuring AE. Overall, the results of the study should encourage medical educators that teaching empathy is effective and that most of the key interventions components are largely similar between CE and AE.

In terms of future research, this study emphasizes the need for more research to examine the multidimensional skills that constitute empathy, rather than just CE. For example, there is very limited knowledge on how patient outcomes are impacted as a function of the type of empathy that physicians demonstrate [57]. That is, what type of empathy is most effective in certain clinical contexts? Most of the studies reviewed either took place in academic settings or did not involve patients (27/36 studies, 75%). An important area of study will include understanding whether certain measurements of empathy are better predictors of patient outcome in real-world settings. Part of this endeavor will involve studying empathy in ecologically-valid contexts, with actual patients, to truly test the effectiveness of empathy interventions.

In order to advance medical education overall, future research must be strengthened and guided by clear conceptualization and the rigorous measurement of empathy. For example, developing a standardized, objective measure of empathy to be incorporated into the OSCE examination may help further the study of empathy in a standardized manner, within a common clinical training context. It is ideal that the medical education field finds greater consensus on viewing empathy as a clinical competence, which can encourage greater adherence to high-quality, well-validated empathy measures. This would alleviate some issues of measurement heterogeneity.

It should of course be acknowledged that understanding influences on the development and expression of highly complex interpersonal skills, such as empathy, needs the integration of multiple paradigms and methods [58]. This paper uses quantitative methods within a positivist frame, firstly because of our own training and secondly because of the research goal (replicable and generalizable knowledge, across populations). We also work in practice and policy in a diverse metropolitan area. We adapt findings generated by quantitative methods for the settings and populations that we serve. This involves a process of co-construction and to achieve this we draw on qualitative and descriptive studies [59,60]. We operated within the implementation science framework which argues that to improve population health the integration of paradigms and mixed methods is essential [61]. While challenges persist in evaluating empathy through structured research designs, we maintain that rigorous investigation remains essential. Advancing our understanding of empathy interventions and their effectiveness is critical to supporting physician development and, ultimately, improving patient care.

Disclosure statement

Funding

References

[1] Kelley JM, Kraft-Todd G, Schapira L, et al. The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. PLOS ONE. 2014;9(4): e94207. doi: 10.1371/JOURNAL.PONE.0094207

[2] Decety J, Fotopoulou A. Why empathy has a beneficial impact on others in medicine: unifying theories. Front Behav Neurosci. 2014;8:457. doi: 10.3389/FNBEH.2014.00457

[3] Derksen F, Bensing J, Lagro-Janssen A. Effectiveness of empathy in general practice: a systematic review. Br J Gen Pract. 2013;63(606):e76–e84. doi: 10.3399/BJGP13X660814

[4] Thirioux B, Birault F, Jaafari N. Empathy is a protective factor of burnout in physicians: new neuro-phenomenological hypotheses regarding empathy and sympathy in care relationship. Front Psychol. 2016;7(MAY). doi: 10.3389/FPSYG.2016.00763

[5] Wilkinson H, Whittington R, Perry L, et al. Examining the relationship between burnout and empathy in healthcare professionals: a systematic review. Burnout Res. 2017;6:18–29. doi: 10.1016/j.burn.2017.06.003

[6] Ogle J, Bushnell JA, Caputi P. Empathy is related to clinical competence in medical care. Med Educ. 2013;47(8):824–831. doi: 10.1111/MEDU.12232

[7] Decety J. Empathy in medicine: what it is, and how much we really need it. Am J Med. 2020;133 (5):561–566. doi: 10.1016/j.amjmed.2019.12.012

[8] Chua JYX, Ang E, Lau STL, et al. Effectiveness of simulation-based interventions at improving empathy among healthcare students: a systematic review and meta-analysis. Nurse Educ Today. 2021;104:105000. doi: 10.1016/J.NEDT.2021.105000

[9] Fragkos KC, Crampton PES. The effectiveness of teaching clinical empathy to medical students: a systematic review and meta-analysis of randomized controlled trials. Acad Med. 2020;95(6):947–957. doi: 10.1097/ACM.0000000000003058

[10] Paulus CM, Meinken S. The effectiveness of empathy training in health care: a meta-analysis of training content and methods. Int J Med Educ. 2022;13:1. doi: 10.5116/IJME.61D4.4216

[11] Håkansson Eklund J, Summer Meranius M. Toward a consensus on the nature of empathy: a review of reviews. Patient Educ Couns. 2021;104(2):300–307. doi: 10.1016/J.PEC.2020.08.022

[12] Satterfield JM, Hughes E. Emotion skills training for medical students: a systematic review. Med Educ. 2007;41(10):935–941. doi: 10.1111/J.1365-2923.2007.02835.X

[13] Hojat M, Gonnella JS, Mangione S, et al. Empathy in medical students as related to academic performance, clinical competence and gender. Med Educ. 2002;36 (6):522–527. doi: 10.1046/J.1365-2923.2002.01234.X

[14] Casas RS, Xuan Z, Jackson AH, et al. Associations of medical student empathy with clinical competence. Patient Educ Couns. 2017;100(4):742–747. doi: 10.1016/J.PEC.2016.11.006

[15] Jang SS, Seo JH, Cho GJ, et al. Correlation of communication skills for emotional empathy and academic achievement on clinical performance examinations. Korean J Med Educ. 2010;22 (2):121–130. doi: 10.3946/KJME.2010.22.2.121

[16] Arora S, Ashrafian H, Davis R, et al. Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies. Med Educ. 2010;44(8):749–764. doi: 10.1111/J.1365-2923.2010.03709.X

[17] Moudatsou M, Stavropoulou A, Philalithis A, et al. The role of empathy in health and social care professionals. Healthcare. 2020;8(1):26. doi: 10.3390/healthcare8010026

[18] Cox CL, Uddin LQ, Di Martino A, et al. The balance between feeling and knowing: affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. Soc Cogn Affect Neurosci. 2012;7 (6):727–737. doi: 10.1093/SCAN/NSR051

[19] Thompson NM, van Reekum CM, Chakrabarti B. Cognitive and affective empathy relate differentially to emotion regulation. Affec Sci. 2022;3(1):118–134. doi: 10.1007/S42761-021-00062-W

[20] Healey ML, Grossman M. Cognitive and affective perspective-taking: evidence for shared and dissociable anatomical substrates. Front Neurol. 2018;9 (JUN). doi: 10.3389/FNEUR.2018.00491

[21] Walter H. Social cognitive neuroscience of empathy: concepts, circuits, and genes. Emot Rev. 2012;4 (1):9–17. doi: 10.1177/1754073911421379

[22] Yu CL, Chou TL. A dual route model of empathy: a neurobiological prospective. Front Psychol. 2018;9 (NOV):2212. doi: 10.3389/FPSYG.2018.02212

[23] McDonald KR, Pearson JM, Huettel SA. Dorsolateral and dorsomedial prefrontal cortex track distinct properties of dynamic social behavior. Soc Cogn Affect Neurosci. 2020;15(4):383–393. doi: 10.1093/SCAN/NSAA053

[24] Halpern J. From idealized clinical empathy to empathic communication in medical care. Med Health Care And Philos. 2014;17(2):301–311. doi: 10.1007/S11019-013-9510-4

[25] Smith KE, Norman GJ, Decety J. The complexity of empathy during medical school training: evidence for positive changes. Med Educ. 2017;51(11):1146–1159. doi: 10.1111/MEDU.13398

[26] Zhou YC, Tan SR, Tan CGH, et al. A systematic scoping review of approaches to teaching and assessing empathy in medicine. BMC Med Educ. 2021;21 (1):1–15. doi: 10.1186/S12909-021-02697-6

[27] Sulzer SH, Feinstein NW, Wendland CL. Assessing empathy development in medical education: a systematic review. Med Educ. 2016;50(3):300–310. doi: 10.1111/MEDU.12806

[28] Hojat M, Maio V, Pohl CA, et al. Clinical empathy: definition, measurement, correlates, group differences, erosion, enhancement, and healthcare outcomes. Discov Heal Syst. 2023;2(1):1–17. doi: 10.1007/S44250-023-00020-2

[29] Guidi C, Traversa C. Empathy in patient care: from 'clinical empathy' to 'empathic concern'. Med Health

Care And Philos. 2021;24(4):573–585. doi: 10.1007/s11019-021-10033-4

[30] Halpern J. What is clinical empathy? J Gen Intern Med. 2003;18(8):670–674. doi: 10.1046/j.1525-1497.2003.21017.x

[31] Davis MH, Begovic E. Empathy-related interventions. In: John Wiley & Sons, Ltd; 2014. p. 111–134. doi: 10.1002/9781118315927.CH6

[32] Hattie J. Visible learning : a synthesis of over 800 meta-analyses relating to achievement. Routledge; 2008. doi: 10.4324/9780203887332

[33] Dodell-Feder D, Tamir DI. Fiction reading has a small positive impact on social cognition: a meta-analysis. J Exp Psychol Gen. 2018;147(11):1713–1727. doi: 10.1037/XGE0000395

[34] Martingano AJ, Hererra F, Konrath S. Virtual reality improves emotional but not cognitive empathy: a meta-analysis. Technol Mind, Behav. 2021;2(1). doi: 10.1037/TMB0000034

[35] Teding van Berkhout E, Malouff JM. The efficacy of empathy training: a meta-analysis of randomized controlled trials. J Couns Psychol. 2016;63(1):32–41. doi: 10.1037/cou0000093

[36] Sokolovic N, Rodrigues M, Tricco AC, et al. Teaching parents to be responsive: a network meta-analysis. Pediatrics. 2021;148(2):e2020033563. doi: 10.1542/peds.2020-033563

[37] Kiosses VN, Karathanos VT, Tatsioni A. Empathy promoting interventions for health professionals: a systematic review of RCTs. J Compassionate Heal Care. 2016;3(1):1–22. doi: 10.1186/S40639-016-0024-9

[38] Ngo H, Sokolovic N, Coleman A, et al. Teaching empathy to mental health practitioners and trainees: pairwise and network meta-analyses. J Consult Clin Psychol. 2022;90 (11):851–860. doi: 10.1037/CCP0000773

[39] Berkhof M, van Rijssen HJ, Schellart AJM, et al. Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews. Patient Educ Couns. 2011;84(2):152–162. doi: 10.1016/J.PEC.2010.06.010

[40] McGowan J, Sampson M, Salzwedel DM, et al. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40–46. doi: 10.1016/J.JCLINEPI.2016.01.021

[41] Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions: Cochrane book series. John Wiley and Sons; 2008. doi: 10.1002/9780470712184

[42] Elliott R, Bohart AC, Watson JC, et al. Empathy. Psychotherapy. 2011;48(1):43–49. doi: 10.1037/a0022187

[43] Hong H, Han A. A systematic review on empathy measurement tools for care professionals. Educ Gerontol. 2020;46(2):72–83. doi: 10.1080/03601277.2020.1712058

[44] Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to meta-analysis. John Wiley and Sons; 2009. doi: 10.1002/9780470743386

[45] Hedges LV, Borenstein M, Higgins JPT, et al. Multiple Outcomes or Time-Points within a Stud. In: Introduction to meta-analysis. Chichester, UK: John Wiley & Sons; 2011.

[46] Cohen J. Statistical Power Analysis for the Behavioral Sciences. In: Statistical Power Analysis for the Behavioral Sciences. United Kingdom: Taylor & Francis Group; 1988.

[47] Hattie JAC, Donoghue GM. Learning strategies: a synthesis and conceptual model. Npj Sci Learn. 2016;1(1):1–13. doi: 10.1038/npjscilearn.2016.13

[48] Anderson T, Perlman MR, McCarrick SM, et al. Modeling therapist responses with structured practice enhances facilitative interpersonal skills J Clin Psychol. 2020;76(4):659–675. doi: 10.1002/jclp.22911

[49] Hill CE, Lent RW. A narrative and meta-analytic review of helping skills training: time to revive a dormant area of inquiry. Psychotherapy: Theory, Res, Pract, Train. 2006;43(2):154–172. doi: 10.1037/0033-3204.43.2.154

[50] Kolb DA, Boyatzis RE, Mainemelis C. Experiential learning theory: previous research and new directions. In: Perspectives on thinking, learning, and cognitive styles. Taylor and Francis; 2001. p. 227–247. doi: 10.4324/9781410605986-9

[51] Allen V, MacCann C, Matthews G, et al. Regulating emotions related to testing. Routledge Handbooks. Published online 2014 Apr 16:348–367. doi: 10.4324/9780203148211-20

[52] Bakermans-Kranenburg MJ, Van IJzendoorn MH, Juffer F. Less is more: meta-analyses of sensitivity and attachment interventions in early childhood. Psychol Bull. 2003;129(2):195–215. doi: 10.1037/0033-2909.129.2.195

[53] Wetzel E, Böhnke JR, Brown A. Response biases. ITC Int Handb Test Assess. [cited 2016 Jun]:349–363. doi: 10.1093/MED:PSYCH/9780199356942.003.0024

[54] Runze J, van IJzendoorn MH. Response bias is genetically biased: another argument for Kagan's philippic against questionnaires in developmental psychology. Dev Psychol. 2024;60(11):2008–2015. doi: 10.1037/DEV0001614

[55] Jami PY, Walker DI, Mansouri B. Interaction of empathy and culture: a review. Curr Psychol. 2024;43 (4):2965–2980. doi: 10.1007/S12144-023-04422-6

[56] Batt-Rawden SA, Chisolm MS, Anton B, et al. Teaching empathy to medical students: an updated, systematic review. Acad Med. 2013;88(8):1171–1177. doi: 10.1097/ACM.0b013e318299f3e3

[57] Lelorain S, Gehenne L, Christophe V, et al. The association of physician empathy with cancer patient outcomes: a meta-analysis. Psychooncology. 2023;32 (4):506–515. doi: 10.1002/PON.6108

[58] Omodan BI. Research paradigms and their methodological alignment in social sciences: a practical guide for researchers. Routledge; 2024 [cited 2025 Apr 25]. Available from: https://www.routledge.com/Research-Paradigms-and-Their-Methodological-Alignment-in-Social-Sciences-A-Practical-Guide-for-Researchers/Omodan/p/book/9781032776248

[59] Barak LC, Kuijpers G, Hoeijmakers L, et al. Clinical empathy training: an explorative qualitative study in search of the barriers and facilitators of implementation. Published online 2022 Mar 7. doi: 10.21203/RS.3.RS-1357849/V1

[60] Brodahl KØ, Finset A, Storøy HLE, et al. Medical students' expressions of empathy: a qualitative study of verbal interactions with patients expressing emotional issues in a medical interview. Patient Educ Couns. 2021;104(12):2936–2943. doi: 10.1016/J.PEC.2021.03.042

[61] Weiner BJ, Lewis CC, Sherr KH. Practical implementation science : moving evidence into action. (Weiner BJ, Lewis CC, Sherr KH, eds.). Springer Publishing Company, Incorporated; 2023 [cited 2025 Apr 25]. Available from: https://books.google.com/books/about/Practical_Implementation_Science.html?id=eRe0zgEACAAJ