Critically Appraised Topic

Diversity Training

1.   Background

Atrain approached the Center for Evidence-Based Management (CEBMa) to undertake a review to understand what is known in the scientific literature about the effect of diversity training for enhancing diversity and inclusion in the workplace. This CAT will present an overview of the best available evidence.

2.   Main questions: What will the review answer?

1. *What is diversity training?*

2. *How is diversity training supposed to support diversity and inclusion?*

3. *What is the effect of diversity training on diversity and inclusion?*

3.   Search process: How was the research evidence sought?

The following three databases were used to identify studies: ABI/INFORM Global from ProQuest and PsycINFO from Ovid. The following generic search filters are applied to all databases during the search:

1. Scholarly journals, peer-reviewed
2. Published in the period 1980 to 2017
3. Articles in English

A search was conducted using combinations of different search terms. We conducted 12 different search queries and screened the titles and abstracts of 29 studies. An overview of all search terms and queries is provided in Annex I.

4.   Selection process: How were the studies selected?

Study selection has taken place in two phases. Firstly, the titles and abstracts of the studies identified were screened for their relevance to the review question. In case of doubt or lack of information, the study was included. Duplicate publications were removed. This first phase yielded 12 studies. Secondly, studies were selected based on the full text of the article according to the following inclusion criteria.

1. Type of studies: Focusing on quantitative, empirical studies.

2. Measurement: Only studies in which the impact on diversity and inclusion was measured.

3. Context: Only studies related to workplace settings.

4. Level of trustworthiness: Only studies that were graded level D or above (see below).

This second phase yielded 9 secondary studies (meta-analyses) An overview of the selection process is provided in Annex II.

5.   Data extraction: What data were extracted?

Data extraction involved the collation of the results of the studies included. From each study, information relevant to the review question was extracted, such as year of publication, research design, sample size, population (e.g., industry, type of employees), possible moderators or mediators, main findings, effect sizes, and limitations.

6.   Critical appraisal: How was the quality of the studies included judged?

The classification system of Shadish, Cook and Campbell (2002), and Petticrew and Roberts (2006) was used to determine the methodological appropriateness of the research design of the studies included on the basis of a systematic assessment.

To determine the magnitude of an effect Cohen's rule of thumb (Cohen, 1988) was applied. According to Cohen (1988) a 'small' effect is an effect that is only visible through careful examination. A 'medium' effect, however, is one that is 'visible to the naked eye of the careful observer'. Finally, a 'large' effect is an effect that anyone can easily see because it is substantial.

7.   Outcome of the critical appraisal

The overall quality of the studies included was high. Most of the meta-analyses were based on controlled studies, and where therefore qualified as level A or higher.

8.   Main findings

Question 1: What is diversity training?

Diversity training has been defined as "instructional programs aimed at facilitating positive intergroup interactions, reducing prejudice and discrimination, and enhancing the skills, knowledge, and motivation of participants to interact with diverse others" (Berzukova et al., 2016). Put differently, it aims to reduce people's bias, prejudice and discrimination against others. The main goal of diversity training is to create a positive, inclusive work environment by helping employees to recognize and be tolerant of differences among co-workers. The training can target different types of "diversity", such as gender, race, ethnicity, demographic background, ability, health status (i.e. disability), and sexual orientation. It can be divided into awareness-based and skill-based training. Awareness-based training makes employees conscious of their biases,

prejudices, and cultural assumptions with respect to minorities. This training often uses case studies and experiential exercises.

Skill-based training develops employees' proficiency in handling diversity in the workplace. Various tools are used to improve employees' interpretations of cross-cultural differences, communication with people from different cultures, and adaptability.

Question 2: How is diversity training supposed to work?

The notion that diversity training reduces prejudice and discrimination against others is based on two theories: social identity theory (Tajfel, 1982) and the contact hypothesis (Allport, 1954). Social identity theory posits that people classify themselves and others into social categories, according to groups such as race, ethnicity, and gender. This classification shapes the way individuals interact with others. According to this theory people wish to belong to groups that enjoy distinct and positive identities, and therefore individuals who belong to groups with greater perceived social status will accept and include people they consider to be like them, while excluding those they perceive to be different. The contact hypothesis assumes that positive contact between members of different groups reduces prejudice, raises awareness, and changes group interactions. Put differently, if members of different groups have the opportunity to communicate with each other, they will learn to understand and appreciate different points of views, cultural habits, etc., and as a result prejudice should diminish.

Measuring the effect of (diversity) training

Researchers have grouped outcomes of training into cognitive, behavioral and attitudinal learning, as well as on trainee reactions. In the case of diversity training, cognitive learning refers to the extent to which trainees acquire knowledge about other cultures and problems or issues amongst different groups. Behavioral learning concerns the development of skills and behaviors. Attitudinal learning refers to the development of trainees' attitudes towards diversity.

Question 3: What is the effectiveness of diversity training?

1. **In general, diversity training elicits strong emotional reactions, and most participants see the training as worthwhile (Level AA).**

A recent meta-analysis indicates that diversity training elicits intense emotional responses, and participants see the training as effective and worthwhile (Bezrukova et al., 2016). Training reaction is an antecedent of learning that leads to the desired behavior in the short term. However, some

people might like the training for reasons that are not directly related to its content (e.g. the trainers' sense of humor), and as a result their diversity-related attitudes and behaviors will not change.

2. Reactions to diversity training and attitudinal learning appear to decay, whereas cognitive knowledge is maintained over time (Level AA).

A meta-analysis found that cognitive learning (e.g. knowledge about different cultures) persists in the long run (Berzukova et al., 2016). After training, cues in the workplace or elsewhere reinforce cognitive responses that trainees have learned. These cues also come from other sources outside the immediate workplace or school (e.g. "Wow, this is what we learned about Italian culture in that diversity training half a year ago…"). In contrast, several meta-analyses and systematic reviews have consistently shown that participants' attitudes are less subject to change after training compared with cognitions and behaviors (Alhejji et al., 2015; Berzukova et al., 2016; Kalinoski et al., 2013; Sit et al., 2017). Most notably, attitudes may gravitate back to the original evaluative judgments after the diversity training ends if negative attitudes that a person had before the training are reinforced. In this sense, environmental prompts can even provoke "backfire effects" that reverse or retard skill development.

3. Diversity training might lead to positive outcomes beyond effects on cognitions, behaviors, and attitudes (Level A).

Several systematic reviews on diversity-related efforts in healthcare settings have investigated if the positive effects extend to higher-level outcomes. While it was found that health care professionals might improve their ability to deal with diversity following training (Chipps et al., 2008; Govere et al., 2016; Lie et al., 2010), it is not clear if this has an impact on patient satisfaction (Govere et al., 2016) or patient care (Chipps et al., 2008). However, it should be noted that generally primary studies in this context are often of poor methodological quality. Overall, studies have shown that arguments for diversity training on a business case perspective (e.g. improves performance) are not well founded (Alhejji et al., 2015).

4. The effects of diversity training are moderated by features related to context, design, and participants' characteristics.

Context

a. Positive reactions to training are considerably stronger in an educational setting compared with an organizational setting (Level AA).

A meta-analysis based on 236 high-quality studies demonstrated that the diversity training in an educational setting (i.e. higher education) elicits stronger effects than in an organizational setting (Berzukova et al., 2016). A possible explanation may be that employees in organizations see diversity training as an "add-on" practice, something that "takes time away from work" and is secondary to the purpose of the organization. Instead, diversity training in an educational setting is usually part of the mission of such institutions, so it may be perceived as an opportunity to learn about diversity and prejudice and apply concepts through experiential learning (Sit et al., 2017).

b. Training that is part of a larger diversity program tends to have better outcomes for both attitudes and behavior (Level AA).

Companies rarely make diversity training part of a broader institutionalized effort. The meta-analysis by Berzukova et al. (2016) indicates that this is a key mistake, since integrated or embedded training leads to strong behavioral learning and moderately strong changes in attitudes. The critical importance of broader systemic and system changes as opposite to a standalone approach was also highlighted in the human service organizations (Lie et al., 2010; Mor Barak et al., 2016). In fact, employees may be more motivated to learn when managers commit to diversity efforts above and beyond that of a single initiative. In addition, the components of the larger program could strengthen one another. For example, a social-networking group of minority professionals, supported by the organization, is a follow-up outcome of a diversity course and also serves as a mentoring source.

c. **Training with mandatory attendance has a stronger positive effect on behavioral learning, whereas voluntary attendance has a stronger effect on reactions (Level AA).**

Attendance requirements do not affect all outcomes, but mandatory diversity training was found to have a moderate to strong effect on behavioral learning, whereas participants' reactions were considerably stronger when attendance was voluntary (Berzukova et al., 2016). A possible explanation is that people who willingly take training may already have an interest in the issue and are thus more likely to enjoy the training. However, a voluntary approach seems not to lead to the strongest effects in diversity training. One reason for this could be that under the voluntary scenario, people participating in training already want to be there and are not necessarily the ones who would benefit most from changes in cognitive, attitudinal, or behavioral outcomes. In addition, a mandatory attendance policy might signal that the organization is truly committed to facilitating positive intergroup relationships, thus enhancing trainees' motivation to learn.

Design

d. **Diversity training has stronger positive effects when it provides longer and distributed opportunities to learn (Level AA)**

The advantage of longer training interventions does seem to transfer to more positive reactions, and better diversity knowledge, attitudes, and skills, as reflected in the strong positive relationship between the hours of training and these outcomes (Berzukova et al., 2016). It seems that the more time participants spend together leads to them getting on better and makes intergroup encounters comfortable and feel "right". Also, a recent meta-analysis based on 65 studies found that training that provides (distributed) practice, yields better attitudinal outcomes (Kalinoski et al., 2013). The finding that massed practice fares worse than distributed learning was confirmed by a recent systematic review of studies conducted in educational settings (Sit et al., 2017).

e. **Diversity training that provides greater opportunities for cooperative contact and social interaction within the training pool yields mixed findings (Level AA).**

Diversity training can either focus on one group (e.g. race) or target multiple groups (e.g. race, gender, sexual orientation, etc.). The group-specific approach (i.e. involving only the minority/majority group) has been criticized as leading to intergroup differentiation and polarized attitudes. Several meta-analyses, however, have demonstrated that the outcomes of these two types of training are mixed. For example, several studies found that training that targets multiple groups and in which trainees interact with members of other groups has greater effects on attitudes compared with training that focuses on one group (Kalinoski et al., 2013), but an equal number of studies found no differences (Berzukova et al., 2016).

f. Diversity training has stronger positive effects when it emphasizes both awareness and behavioral components (Level AA).

Most effective types of diversity training programs were primarily designed to increase both diversity awareness and skills (Berzukova et al., 2016). Awareness training focuses on getting participants to be more aware of their own and other cultural assumptions, values, and biases. Skill-building (behavioral) training educates participants on monitoring their own actions and appropriate responses to specific differences, such as identifying and overcoming interracial communication barriers. Overall, several studies show that combining both awareness and behavioral components enables people to better understand their behavior (being aware of why they are doing what they are doing; e.g. Sit et al., 2017).

g. Diversity training that combines multiple instructional methods does not seem to have stronger effects (Level AA).

Diversity training can use several different instructional methods or just one. It's reasonable to assume that training that "touches all the bases," combining multiple methods (e.g. lectures, simulation exercises, group activities and discussions, etc.) leads to better outcomes. However, the most updated meta-analysis found that this has an effect only on how much people like training, i.e. on their reactions, and it does not affect cognitive, attitudinal, or behavioral learning (Berzukova et al., 2016). Nonetheless, findings from educational settings have suggested that multi-methods delivery modes seem cross-culturally effective (Sit et al., 2017). More generally, active methods and face-to-face methods still seem to lead to more positive attitudes than passive ones and online training (Kalinoski et al., 2013).

h. Diversity training that focuses on one single aspect of demographic diversity has a stronger positive effect than aiming at multiple aspects (Level A).

While focusing on multiple aspects of diversity training (e.g. generic diversity training, multicultural, or sexual harassment) yielded moderate effects on cognitive learning, it was found that dealing with one aspect of demographic diversity (e.g. race) at a time leads to considerably stronger effects on cognitive learning (Kalinoski et al., 2013). However, it should be noted that the majority of studies upon which this finding is based considered only race. Instead, for instance, a systematic review has concluded that there is a lack of empirically validated diversity training programs that specifically focus on disability (Philips et al., 2016). Therefore it is possible that the evidence for diversity-specific interventions may be missing.

Trainee and trainer characteristics

i. A larger proportion of women in the training group leads to stronger reactions (e.g. liking), whereas a more diverse training group might have greater effects on cognitive learning (Level AA/A).

Since women have a history of discrimination, the finding that they tend to be more receptive and welcoming of diversity training makes sense. Findings regarding the race composition of the training pool are, however, mixed: Berzukova et al. (2016) did not find an effect for a more diverse make-up on the outcomes assessed, whereas Kalinoski et al. (2013) found that people learn more (i.e. cognitive learning) when training involves a higher number of ethnic minorities. The latter finding, however, aligns with the hypothesis that social interdependence and contact under optimal conditions might reduce prejudice.

j. Characteristics of the trainer affects outcomes at the affective and cognitive levels (Level A).

It was found that trainees' motivation is higher and effects on attitudes are stronger when a direct manager/supervisor delivers the training, as opposed to an internal trainer belonging to a minority group, a diversity and inclusion manager, or a HR generalist. However, it was also found that cognitive learning (e.g. knowledge about other cultures) was strongly affected when the trainer belonged to a minority group (Kalinoski et al., 2013).

Other findings

Studies on the effects of de-biasing training are almost absent.

Simply being aware that human judgment is subject to cognitive biases does not prevent them from occurring. Even Daniel Kahneman, the world's leading authority on this subject, stated: "I've been studying this stuff for 45 years, and I'm no better than when I started. I make extreme predictions. I'm over-confident. I fall for every one of the biases." Not surprisingly, studies on the effect of de-biasing training are almost absent. However, a recent study found that a single training intervention (i.e. playing a computer game or watching an instructional video) has sparked debiasing effects that persist across a variety of contexts affected by the same bias (e.g. blind spot bias, confirmation bias, fundamental attribution error, anchoring, social projection, representativeness; Morewedge et al., 2015). Here, games that incorporated personalized feedback and practice yielded the strongest effect. It is suggested by the authors that this type of simple intervention can be used alongside improved incentives, information presentation, and nudges to reduce costly errors associated with biased judgments and decisions. It should be noted, however, that no single study can be considered to be strong evidence – it is merely indicative.

9. Conclusion

Diversity training may have small to moderate positive effects on diversity and (perception of) inclusion, but only when a wide range of moderators are taken into account.

10. Limitations

This CAT aims to provide a balanced assessment of what is known in the scientific literature about the effects of diversity training by using the systematic review method to search and critically appraise empirical studies. However, in order to be 'rapid', concessions were made in relation to the breadth and depth of the search process. As a consequence, some relevant studies may have been missed.

A second limitation concerns the critical appraisal of the studies included, which did not incorporate a comprehensive review of the psychometric properties of the tests, scales and questionnaires used.

Finally, this CAT focused only on meta-analyses and high-quality single studies, i.e. studies with a control group and/or a before- and after-measurement. For this reason, cross-sectional studies were excluded. As a consequence, new, promising findings that are relevant for practice may have been missed.
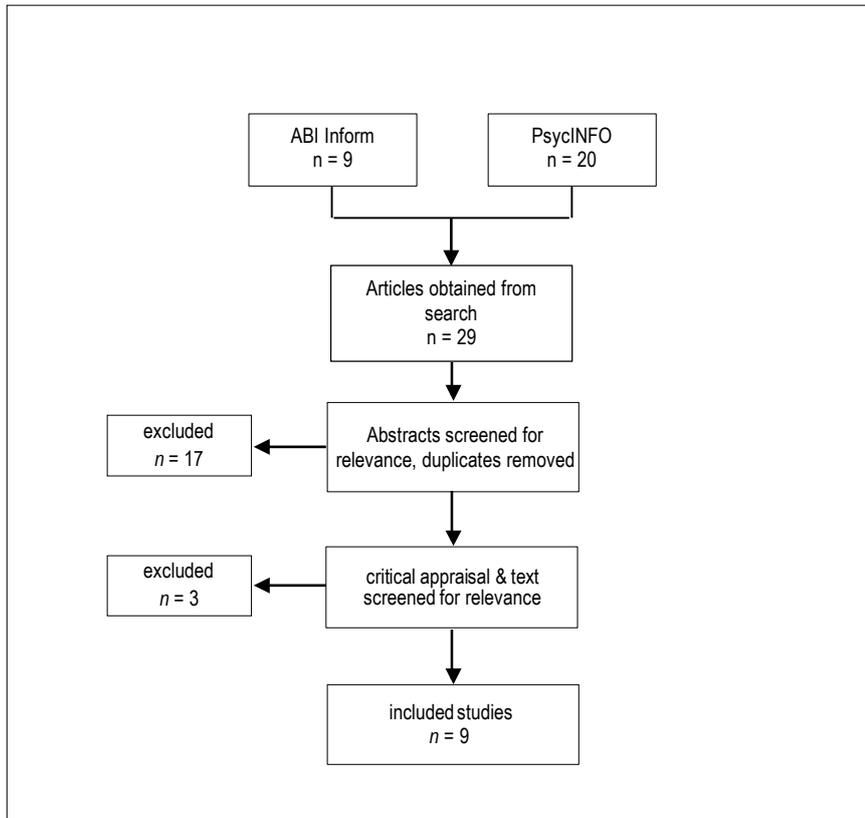
References

Alhejji, H., et al. (2016). Diversity training programme outcomes: A systematic review." *Human Resource Development Quarterly*, 27(1), 95–149.

Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A Meta-Analytical Integration of Over 40 Years of Research on Diversity Training Evaluation. *Psychological Bulletin*

Chipps, J. A., Simpson, B. & Brysiewicz, P. (2008). The Effectiveness of Cultural-Competence Training for Health Professionals in Community-Based Rehabilitation: A Systematic Review of Literature. *Worldviews on Evidence-Based Nursing,* 5(2), 85–94.

Govere, L., & Govere, E.M. (2016). How effective is cultural competence training of healthcare providers on improving patient satisfaction of minority groups? A systematic review of literature. *Worldviews on Evidence-Based Nursing,* 13(6), 402–410.

Kalinoski, Z. T., et al. (2013). A meta-analytic evaluation of diversity training outcomes." *Journal of Organizational Behavior,* 34(8), 1076–1104.

Lie, D. A., et al. (2011). Does cultural competency training of health professionals improve patient outcomes? A systematic review and proposed algorithm for future research. *Journal of general internal medicine,* 26(3), 317–325.

Morewedge, C. K., et al. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences,* 2(1), 129–140.

Mor Barak, M. E., et al. (2016). The promise of diversity management for climate of inclusion: a state-of-the-art review and meta-analysis. *Human Service Organizations: Management, Leadership & Governance,* 40 (4), 305–333.

Phillips, B. N., et al. (2016). Disability Diversity Training in the Workplace: Systematic Review and Future Directions. *Journal of occupational rehabilitation,* 26(3), 264–275.

Sit, A., Mak, A. S., & Neill, T. J. (2017). Does cross-cultural training in tertiary education enhance cross-cultural adjustment? A systematic review. *International Journal of Intercultural Relations,* 57, 1–18.

Annex I
Search terms & hits Diversity Training

Search terms	ABI	PsycINFO
S1: TI(divers*) AND TI(training)	94	193
S2: AB("diversity training")	128	434
S3: TI("cultur*") AND TI(training)	216	498
S4: AB("cultur* training")	163	236
S5: AB(cultur* awareness) AND AB(training)	231	113
S6: TI("diversity education") OR AB("diversity education")	27	63
S7: TI(bias*) AND TI(training)	11	145
S8: AB("bias* training")	6	12
S9: TI(debias*)	46	59
S10: S1 OR …. S9	759	1,470
S11: TI(meta-analy*) OR AB(meta-analy*) OR TI("systematic review") OR AB("systematic review")	-	-
S12: S10 AND S11	9	20
Total	**29**	

ANNEX II   STUDY SELECTION DIVERSITY TRAINING

ABI Inform n = 9	PsycINFO n = 20

Articles obtained from search
n = 29

excluded
n = 17

Abstracts screened for relevance, duplicates removed

excluded
n = 3

critical appraisal & text screened for relevance

included studies
n = 9

ANNEX III   MATRIX DIVERSITY TRAINING

Author & year	Design + sample size	Sector / Population	Main findings	Limitations	Level
Bezrukova et al. (2016)	meta-analysis of k = 260 experimental, quasi, and non-experimental studies involving N= 29,407 subjects	Adult population 89% in the USA 80% from educational settings	**Main effect** Diversity training has an overall mean effect of g = .38 on training outcomes. **Training outcomes** 1. Diversity training has a larger effect on participants' reactions (g = .61) than on their cognitive learning (g = .57), behavioral learning (g = .48) and attitudinal learning (g = .30). 2. Cognitive learning is maintained over time, while reactions and attitudinal learning decay. **Training input(s)** **A. Context** 1. Training setting doesn't have a significantly different effect on the outcomes, only on reactions (g = .80 for educational setting vs. g = .28 in organizations), meaning that other factors might influence how employees receive this type of training in organizational settings. 2. Training as part of a more complex diversity initiative had a greater effect overall than when it was standalone, specifically on attitudinal learning (g = .47) and behavioral learning (g = .86). 3. Whether attendance was mandatory of voluntarily didn't make a difference on the overall effect the training had on outcomes. However, mandatory training had a stronger effect on behavioral learning (g = .63), while voluntary training on reactions (g = .71) 　**B. Design** 1. The focus of the training on one group, several groups or adopting an inclusive approach did not make any difference on its effect on the outcomes. 2. Longer trainings have a greater effect on all outcomes. 3. Awareness based trainings have a lower overall effect (g = .31) than other types of training, especially on attitudinal and behavioral learning. The most effective are combined trainings. 4. Using many different instructional methods doesn't make a difference in the overall effect, and it only has a higher effect on reactions (people like diverse methods; g = .73). 　**C. Trainee characteristics** 1. Age of participants doesn't make a difference in the effect of the training on outcomes. 2. More women in the training groups only had an effect on reactions, but not overall or on the other outcomes. 3. How many white people there are in the training group also doesn't affect the training's impact on the outcomes.	part of the studies used explicit measures of attitudes and behaviors, which are prone to desirability bias	AA
Alhejji et al. (2015)	Systematic review of k = 61 cross-sectional surveys	Public or private, manufacturing or service, profit or not, SME, MNC Mainly in USA and Canada	Three theories are used to investigate diversity training: business case, social justice, and learning outcomes. Findings mainly within the **learning perspective**: (self-report) enhanced employees' knowledge and awareness of diversity issues (k = 36); (self-report) enhanced diversity skills and behaviours on diversity issues (k = 9); (self-report) changes in attitudes towards diversity (k = 5). Few studies on **business case perspective**. Two studies reported organizational performance impacts. One study found increased sales, customer satisfaction, and productivity gains as measured with archival data on employees' annual survey and branch performance data. Very few studies on **social justice perspective**, reportedly improved relationships, enhanced tolerance towards minorities, and improved confidence to work with diverse groups. Call for integrating theoretical perspectives, appropriate research designs, and nuanced study of antecedents of diversity training outcomes.	Small sample sizes Poor use of diversity-training measures Self-report Little longitudinal investigations	C

			Main findings		
Kalinoski et al. (2013)	Meta-analysis of k = 65 studies (control + pre-test post-test, control + only post-test, single group pre-post test) with N = 8465	Adult and student populations from wide range of industries	**Main findings** Findings revealed sizable effects on affective-based (d = .30), cognitive-based (d = .71), and skill-based outcomes (d = .47). Mixed evidence was found for attitude change (e.g., being more tolerable towards diversity as a result of the training), but larger effects for cognitive and skills based outcomes. Overall, study addressed the psychological effects of diversity training on change as opposed to how training actually worked in organizations. **Moderators** Interesting boundary conditions were found for the effects on affective-based outcomes → greater opportunities for social interactions, namely, higher task interdependence (d = .41) vs lower (d = .14), active and passive methods (d = .37) vs only passive (d = .07), online (d = .08; ns) vs face to face (d = .32), less than four hours duration (d = .11; ns) vs more than four hours d = .52, .49, .46), distributed (d = .45) vs massed practice (d = .21). Again for affective-based outcomes, higher trainee motivation was related to: training delivered by direct manager/supervisor (d = .44) vs internal other vs diversity/inclusion manager, HR generalist (d = .05); participant employee (d = .50) vs student (d = .22); study setting field (d = .35) vs lab (d = .07; ns). Similar findings on motivational factors related cognitive based outcomes, with the exception of internal 'other' trainer yielding larger effect (d = 1.61) than training from direct manager (d = .51). **Control variables** Needs assessment (little information present, ns), voluntary vs mandatory (ns), awareness (e.g., cognitive biases) vs awareness + skills training (ns), compliance/legal content in training (ns), differences vs differences + similarities (ns). Training (d = .26) vs education (d = .45) on affective-based outcomes; subject matters focused on a single attribute of demographic diversity (e.g., race) yielded larger effect size (d = 1.28) on cognitive-based outcomes than generic (d = .52), multicultural (d = .54), and sexual harassment (d = .54); more or less diverse composition of participants pools on cognitive based outcomes: such as greater effects when pools composed by less than 40% of Caucasian (d = 1.28) compared with more than 60% Caucasian (d = .42), stronger effects when women make up was greater than 60% (d = .71) vs lower than 40% (d =.38); finally, larger effects when trainer was Caucasian (d = 1.61), than for non-Caucasian (d = .54) (interpret with caution because based on four studies). Finally, on cognitive-based outcomes, other findings were: self-reports (d = .85) vs non self-reports (d = .60); end-of-training evaluation (d = .82) vs time lag greater than one month (d = .39); single group repeated measures (d = 1.18); training vs control post-test only (d = .45); training vs control repeated measures (d = .66). Published (d = .38) vs unpublished (d = .13) on affective-based outcomes.	Search for unpublished was limited Quality not assessed, but controlled for study rigor effect	A
Philips et al. (2016)	Systematic review of k = 3 studies, two with pre-post comparison for treatment group only, one with randomised controlled trial	Employer, employees, HR professionals in organisations, excluding disability care orgs	Insufficient evidence to make generalizable claims on the effectiveness of disability specific diversity training interventions. One intervention found significant gains in attitudes, knowledge and behavioural intentions between pre- and post-training immediately after training. Qualitative follow up between four to eight months highlight that behavioural intentions that were enacted concerned (in order of frequency): individual actions, teaching others, building partnerships and changing organisational policies and practices. Previous experience with disability and working in a supportive environment facilitated action following training. The other two studies focussed on training supervisors to support employees with musculoskeletal pain or disability due to workplace injuries. One study found no significant supervisor's confidence levels to deal with disabled employee. Other study found significant declines in injury claim rates by employees for the group whose supervisors were trained. Study provides guidelines for designing training, defining content, accounting for participants factors.	Relatively low quality of studies based on rigorous criteria – no confounding check (Cochrane guidelines)	A
Govere et al. (2016)	Systematic review of k = 6 of which one experimental and five non-experimental study	Healthcare providers in USA (k = 5), UK (k = 1), Canada (k = 1), such as dental practitioners, physicians, nurses	Six studies revealed that cultural competence training intervention significantly increased the cultural competence level of healthcare providers. Five studies demonstrated that cultural competence training of healthcare providers was significantly associated with increased patient satisfaction. Two studies that were less biased (i.e., powered samples, superior experimental design, and probability sampling) found no significant effect of cultural competence training on patient satisfaction.	primary studies quality, e.g., use of not context specific measures, and not valid	B
Sit et al. (2017)	Systematic review of k = 35 studies including k 46 =	Tertiary education students, e.g., medical, counselling, and	Review based on psychological theory of cross-cultural adjustment (namely, ABC model by Ward et al. 2001). Study found that experiential and multi-method delivery modes are more cross-culturally effective (e.g., role plays and individual or group exercises). Nearly two thirds of training programmes used three or more methods. Distributed delivery was more effective. According with ABC model which posit that successful acculturation is a robust process	Publication bias Quality of primary studies	A

	cultural competence intervention groups, median group size 32 and total 2834 students	post-graduates	encompassing affective, behavioural and cognitive change, trainings that used both cognitive change and behavioural modification were more effective than cognitive alone and didactic alone. Programs were more effective in facilitating cross-cultural knowledge and inducing behavioural adjustment, than in fostering cross-cultural cognitive and emotional adjustments. Study also found that benefit in multicultural competence appear to influence students' academic performance and preparation for multicultural careers.		
Lie et al. (2010)	Systematic review of k = 7 studies: two quasi-randomised, two clusters randomised, three pre-test post-test field studies	Physicians, mental health professionals, and multiple health professionals and students	Included studies, albeit of limited quality, showed a trend for cultural-competence training as resulting in a positive impact on patient outcomes. It suggests that CC training as a standalone strategy may not suffice to improve outcomes, so it should fall within broader systemic and system changes. Provides a framework for conducting highly rigorous and relevant research in training settings.	Low to moderate primary study quality	A/B
Chipps et al. (2008)	Systematic review including k = 5 single studies and k = 1 systematic review	Focused on Western context (i.e., USA)	Based on Beach et al.'s (2005) review and the five included studies, reasonable evidence exists to indicate that cultural-competence training can increase the knowledge, attitudes, and skills of health professionals. Some evidence seems to indicate that cultural-competence training affects patient satisfaction, but little evidence that it improves patient care.	Primary studies seriously flawed	B
Mor Barak et al. (2016)	Meta-analysis of k = 30 studies with N = 496,740 workers	Human services organisations in USA and other locations	**Main findings** Results show that although diversity is associated with both beneficial and detrimental outcomes, diversity management efforts that promote a climate of inclusion are consistently associated with positive outcomes (p = .46) (N = 290,854). Organisations should move beyond increasing diversity to fostering inclusion, several practical implications are provided. Note: study mention relevant socio-psychological theories on diversity, different types of diversity, diversity management efforts, and both beneficial and detrimental outcomes.		C

Excluded

Morris et al. (2001)	Meta-analysis of k = 16 studies on expatriate adjustment and k = 25 on expatriate performance	American expats	**Not relevant for research question** Cross culture training for American expatriates showed effects weaker than expected	Performance r = .26 Adjustment r = .13	Moderators not well understood due to primary research limitations	N/A
Martinez et al. (2015)	Systematic review of k = 40 studies	N/A	**Not relevant for research question** Investigated how immigration policies and laws may affect both access to health services and health outcomes among undocumented immigrant. Negative relationships were found with access to health services and mental health outcomes.	N/A	N/A	N/A
Dong et al. (2010)	Meta-analysis	N/A	**Not relevant for research question** Cross cultural issues and management in Chinese organisations. States concerns about research on cross-cultural management being conducted in Western countries, and lack of it in Eastern ones.	N/A	N/A	N/A