



The impact of diversity training on workplace behavior – A meta-analysis

Journal:	<i>Journal of Workplace Learning</i>
Manuscript ID	JWL-03-2025-0051.R1
Manuscript Type:	Research Paper
Methodologies:	Quantitative
Keywords:	Training transfer, training effectiveness, Diversity Management, Learning Transfer

SCHOLARONE™
Manuscripts

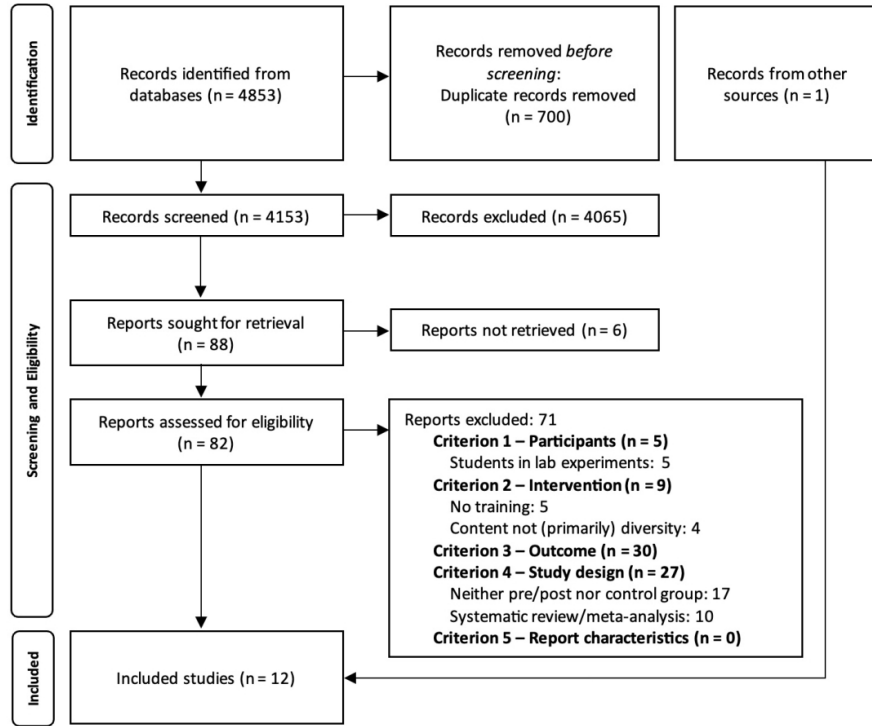


Figure 1. PRISMA flowchart Source: Authors own work

275x225mm (144 x 144 DPI)

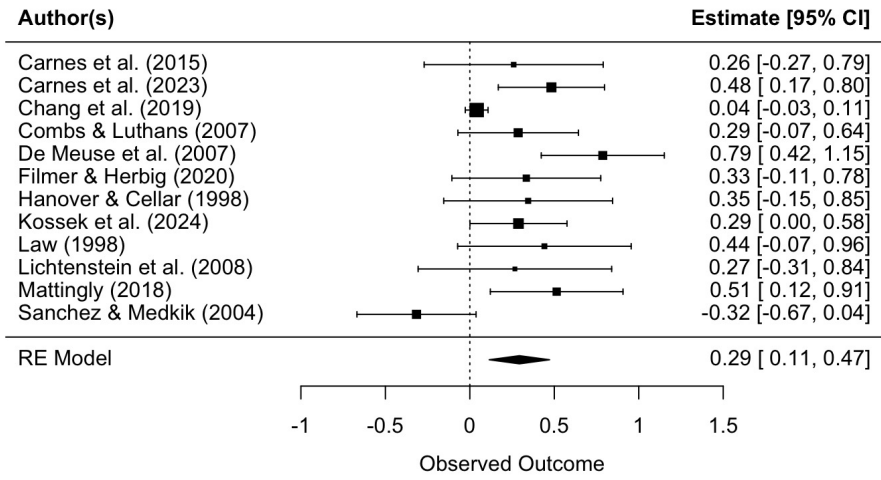


Figure 2. Forest plot Source: Authors own work

492x324mm (72 x 72 DPI)

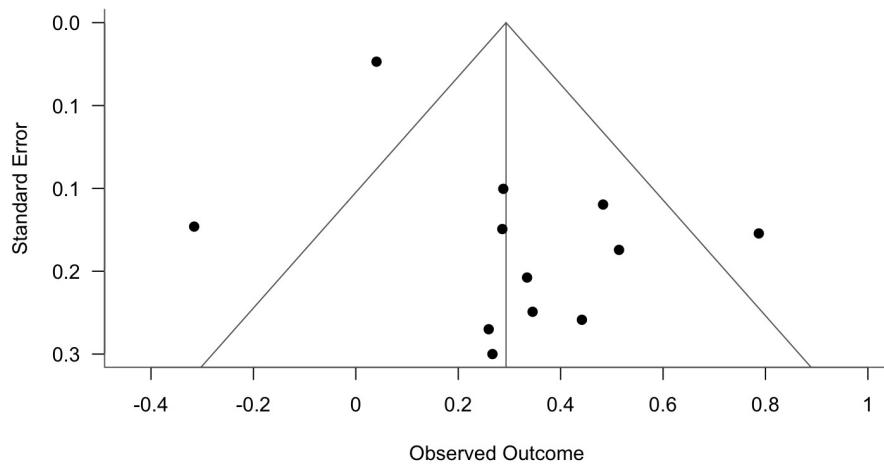


Figure 3. Funnel plot Legend: Note. Black circles represent observed data points. Source: Authors own work

521x324mm (72 x 72 DPI)

Table I

Study	Type ^a	Country	Sample	Intervention	Study design ^b	Transfer measurement(s)	Transfer Measuring Point	Data source
Games <i>et al.</i> (2015)	P	USA	University faculty	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : Yes	PPWC, cluster randomized	Acting on a regular basis to promote gender equity in one's department	3 months after the intervention	Self-report
Carnes <i>et al.</i> (2023)	P	USA	University faculty	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : Yes	PPWC, cluster randomized	Taking action to reduce bias: Acting on a regular basis to promote bias reduction activities in one's division	3 months after the intervention	Self-report
Chang <i>et al.</i> (2019)	P	Various	Staff from a large global organization	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : No	POWC, randomized	Number of women selected to be informally mentored over coffee/per participant; Number of women selected to be recognized as excellent/per participant	3-6 weeks after the intervention	Objective
Combs & Luthans (2007)	P	USA	Staff from three organizations	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : No	POWC, randomized	Difficulty of pro-diversity actions attempted; Mean number of pro-diversity actions attempted; Success with carrying out pro-diversity actions	1 year after the intervention	Self-report
De Meuse <i>et al.</i> (2007)	P	USA	Managers from a production company	<i>Needs analysis</i> : No <i>Attendance policy</i> : Unknown <i>Action planning</i> : Yes	SGPP	What a person does in response to diversity: A person's verbal as well as nonverbal actions	3 months after the intervention	Self-report
Filmer & Herbig (2020)	P	Germany	Nursing staff	<i>Needs analysis</i> : Yes <i>Attendance policy</i> : Voluntary <i>Action planning</i> : No	PPWC, non-randomized	Self-reported behavior in situations with people from other cultures (cultural communication and perspective taking)	6 months after the intervention	Self-report
Hanover & Cellar (1998)	P	USA	Managers from a consumer goods company	<i>Needs analysis</i> : Yes <i>Attendance policy</i> : Mandatory <i>Action planning</i> : Yes	PPWC, non-randomized	Execution of diversity practices: Behaviors believed to contribute to effective diversity management in the workplace	2 months after the intervention	Self-report
Kossek <i>et al.</i> (2024)	P	USA	University faculty	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : Yes	SGPP	Use of inclusive language and behavior with colleagues at the workplace	3 weeks after the intervention	Self-report

	Study	Type ^a	Country	Sample	Intervention	Study design ^b	Transfer measurement(s)	Transfer Measuring Point	Data source
1									
2									
3									
4									
5									
6									
7	Law (1998)	D	USA	Childhood educators	<i>Needs analysis</i> : No <i>Attendance policy</i> : Mandatory <i>Action planning</i> : No	PPWC, non-randomized	Self-reported behavior change in interactions with people from different cultures	4 weeks after the intervention	Self-report
8									
9									
10									
11	Lichtenstein <i>et al.</i> (2008)	P	USA	Rehabilitation professionals	<i>Needs analysis</i> : No <i>Attendance policy</i> : Voluntary <i>Action planning</i> : Yes	POWC, PPWC, non-randomized	Self-reported changes in service; Self-assessment of valuing culture (clear focus on actual behaviors and not attitudes); Self-assessment of dealing with diverse families	2 months after the intervention	Self-report
12									
13									
14									
15									
16	Mattingly (2018)	D	USA	Executives of a global manufacturing company	<i>Needs analysis</i> : No <i>Attendance policy</i> : Unknown <i>Action planning</i> : Yes	POWC, randomized	Participation in gender parity; Execution of ally behaviors (mentor, sponsor; build confidence) and change agent behaviors (includer, catalyst, disruptor)	3-4 months after the intervention	Self-report & third-party
17									
18									
19									
20									
21	Sanchez & Medkik (2004)	P	USA	County government managers	<i>Needs analysis</i> : No <i>Attendance policy</i> : Mandatory <i>Action planning</i> : No	POWC, non-randomized	Coworkers' ratings of discriminatory treatment enacted by trainees towards culturally different others	1 year after the intervention	Third-party
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									
38									
39									
40									
41									
42									
43									
44									
45									
46									

Note. ^aType of literature: P = published in a journal; D = dissertation. ^b Study design: POWC = posttest only with control group; PPWC = pretest-posttest with control group; SGPP = single

group pretest-posttest

Source: Authors own work

Table II

Variable	<i>k</i>	<i>g</i>	Statistics in subsamples			Moderation analysis	
			<i>SE</i>	95% CI	<i>p</i>	<i>F</i>	<i>p</i>
Needs analysis						0.04	.837
Yes	2	0.34	0.23	-0.17-0.85	.166		
No	10	0.29	0.09	0.08-0.49	.010		
Attendance policy						0.88	.374
Mandatory	3	0.09	0.15	-0.27-0.44	.597		
Voluntary	7	0.25	0.09	0.05-0.45	.021		
Goal setting/ Action planning						6.75	.027
Yes	7	0.44	0.09	0.23-0.65	<.001		
No	5	0.10	0.09	-0.11-0.31	.320		
Data source						16.03	.003
Self-report	9	0.40	0.07	0.25-0.56	<.001		
Other	2	-0.04	0.09	-0.24-0.16	.661		
Study design						2.28	.165
PPWC	5	0.38	0.13	0.08-0.69	.018		
POWC	4	0.11	0.12	-0.17-0.38	.393		
SGPP	2	0.51	0.17	0.11-0.92	.019		
Randomization						0.43	.667
Randomized	3	0.23	0.14	-0.10-0.56	.145		
Cluster-randomized	2	0.40	0.20	-0.08-0.87	.087		
Non-randomized	5	0.17	0.14	-0.15-0.50	.239		

Note. Statistics in subsamples are based on the common τ^2 of the subsets. Studies that could not be allocated to a subgroup were omitted from the analysis. *k* = number of studies included in the analysis; *g* = average effect size; *SE* = standard error of the average effect size; 95% CI = minimum and maximum limits of the 95% confidence interval; *F* = test of moderators.

Study design: POWC = posttest only with control group; PPWC = pretest-posttest with control group; SGPP = single group pretest-posttest.

Source: Authors own work

The impact of diversity training on workplace behavior – A meta-analysis

ABSTRACT

Purpose

This meta-analysis examines whether organizational diversity training results in transfer and thereby affects the workplace behavior of trainees. Furthermore, the article investigates the influence of three possible moderators.

Methodology

We conducted a systematic literature search to identify relevant studies ($n = 12$), and we synthesized the effect sizes reported in these studies via a random effects meta-analytic model. We employed subgroup analyses to examine the influence of potential moderator variables.

Findings

The meta-analysis revealed a small yet statistically significant effect ($g = 0.29$) of diversity training on workplace behavior. Subgroup analyses indicated that planning transfer behavior during training might moderate the transfer success of diversity training. No significant effects were found for needs analysis or attendance policy as moderators.

Research limitations/implications

Further primary studies should be conducted to evaluate the workplace behavior of trainees after diversity training. Ideally, future studies should go beyond solely relying on self-reports of trainees.

Practical implications

The results suggest that diversity training can positively impact workplace behavior and can thus be an effective tool for addressing the challenges and opportunities associated with a

1
2
3 diverse workforce. Furthermore, our results provide indications for transfer-enhancing
4
5 training designs.
6

7 **Originality**

8
9
10 Previous research has primarily examined the extent to which trainees achieve learning gains
11
12 after diversity training. However, evidence suggests that learning does not necessarily lead to
13
14 behavioral change in the workplace. Thus, this analysis extends our knowledge of diversity
15
16 training effectiveness.
17

18
19 *Keywords: Training transfer, Training effectiveness, Diversity management, Learning*
20
21 *transfer*
22

23
24 *Article classification: Research paper*
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Workforces are increasingly characterized by diversity (Roberson *et al.*, 2024), which in turn is driven by factors such as digitalization, globalization, and migration trends. Employee diversity encompasses a range of dimensions, including but not limited to ethnicity, religion, race, gender, and sexual orientation (Gardenswartz and Rowe, 1998). In the context of such multifaceted diversity characteristics among employees, ensuring a sense of belonging within the organization is crucial for preventing certain groups from becoming marginalized. Thus, creating an inclusive work environment where people from various backgrounds are able to work together productively has become a priority in organizations, and effective diversity management has emerged as a critical skill (Davies *et al.*, 2011; Groenewald, 2024).

A common approach that organizations employ to address the diversity of their workforce and customers is diversity training (DT) (Wang *et al.*, 2024). Dobbin and Kalev (2018) stated that almost all Fortune 500 companies implement DT. However, carrying out such training is costly. Organizations in the United States invest approximately 8 billion U.S. dollars annually in DT (Kirkland and Bohnet, 2017). Given such high expenditures, the question arises of how effective DT is.

Previous research has focused primarily on assessing the learning outcomes associated with DT (Bezrukova *et al.*, 2024). However, learning success is of limited value unless it leads to meaningful behavioral change in the workplace. Thus, organizations can only derive the benefits of training when trainees successfully apply the acquired skills, knowledge, and attitudes in the workplace, thereby achieving training transfer (Baldwin and Ford, 1988). For example, trainees may develop a theoretical understanding of key diversity principles and strategies for fostering an inclusive work environment; however, without the actual implementation of these behaviors, the organizational impact remains limited.

To date, little is known about whether DT can truly impact workplace behavior of trainees, despite this being the major goal of DT (Bendick *et al.*, 2001). This paper contributes to the DT literature in two ways. First, we examine the impact of DT on the workplace behavior of trainees via a meta-analysis. Second, by analyzing potential moderators of DT effectiveness, we provide a foundation for explaining why some programs yield better results than others. These insights can then inform the design of future DT by highlighting strategies that facilitate the transfer of learning to the workplace.

Effectiveness of diversity training programs

The conceptualization of DT used herein is based on Bezrukova *et al.* (2016), who built on Pendry *et al.* (2007) and defined DT as “a distinct set of 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” (p. 1228).

In the past, two comprehensive meta-analyses regarding the effectiveness of DT have been published (Bezrukova *et al.*, 2016; Kalinoski *et al.*, 2013). Both classified the outcome measures of DT based on Kraiger *et al.*'s (1993) framework for training evaluation. Kraiger *et al.* (1993) proposed that evaluation consists of two questions: (1) Are training objectives achieved (i.e., learning)? (2) Does the achievement of these goals result in better work performance (i.e., transfer)? In their evaluation model, Kraiger *et al.* focused on the first question and thus on “issues related to specification and measurement of learning” (1993, p. 311). They distinguished three types of learning outcomes: cognitive, skill-based, and affective.

Kalinoski *et al.* (2013) reported a small to medium effect regarding those learning outcomes ($d = 0.43$), and Bezrukova *et al.* (2016) reported similar results ($g = 0.38$). In addition to Kraiger *et al.*'s (1993) learning outcomes, Bezrukova *et al.* (2016) included

1
2
3 trainees' reactions to training in the overall effect size. DT's effect on the transfer level,
4
5 however, was not explicitly investigated. While those studies, in general, support the
6
7 effectiveness of DT, the effect sizes of the original studies vary (Bezrukova *et al.*, 2024). Not
8
9 all DT lead to the anticipated effects, and sometimes, they can even lead to adverse outcomes
10
11 and thus increase prejudice - a phenomenon called backfire or backlash (Leslie *et al.*, 2025).
12
13

14
15 A similar picture emerges when examining recent literature reviews about DT. Devine
16
17 and Ash (2022) highlighted that evidence regarding the effectiveness of DT is conflicting.
18
19 Wang *et al.* (2024) reported that the majority of the reviewed DT yielded significant results.
20
21 However, most results considered short-term changes such as knowledge and awareness.
22
23 Thus, Devine and Ash (2022, p. 403) concluded that researchers "too often use proxy
24
25 measures for success" when evaluating DT. To date, the impact of DT has been commonly
26
27 evaluated through trainees' reactions and their affective, cognitive, and skill-based learning
28
29 (Bezrukova *et al.*, 2024). However, the primary objective of DT is to affect changes in
30
31 workplace behavior (Bendick *et al.*, 2001), and findings indicate that successfully acquiring
32
33 learning content does not necessarily lead to the subsequent application of the learned
34
35 material in the job context, i.e., the transfer of training (Alliger *et al.*, 1997; Baldwin and
36
37 Ford, 1988). With respect to DT, positive transfer could, for example, imply that trainees
38
39 refrain from discriminatory behavior or wording or that they undertake efforts to support
40
41 members of minorities at their workplace (Carnes *et al.*, 2015; Sanchez and Medkik, 2004).
42
43 Our analysis thus separates itself from previous work by focusing on the transfer of DT.
44
45 Research question 1: *Does DT have a positive effect on the workplace behavior of trainees?*
46
47
48
49
50
51

52 **Moderator variables for the transfer success of diversity training**

53
54
55 As discussed, previous research has shown that the effectiveness of DT varies, with some
56
57 programs being more impactful than others (Bezrukova *et al.*, 2024). Understanding the
58
59
60

1
2
3 factors that increase the probability of successful transfer is essential for creating effective
4
5 DT.
6

7
8 According to Baldwin and Ford (1988), the transfer success of training depends on
9
10 three areas: training design, work environment, and trainee characteristics. Since the
11
12 publication of their framework, researchers have attempted to identify transfer-promoting
13
14 factors for different training programs within these areas (e.g., Lacerenza *et al.*, 2017). Some
15
16 scholars have focused specifically on training in so-called soft skills, as transfer success in
17
18 this type of training may be particularly difficult (Hamzah *et al.*, 2024). Soft skills are defined
19
20 as skills related to interaction with oneself or with others (Laker and Powell, 2011) and
21
22 therefore include DT. Laker and Powell (2011) argued that the transfer of soft skills is
23
24 associated with more obstacles than the transfer of hard skills (e.g., handling software). They
25
26 propose that soft-skill training is associated with less precise identification of training
27
28 objectives and more trainee resistance because trainees might be less aware of their training
29
30 needs. Furthermore, soft-skill training content could challenge trainees' beliefs and lead to
31
32 anxiety and resistance (Boere *et al.*, 2023; Laker and Powell, 2011), which might be
33
34 especially plausible for DT given its propensity to elicit stronger emotional reactions than
35
36 other training programs and even backlash (Bezrukova *et al.*, 2016). For those reasons, we
37
38 investigate three training design moderators that could be associated with the effectiveness of
39
40 DT as a potentially delicate soft-skills training.
41
42
43
44
45

46 47 *Needs analysis*

48
49 Conducting a needs analysis helps practitioners design programs aligned with organizational
50
51 and trainee requirements (Lacerenza *et al.*, 2017), which are often less obvious for soft skills
52
53 than for hard skills. For example, in a meta-analysis by Lacerenza *et al.* (2017), the transfer
54
55 of leadership training - a typical soft-skill training - was more effective when training was
56
57 built on needs analysis. Similar conclusions might hold true for DT. When diversity experts
58
59
60

1
2
3 were asked to identify the components of effective DT, 67% stated that programs need to
4
5 meet the organization's needs (Wentling and Palma-Rivas, 1999). A needs analysis enables a
6
7 more learner-centric approach to DT, where training content is tailored to individual needs
8
9 (Roberson *et al.*, 2024). For example, a needs analysis could provide information on whether
10
11 participants belong to the social majority or to certain marginalized groups and what their
12
13 previous experiences and attitudes toward various aspects of diversity are. Thus, training that
14
15 builds on a needs analysis and subsequently addresses trainees' pretraining beliefs and
16
17 knowledge could be more effective and reduce the likelihood of backlash (Roberson *et al.*,
18
19 2024; Sanchez and Medkik, 2004). These theoretical and empirical considerations led us to
20
21 investigate needs analysis as a potential moderator.
22
23

24
25
26 Research question 2: *Does a needs analysis affect the effectiveness of DT?*

27 28 29 *Attendance policy*

30
31 When planning training, an important question is whether attendance should be mandatory or
32
33 voluntary. Mandatory training might signal to trainees that the program is important
34
35 (Bezrukova *et al.*, 2016). Furthermore, to train entire teams, it is usually necessary to make
36
37 training mandatory. The support of colleagues is an important prerequisite for transfer
38
39 (Blume *et al.*, 2010). Therefore, scholars have suggested training entire teams so that their
40
41 respective knowledge, skills, and attitudes are acquired collectively and that trainees can
42
43 support each other's transfer attempts at the workplace (Salamon *et al.*, 2022). On the other
44
45 hand, mandatory DT might cause trainees to perceive the training as a form of punishment,
46
47 thus leading to resentment (Sanchez and Medkik, 2004); in such cases, a mandatory
48
49 participation policy can lead to backlash (Salamon *et al.*, 2021).
50
51
52

53
54 In contrast, voluntary participation ensures that trainees are motivated and thus more
55
56 likely to transfer training content (Lacerenza *et al.*, 2017). Lacerenza *et al.* (2017) reported
57
58 that although voluntary participation may decrease overall attendance, it can increase transfer
59
60

1
2
3 success for those who attend. However, a concern with voluntary DT is that it might
4
5 primarily attract individuals who are already supportive and least likely to create barriers for
6
7 marginalized groups (Kulik *et al.*, 2007).
8
9

10 While transfer research, in general, has produced inconsistent results regarding
11
12 whether mandatory or voluntary training leads to more positive transfer outcomes (Salamon
13
14 *et al.*, 2021), Bezrukova *et al.* (2016) and Kalinoski *et al.* (2013) reported no significant
15
16 difference in learning outcomes as a function of attendance policy for DT. The debate on
17
18 whether DT should be compulsory or voluntary is still ongoing (Bezrukova *et al.*, 2024), and
19
20 we consider attendance policy a potential moderator that is worth investigating, especially in
21
22 light of the backlash phenomenon specific to DT.
23
24

25
26 Research question 3: *Does the effectiveness of DT differ depending on the attendance policy?*
27

28 *Goal setting/action planning*

29

30
31 Scholars recommend setting specific but challenging transfer goals regarding trainees'
32
33 behavior in the workplace (Burke and Hutchins, 2007). The concept of action planning is
34
35 related to goal setting but goes one step further. Regarding health behavior change, Bailey
36
37 (2019) defined action plans as specifying “*where, when, and how* a goal will be implemented
38
39 [...] and help[ing] individuals plan the specific actions they will take to achieve their
40
41 overarching goal" (p. 617). Foxon (1994) argued that creating action plans helps establish a
42
43 cognitive connection between the training environment and the workplace. This is
44
45 particularly important for soft-skill training, such as DT, where the application context is less
46
47 clear than that of hard-skill training (Laker and Powell, 2011).
48
49
50

51
52 In their literature review, Burke and Hutchins (2007) stated that both goal setting and
53
54 action planning have received conceptual and empirical support but that more empirical
55
56 research is needed. Blume *et al.* (2010) reported a marginal relationship between goal setting
57
58 and the transfer of training ($p = .08$) in their meta-analysis. A review by Botke *et al.* (2018),
59
60

1
2
3 which focused on soft-skills training only, suggested that goal setting can promote transfer in
4
5 this context.
6

7
8 Given these mixed previous findings and the transfer-promoting potential that goal
9
10 setting and action planning could have for DT, we decided to include this category in our
11
12 moderation analysis. Since goal setting and action planning are not always clearly
13
14 differentiated in intervention studies, we combined them into one category.
15

16
17 Research question 4: *Does goal setting/action planning impact the effectiveness of DT?*
18
19

20 21 **Methods**

22 23 *Eligibility criteria and selection of studies*

24
25 The eligibility criteria were derived based on the PICOS approach (participants, intervention,
26
27 comparison, outcome, study design) (Amir-Behghadami and Janati, 2020). For a study to be
28
29 included in our meta-analysis, it had to meet the following criteria: *participants* were
30
31 employees in a workplace setting; the *intervention* involved a training program with a clear
32
33 focus on diversity (both generic DT and programs targeting specific diversity aspects (e.g.,
34
35 gender) were included); the *outcome* variable was transfer, i.e., on-the-job behavior that was
36
37 measured at least one week after training; and the *study design* involved a pretest–posttest
38
39 design or a control group. As a comparison group was not a necessary condition for all
40
41 studies, no criterion was developed for the *comparison* aspect of the framework. Only reports
42
43 in English or German were included.
44
45
46
47

48 49 *Literature search*

50
51 We systematically searched the PsycARTICLES, PsycINFO, PSYINDEX, Business Source
52
53 Premier, and ERIC databases. The following search string was used: (employ* OR work* OR
54
55 business) AND (diversity OR cultur* OR bias OR "social inclusion" OR discrimination OR
56
57 "prejudice reduction" OR gender OR "sexual harassment" OR LGBT OR Rac* OR
58
59
60

1
2
3 homophob* OR transphob* OR wom?n) AND (training OR intervention OR workshop)
4
5 AND (evaluat* OR assess* OR impact OR effect*). In addition to these English search terms,
6
7 German equivalents were used. A total of 4853 articles were identified through the search.
8
9 After the articles were assessed for eligibility via the Rayyan software (Ouzzani *et al.*, 2016),
10
11 11 articles remained. Correspondence with one of the authors yielded an additional eligible
12
13 study. Thus, in total, 12 articles met the inclusion criteria. The PRISMA flowchart (Page *et*
14
15 *al.*, 2021) is presented in Figure 1.
16
17

18
19 **Figure 1: PRISMA flowchart**

20
21 *Coding procedures*

22
23 The two authors independently coded the included studies. The coding system included the
24
25 following characteristics: publication year, author(s), type of literature, country, needs
26
27 analysis, attendance policy, goal setting/action planning, transfer conceptualization, source of
28
29 transfer rating, study design, and randomization. Further information is provided in the
30
31 supplementary material ([Link](#)). Interrater reliability was assessed on an item-to-item basis
32
33 with interrater agreement statistics. Interrater agreement was high, with agreement rates
34
35 ranging from 90% to 100% per category. Disagreements were resolved by discussion among
36
37 the coders.
38
39
40
41

42
43 *Computation and analysis of effect sizes*

44
45 Hedges' *g* was calculated to assess the effectiveness of DT. This statistic is a standardized
46
47 mean effect size measure that corrects for the bias in Cohen's *d* that has been observed in
48
49 small samples (Hedges and Olkin, 1985). The interpretation of Hedges' *g* is the same as that
50
51 of Cohen's *d*, and effect sizes around .20 are considered small, those around .50 are
52
53 considered medium, and effect sizes near .80 are considered large (Cohen, 1988). Effect sizes
54
55 were coded uniformly, and a positive difference indicates positive training outcomes.
56
57
58
59
60

1
2
3 In the *posttest only with control group designs (POWC)*, where the outcome was
4 measured solely after the intervention, the effect size was calculated as the difference
5 between the means of the control and treatment group divided by the within-groups standard
6 deviation pooled across groups (Borenstein *et al.*, 2021). In *single-group pretest–posttest*
7 *designs (SGPP)*, where the outcome was measured in a single group before and after the
8 treatment, the effect size was calculated as the difference between the posttest and the pretest
9 divided by the standard deviation of the pretest (Becker, 1988). The variance of effect sizes in
10 such study designs depends on the correlation of the pre- and posttest scores. If such
11 information was unavailable, authors were contacted or formulas were used to infer the
12 correlation from other reported statistics (Morris and DeShon, 2002). When this was
13 impossible, we used the mean correlation between pre- and posttest scores across all studies
14 to fill in missing values. If means were reported for different pretest and posttest sample sizes
15 due to dropouts, the posttest sample size was used for variance calculations. In *the*
16 *pretest–posttest with control group designs (PPWC)*, where the outcome was measured
17 before and after the treatment in the control and treatment group, we first calculated an effect
18 size for each group and subsequently subtracted the effect size of the control group from the
19 effect size of the treatment group to obtain the overall effect size. These procedures ensured
20 that effect sizes from different study designs were within the same metric (Morris and
21 DeShon, 2002).

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

If an effect size was not based on means but on *binary data*, a log odds ratio was calculated and then transformed to the standardized mean difference (Borenstein *et al.*, 2021). Two studies (Carnes *et al.*, 2015, 2023) had a *cluster-randomized design*, where means and standard deviations were only available for the clusters. In both studies, effect sizes were calculated based on the cluster-level data (Chaplin and Dwan, 2023). All effect sizes were computed in *R* with the package *metafor* (Viechtbauer, 2010).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In a few of the included studies, multiple outcomes were measured to assess the transfer of DT. Individual effect sizes were calculated, and then the mean of the scores was used to create one single effect size per study (Borenstein *et al.*, 2021). This prevented us from assigning more weight to these studies, given that these outcomes did not provide independent information (Borenstein *et al.*, 2021). The variance of the aggregated effect size depends on the correlation of the variables (Borenstein *et al.*, 2021). Such correlations are seldom reported in primary studies. One of the included studies, however, reported correlations between transfer measurements ($r = 0.59$). We used that mean correlation for all our studies and conducted sensitivity analyses with other values ($r = 0.2$; $r = 0.8$; $r = 1$) to determine whether the results were robust. Another approach to handle the dependency of data would have been the recently introduced multilevel model with robust variance estimation (Pustejovsky and Tipton, 2022). We decided not to use a multilevel model since the number of reported effect sizes was distributed very unevenly across our studies, with more than half of the effect sizes reported in one dissertation. This uneven distribution would have led to a tenuous multilevel model.

Since the studies exhibited diverse characteristics, we assumed a random-effects model. Furthermore, we conducted the Knapp–Hartung method to obtain robust confidence interval estimates and significance tests (IntHout *et al.*, 2014).

Additional analyses

The hypothesized moderators were examined via subgroup analyses based on a mixed-effects model. A mixed-effects model means that a random-effects model was used within subgroups and a fixed-effect model was used across subgroups (Borenstein *et al.*, 2021). In such models, the residual between-studies variance – i.e., the heterogeneity in true effect sizes not explained by the moderator variable – can either be separately estimated for each category in the moderator or pooled across all moderator categories (Rubio-Aparicio *et al.*, 2020). If the

1
2
3 number of studies per category is small, it is advised not to use separate estimates of Tau^2
4
5 (Borenstein *et al.*, 2021; Rubio-Aparicio *et al.*, 2020). Thus, our models assume that
6
7 heterogeneity is the same within each subset. We implemented meta-regressions in *R* with
8
9 dummy variables for the different subgroups of the categorical moderators. Omnibus tests
10
11 were used to assess whether there were significant subgroup differences. We applied the
12
13 Knapp–Hartung method, which means that the moderator tests were based on an
14
15 F distribution.
16
17

18
19 We performed publication bias analyses by examining a funnel plot, which illustrates
20
21 effect sizes against their standard errors. An asymmetric funnel plot may suggest bias, often
22
23 stemming from the underrepresentation of studies with nonsignificant results (Borenstein *et*
24
25 *al.*, 2021). We conducted rank correlation and regression tests to further investigate
26
27 asymmetry. All analyses were conducted with *metafor* in *R* (Viechtbauer, 2010).
28
29
30

31 32 **Results**

33 34 35 *Study characteristics*

36
37 Table 1 provides a systematic overview of the studies' characteristics. All but two studies
38
39 were published in academic journals. Ten studies were conducted in the United States, one in
40
41 Germany, and one involved trainees from multiple countries. Only two interventions were
42
43 developed based on a needs analysis. Moreover, nine studies measured behavior in the
44
45 workplace solely by relying on trainee self-reports.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table I. *Characteristics of the included studies*

Journal of Workplace Learning

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Overall transfer effectiveness of diversity training

The analysis revealed a small statistically significant positive effect of DT on trainees' workplace behavior, $g = 0.29$, 95% CI [0.11–0.47]. The observed outcomes ranged from -0.32 to 0.79 . Figure 2 shows a forest plot illustrating the observed outcomes and the estimate based on the random-effects model.

Figure 2. *Forest plot*

According to the Q-test, which is a statistical test for the estimation of heterogeneity, the true outcomes appeared to be heterogeneous ($Q(11) = 38.86$, $p < 0.0001$). We can thus reject the null hypothesis that the true effect size is the same across all studies. The I^2 statistic was 69.16%, meaning that 69.16% of the variance was due to true effects rather than sampling error.

Publication bias

Analyses were conducted to assess whether the results were influenced by publication bias. First, a funnel plot of the estimates was constructed (Figure 3).

Figure 3. *Funnel plot*

However, visual interpretation of a funnel plot tends to be subjective and unreliable; therefore, several methods have been developed to quantify publication bias in a funnel plot (Borenstein *et al.*, 2021). Neither the rank correlation nor the regression test indicated asymmetry in the funnel plot ($p = 0.74$ and $p = 0.30$, respectively).

Moderator analysis

Heterogeneity analysis revealed inconsistencies in the effects of DT on workplace behavior. Thus, characteristics of the training design were analyzed to determine whether they influenced the effectiveness of the training.

As shown in Table 2, no significant subgroup effect was found for the potential moderator needs analysis ($F = 0.04$, $p = .837$). The covariate distribution for this subgroup

1
2
3 analysis was uneven, with only two studies having conducted a needs analysis. We also found
4
5 no significant difference between subgroups as a result of attendance policy ($F = 0.88$,
6
7 $p = .374$). Regarding the third moderator, there was a significant subgroup difference
8
9 ($F = 6.75$, $p = .027$), indicating that goal setting/action planning may impact the transfer
10
11 success of DT.
12
13

14
15 In addition to examining the empirically derived moderators, we investigated whether
16
17 study design characteristics impacted the outcomes of the studies. First, we analyzed whether
18
19 it made a difference who reported the workplace behavior. Analyses indicated that studies
20
21 that relied entirely on self-reports yielded greater effect sizes than studies that used other
22
23 measurements ($F = 16.03$, $p = .003$). Furthermore, different study designs may have different
24
25 sources of bias, so a moderation analysis with study design as a moderator was conducted.
26
27 However, the moderating effect of the study design was not significant ($F = 2.28$, $p = .165$).
28
29 Finally, in controlled studies, we compared outcomes between those studies that used random
30
31 assignment of trainees and those that did not. Our analysis revealed no significant moderation
32
33 effect of randomization ($F = 0.43$, $p = .667$).
34
35
36

37
38 **Table 2.** *Moderation analysis*
39
40

41 **Discussion**

42
43

44 This study investigated (1) whether DT has a positive effect on the workplace behavior of
45
46 trainees and whether (2) needs analysis, (3) attendance policy and (4) goal setting/action
47
48 planning moderate the transfer of DT. Our analysis revealed a small but significant effect of
49
50 DT on trainees' workplace behavior. The finding that this effect size is smaller than the effect
51
52 sizes reported in previous meta-analyses that examined the learning outcomes of DT concurs
53
54 with previous findings that learning does not necessarily lead to transfer (Alliger *et al.*, 1997;
55
56 Blume *et al.*, 2010).
57
58
59
60

1
2
3 Since our results are based on heterogeneous data, we investigated whether specific
4 intervention features moderated the effectiveness of DT. We found no significant effect for
5 needs analysis. However, the covariate distribution was largely uneven in that analysis, which
6 is why the results should be interpreted with caution.
7
8
9

10
11 Another subgroup analysis revealed no impact of attendance policy on the
12 effectiveness of DT. This result is consistent with previous meta-analyses that examined
13 whether attendance requirements impacted learning outcomes after DT (Bezrukova *et al.*,
14 2016; Kalinoski *et al.*, 2013). The respective advantages of the different attendance policies
15 may counterbalance each other, resulting in the absence of significant subgroup differences.
16
17
18
19
20
21
22
23

24 Concerning our third proposed moderator, we found that goal setting/action planning
25 is significantly related to workplace behavior. This finding corroborates previous arguments
26 and empirical findings regarding the transfer-promoting function of such measures (Botke *et al.*
27 *et al.*, 2018; Burke and Hutchins, 2007; Foxon, 1994).
28
29
30
31
32

33 In summary, our meta-analysis allows for a cautiously positive interpretation of the
34 existing research regarding the transfer of DT. Taken together, the reviewed diversity
35 interventions had a small yet statistically significant effect on—mostly self-reported—
36 workplace behavior. In this context, goal setting/action planning seems to be an effective
37 measure to promote transfer.
38
39
40
41
42
43

44 *Limitations*

45
46 Our meta-analysis has several limitations. First, the number of included studies was rather
47 small. A common criticism is that DT is often only evaluated with respect to proximal
48 criteria, such as reactions and learning, with workplace behavior seldom being assessed
49 (Devine and Ash, 2022; Gill and Olson, 2023). The small number of studies reduces the
50 generalizability of our findings. Although moderation analyses are generally deemed
51 acceptable when there are at least ten studies in the meta-analysis (Deeks *et al.*, 2023), the
52
53
54
55
56
57
58
59
60

1
2
3 small sample sizes of studies per category reduce the analyses' power and confidence in the
4
5 conclusions, especially when covariates are distributed unevenly.
6
7

8 Second, the included studies had several limitations affecting internal validity. Nine
9
10 studies relied entirely on the self-reports of trainees. We found that studies using other (e.g.,
11
12 third-party) evaluations reported smaller effect sizes than those relying on self-reports. This
13
14 finding suggests that self-reports regarding such a sensitive topic as diversity may be inflated.
15
16 Kalinoski *et al.* (2013) reported similar results for self-reports regarding cognitive learning
17
18 outcomes after DT. Moreover, some studies did not use a control group, and others used
19
20 control groups without random allocation to the respective groups. We conducted moderation
21
22 analyses to assess whether these factors affected our results but found no statistically
23
24 significant subgroup differences.
25
26
27

28 Third, since all but two studies were conducted in the United States, the external
29
30 validity of our findings is limited. Kawasaki and Zou (2023), for instance, argued that
31
32 cultural values might impact the success of DT.
33
34

35 *Implications for research and practice*

36
37 Baldwin and Ford (1988) identified three areas influencing training transfer: training design,
38
39 work environment, and trainee characteristics. In our moderator analyses, we found support
40
41 for the relevance of the training design feature action planning/goal setting, which apparently
42
43 helps bridge the gap between training and the workplace. A theoretical implication of our
44
45 research could be that we should focus more strongly on strategies that link training directly
46
47 to workplace behavior, instead of treating both as two separate events. Training should
48
49 initiate tangible transfer actions and, building on this, enable further learning in the
50
51 workplace. Such strategies could generate positive transfer experiences and feedback, which
52
53 in turn strengthen trainees' transfer motivation and self-efficacy and lead to more transfer
54
55 (Blume *et al.*, 2019).
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In the future, more studies should evaluate not only the learning outcomes but also the transfer success of DT to increase the reliability of the reported results. Moreover, instead of self-reports of trainees, future DT evaluations should place greater emphasis on judgments of those groups that may potentially be discriminated against in organizations (Sanchez and Medkik, 2004). Generally, if self-reports are used, they should be accompanied by other evaluation sources. A multisource evaluation based on feedback from trainees, supervisors, subordinates, peers, and other affiliated groups would be ideal (Kirkpatrick, 1996).

This meta-analysis also has implications for practitioners. First, the findings confirm that DT can be an effective strategy for managing diversity in organizations. Our results further indicate that programs where trainees formulate goals and/or action plans might be more effective. Specifying when and how trainees plan to implement specific pro-diversity behavior might increase the likelihood of these behaviors being performed in the workplace. Therefore, practitioners should consider how they can incorporate these elements into their training. For instance, trainers can use the end of a training session to develop transfer goals or plans with trainees, thereby bridging the gap between training and the workplace. Additional strategies, such as homework assignments or technology-driven reminders, might further reinforce the connection between training and the job. Building on this, supervisors and peers can support transfer by providing feedback on goal attainment in the workplace.

Moreover, although we found no moderator effect for attendance policy, communication appears essential when practitioners choose to implement mandatory DT. Sanchez and Medkik (2004) reported negative workplace behavior following mandatory DT, partly because trainees perceived their participation as a punishment for prior misconduct—even though the training was required for all supervisors. To prevent such backlash, practitioners should clearly communicate the rationale for compulsory participation.

BIBLIOGRAPHY

References marked with an asterisk indicate studies included in the meta-analysis.

Alliger, G.M., Tannenbaum, S.I., Bennett, W., Traver, H. and Shotland, A. (1997), “A meta-analysis of the relations among training criteria”, *Personnel Psychology*, Vol. 50 No. 2, pp. 341–358, doi: 10.1111/j.1744-6570.1997.tb00911.x.

Amir-Behghadami, M. and Janati, A. (2020), “Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews”, *Emergency Medicine Journal*, Vol. 37 No. 6, 387, doi: 10.1136/emered-2020-209567.

Bailey, R.R. (2019), “Goal setting and action planning for health behavior change”, *American Journal of Lifestyle Medicine*, Vol. 13 No. 6, pp. 615–618, doi: 10.1177/1559827617729634.

Baldwin, T.T. and Ford, J.K. (1988), “Transfer of training: A review and directions for future research”, *Personnel Psychology*, Vol. 41 No. 1, pp. 63–105, doi: 10.1111/j.1744-6570.1988.tb00632.x.

Becker, B.J. (1988), “Synthesizing standardized mean-change measures”, *British Journal of Mathematical and Statistical Psychology*, Vol. 41 No. 2, pp. 257–278, doi: 10.1111/j.2044-8317.1988.tb00901.x.

Bendick, M., Egan, M. and Lofhjelm, S. (2001), “Workforce diversity training: From anti-discrimination compliance to organizational development”, *Human Resource Planning*, Vol. 24 No. 2, pp. 10–25.

Bezrukova, K., Spell, C. and Perry, J. (2024), “Organizational diversity training programs”, *Current Opinion in Psychology*, Vol. 60, 101907, doi: 10.1016/j.copsyc.2024.101907.

Bezrukova, K., Spell, C.S., Perry, J.L. and Jehn, K.A. (2016), “A meta-analytical integration of over 40 years of research on diversity training evaluation”, *Psychological Bulletin*, Vol. 142 No. 11, pp. 1227–1274, doi: 10.1037/bul0000067.

1
2
3 Blume, B.D., Ford, J.K., Baldwin, T.T. and Huang, J.L. (2010), “Transfer of training: A
4 meta-analytic review”, *Journal of Management*, Vol. 36 No. 4, pp. 1065–1105, doi:
5
6 10.1177/0149206309352880.
7

8
9 Blume, B.D., Ford, J.K., Surface, E.A. and Olenick, J. (2019), “A dynamic model of training
10 transfer”, *Human Resource Management Review*, Vol. 29 No. 2, pp. 270–283, doi:
11
12 10.1016/j.hrmr.2017.11.004
13

14
15 Boere, N.A., de Jong, B., Jansen in de Wal, J. and Cornelissen, F. (2023), “Does training
16 content matter? Differences between soft- and hard-skill trainings in transfer motivation”,
17
18 *Journal of Workplace Learning*, Vol. 35 No. 9, pp. 274–290, doi: 10.1108/JWL-03-2023-
19
20 0046.
21

22
23 Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. (2021), *Introduction to*
24
25 *Meta-Analysis*, 2nd ed., Wiley, Hoboken, NJ.
26

27
28 Botke, J.A., Jansen, P.G.W., Khapova, S.N. and Tims, M. (2018), “Work factors influencing
29
30 the transfer stages of soft skills training: A literature review”, *Educational Research Review*,
31
32 Vol. 24, pp. 130–147, doi: 10.1016/j.edurev.2018.04.001.
33

34
35 Burke, L.A. and Hutchins, H.M. (2007), “Training transfer: An integrative literature review”,
36
37 *Human Resource Development Review*, Vol. 6 No. 3, pp. 263–296, doi:
38
39 10.1177/1534484307303035.
40

41
42 *Carnes, M., Devine, P.G., Baier Manwell, L., Byars-Winston, A., Fine, E., Ford, C.E.,
43
44 Forscher, P., *et al.* (2015), “The effect of an intervention to break the gender bias habit for
45
46 faculty at one institution: A cluster randomized, controlled trial”, *Academic Medicine*, Vol.
47
48 90 No. 2, pp. 221–230, doi: 10.1097/ACM.0000000000000552.
49

50
51 *Carnes, M., Sheridan, J., Fine, E., Lee, Y.-G. and Filut, A. (2023), “Effect of a workshop to
52
53 break the bias habit for internal medicine faculty: A multisite cluster randomized controlled
54
55 study”, *Academic Medicine*, Vol. 98 No. 10, pp. 1211–1219, doi:
56
57
58
59
60

1
2
3 10.1097/ACM.00000000000005271.
4

5 *Chang, E.H., Milkman, K.L., Gromet, D.M., Rebele, R.W., Massey, C., Duckworth, A.L.
6 and Grant, A.M. (2019), “The mixed effects of online diversity training”, *Proceedings of the*
7
8 *National Academy of Sciences*, Vol. 116 No. 16, pp. 7778–7783, doi:
9
10 10.1073/pnas.1816076116.
11
12

13
14 Chaplin, M. and Dwan, K. (2023), “Cluster-randomized controlled trials: A tutorial”,
15
16 *Cochrane Evidence Synthesis and Methods*, Vol. 1 No. 7, e12024, doi: 10.1002/cesm.12024.
17
18

19 Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Routledge,
20
21 New York, NY, doi: 10.4324/9780203771587.
22

23
24 *Combs, G.M. and Luthans, F. (2007), “Diversity training: Analysis of the impact of self-
25
26 efficacy”, *Human Resource Development Quarterly*, Vol. 18 No. 1, pp. 91–120, doi:
27
28 10.1002/hrdq.1193.
29

30
31 Davies, A., Fidler, D. and Gorbis, M. (2011), “Future work skills 2020”, *Institute for the*
32
33 *Future for Apollo Research Institute*, available at:
34
35 [https://uqpn.uq.edu.au/files/203/LIBBY%20MARSHALL%20future_work_skills_2020_full_](https://uqpn.uq.edu.au/files/203/LIBBY%20MARSHALL%20future_work_skills_2020_full_research_report_final_1.pdf)
36
37 [research_report_final_1.pdf](https://uqpn.uq.edu.au/files/203/LIBBY%20MARSHALL%20future_work_skills_2020_full_research_report_final_1.pdf) (accessed 13 May 2024).
38
39

40 Deeks, J.J., Higgins, J.P.T. and Altman, D.G. (2023), “Chapter 10: Analysing data and
41
42 undertaking meta-analyses”, in Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li,
43
44 T., Page, M.J. and Welch, V.A. (Eds.), *Cochrane Handbook for Systematic Reviews of*
45
46 *Interventions*, Cochrane, available at: www.training.cochrane.org/handbook (accessed 10
47
48 December 2024).
49

50
51 *De Meuse, K.P., Hostager, T.J. and O’Neill, K. (2007), “A longitudinal evaluation of senior
52
53 managers’ perceptions and attitudes of a workplace diversity training program”, *Human*
54
55 *Resource Planning*, Vol. 30 No. 2, pp. 38–46.
56
57

58 Devine, P.G. and Ash, T.L. (2022), “Diversity training goals, limitations, and promise: A
59
60

1
2
3 review of the multidisciplinary literature”, *Annual Review of Psychology*, Vol. 73, pp. 403–
4
5 429, doi: 10.1146/annurev-psych-060221-122215.

6
7
8 Dobbin, F. and Kalev, A. (2018), “Why doesn’t diversity training work? The challenge for
9
10 Industry and Academia”, *Anthropology Now*, Vol. 10 No. 2, pp. 48–55, doi:
11
12 10.1080/19428200.2018.1493182.

13
14 *Filmer, T. and Herbig, B. (2020), “A training intervention for home care nurses in
15
16 cross-cultural communication: An evaluation study of changes in attitudes, knowledge and
17
18 behaviour”, *Journal of Advanced Nursing*, Vol. 76 No. 1, pp. 147–162, doi:
19
20 10.1111/jan.14133.

21
22
23
24 Foxon, M. (1994), “A process approach to the transfer of training: Part 2: Using action
25
26 planning to facilitate the transfer of training”, *Australian Journal of Educational Technology*,
27
28 Vol. 10 No. 1, pp. 1–18.

29
30
31 Gardenswartz, L. and Rowe, A. (1998), *Managing Diversity: A Complete Desk Reference*
32
33 *and Planning Guide*, McGraw-Hill, New York, NY.

34
35
36 Gill, L.J. and Olson, M.A. (2023), “A dual-process framework for diversity training to reduce
37
38 discrimination in organizational settings”, *Social Issues and Policy Review*, Vol. 17 No. 1,
39
40 pp. 79–103, doi: 10.1111/sipr.12094.

41
42
43 Groenewald, E. (2024), “Inclusive Workplaces: The Key Strategies for Sustainable Diversity
44
45 Practices”, *International Multidisciplinary Journal of Research for Innovation,*
46
47 *Sustainability, and Excellence*, Vol. 1 No. 1, pp. 156–161.

48
49
50 Hamzah, H.A., Marcinko, A.J., Stephens, B. and Weick, M. (2024), “Making soft skills
51
52 ‘stick’: A systematic scoping review and integrated training transfer framework grounded in
53
54 behavioural science”, *European Journal of Work and Organizational Psychology*, pp. 1–14,
55
56 doi: 10.1080/1359432X.2024.2376909.

57
58 *Hanover, J.M.B. and Cellar, D.F. (1998), “Environmental factors and the effectiveness of
59
60

workforce diversity training”, *Human Resource Development Quarterly*, Vol. 9 No. 2, pp. 105–124, doi: 10.1002/hrdq.3920090203.

Hedges, L.V. and Olkin, I. (1985), *Statistical Methods for Meta-Analysis*, Academic Press, San Diego, CA.

IntHout, J., Ioannidis, J.P. and Borm, G.F. (2014), “The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method”, *BMC Medical Research Methodology*, Vol. 14, 25, doi: 10.1186/1471-2288-14-25.

Kalinoski, Z.T., Steele-Johnson, D., Peyton, E.J., Leas, K.A., Steinke, J. and Bowling, N.A. (2013), “A meta-analytic evaluation of diversity training outcomes”, *Journal of Organizational Behavior*, Vol. 34 No. 8, pp. 1076–1104, doi: 10.1002/job.1839.

Kawasaki, S. and Zou, X. (2023), “A meta-analytic evaluation of diversity training outcomes across cultures”, *Social and Personality Psychology Compass*, Vol. 17 No. 5, e12741, doi: 10.1111/spc3.12741.

Kirkland, R. and Bohnet, I. (2017), “Focusing on what works for workplace diversity”, available at: <https://www.mckinsey.com/featured-insights/gender-equality/focusing-on-what-works-for-workplace-diversity> (accessed 27 May 2023).

Kirkpatrick, D.L. (1996), “Great ideas revisited. Techniques for evaluating training programs. Revisiting Kirkpatrick’s Four-Level Model.”, *Training and Development*, Vol. 50 No. 1, pp. 54–59.

*Kossek, E.E., Buzzanell, P.M., Wright, B.J., Batz-Barbarich, C., Moors, A.C., Sullivan, C., Kokini, K., *et al.* (2024), “Implementing diversity training targeting faculty microaggressions and inclusion: Practical insights and initial findings”, *The Journal of Applied Behavioral Science*, p. 002188632211323, doi: 10.1177/00218863221132321.

Kraiger, K., Ford, K.J. and Salas, E. (1993), “Application of cognitive, skill-based, and

1
2
3 affective theories of learning outcomes to new methods of training evaluation”, *Journal of*
4
5
6 *Applied Psychology*, Vol. 78 No. 2, pp. 311–328, doi: 10.1037/0021-9010.78.2.311.

7
8 Kulik, C.T., Pepper, M.B., Roberson, L. and Parker, S.K. (2007), “The rich get richer:
9
10 predicting participation in voluntary diversity training”, *Journal of Organizational Behavior*,
11
12 Vol. 28 No. 6, pp. 753–769, doi: 10.1002/job.444.

13
14 Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L. and Salas, E. (2017), “Leadership
15
16 training design, delivery, and implementation: A meta-analysis.”, *Journal of Applied*
17
18 *Psychology*, Vol. 102 No. 12, pp. 1686–1718, doi: 10.1037/apl0000241.

19
20
21 Laker, D.R. and Powell, J.L. (2011), “The differences between hard and soft skills and their
22
23 relative impact on training transfer”, *Human Resource Development Quarterly*, Vol. 22 No.
24
25 1, pp. 111–122, doi: 10.1002/hrdq.20063.

26
27
28 *Law, D.Y. (1998), *An Evaluation of a Cultural Diversity Training Program*, Auburn
29
30 University, Auburn, AL.

31
32
33 Leslie, L.M., Kim, Y.L. and Ye, E.R. (2025), “Diversity initiatives: Intended and unintended
34
35 effects”, *Current Opinion in Psychology*, Vol. 61, 101942, doi:
36
37 10.1016/j.copsy.2024.101942.

38
39
40 *Lichtenstein, D., Lindstrom, L. and Povenmire-Kirk, T. (2008), “Promoting multicultural
41
42 competence: Diversity training for transition professionals”, *The Journal for Vocational*
43
44 *Special Needs Education*, Vol. 30 No. 3, pp. 3–15.

45
46
47 *Mattingly, V.P. (2018), *Glass Kickers: Training Men as Allies to Promote Women in*
48
49 *Leadership*, Colorado State University, Fort Collins, CO.

50
51
52 Morris, S.B. and DeShon, R.P. (2002), “Combining effect size estimates in meta-analysis
53
54 with repeated measures and independent-groups designs.”, *Psychological Methods*, Vol. 7
55
56 No. 1, pp. 105–125, doi: 10.1037/1082-989X.7.1.105.

57
58
59 Ouzzani, M., Hammady, H., Fedorowicz, Z. and Elmagarmid, A. (2016), “Rayyan—a web
60

1
2
3 and mobile app for systematic reviews”, *Systematic Reviews*, Vol. 5, 210, doi:
4
5 10.1186/s13643-016-0384-4.
6

7
8 Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D.,
9
10 Shamseer, L., *et al.* (2021), “The PRISMA 2020 statement: an updated guideline for
11
12 reporting systematic reviews”, *BMJ*, Vol. 372, n71, doi: 10.1136/bmj.n71.
13

14
15 Pendry, L.F., Driscoll, D.M. and Field, S.C.T. (2007), “Diversity training: Putting theory into
16
17 practice”, *Journal of Occupational and Organizational Psychology*, Vol. 80 No. 1, pp. 27–
18
19 50, doi: 10.1348/096317906X118397.
20

21
22 Pustejovsky, J.E. and Tipton, E. (2022), “Meta-analysis with robust variance estimation:
23
24 Expanding the range of working models”, *Prevention Science*, Vol. 23 No. 3, pp. 425–438,
25
26 doi: 10.1007/s11121-021-01246-3.
27

28
29 Roberson, Q.M., Moore, O.A. and Bell, B.S. (2024), “An active learning approach to
30
31 diversity training”, *Academy of Management Review*, Vol. 49 No. 2, pp. 344–365, doi:
32
33 10.5465/amr.2020.0231.
34

35
36 Rubio-Aparicio, M., López-López, J.A., Viechtbauer, W., Marín-Martínez, F., Botella, J. and
37
38 Sánchez-Meca, J. (2020), “Testing categorical moderators in mixed-effects meta-analysis in
39
40 the presence of heteroscedasticity”, *The Journal of Experimental Education*, Vol. 88 No. 2,
41
42 pp. 288–310, doi: 10.1080/00220973.2018.1561404.
43

44
45 Salamon, J., Blume, B.D., Orosz, G. and Nagy, T. (2021), “The interplay between the level of
46
47 voluntary participation and supervisor support on trainee motivation and transfer”, *Human*
48
49 *Resource Development Quarterly*, Vol. 32 No. 4, pp. 459–481, doi: 10.1002/hrdq.21428.
50

51
52 Salamon, J., Blume, B.D., Orosz, G. and Nagy, T. (2022), “The moderating effect of
53
54 coworkers’ training participation on the influence of peer support in the transfer process”,
55
56 *European Journal of Training and Development*, Vol. 47 No. 10, pp. 15–36, doi:
57
58 10.1108/EJTD-07-2021-0102.
59
60

1
2
3 *Sanchez, J.I. and Medkik, N. (2004), “The effects of diversity awareness training on
4 differential treatment”, *Group & Organization Management*, Vol. 29 No. 4, pp. 517–536,
5
6 doi: 10.1177/1059601103257426.
7
8

9
10 Viechtbauer, W. (2010), “Conducting meta-analyses in R with the metafor package”, *Journal*
11
12 *of Statistical Software*, Vol. 36 No. 3, pp. 1–48, doi: 10.18637/jss.v036.i03.
13

14 Wang, M.L., Gomes, A., Rosa, M., Copeland, P. and Santana, V.J. (2024), “A systematic
15
16 review of diversity, equity, and inclusion and antiracism training studies: Findings and future
17
18 directions”, *Translational Behavioral Medicine*, Vol. 14 No. 3, pp. 156–171, doi:
19
20 10.1093/tbm/ibad061.
21
22

23
24 Wentling, R.M. and Palma-Rivas, N. (1999), “Components of effective diversity training
25
26 programmes”, *International Journal of Training and Development*, Vol. 3 No. 3, pp. 215–
27
28 226, doi: 10.1111/1468-2419.00079.
29

30
31 Note: The search process for the literature included in the meta-analysis was adapted from
32
33 parts of the first author's master's thesis.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60