

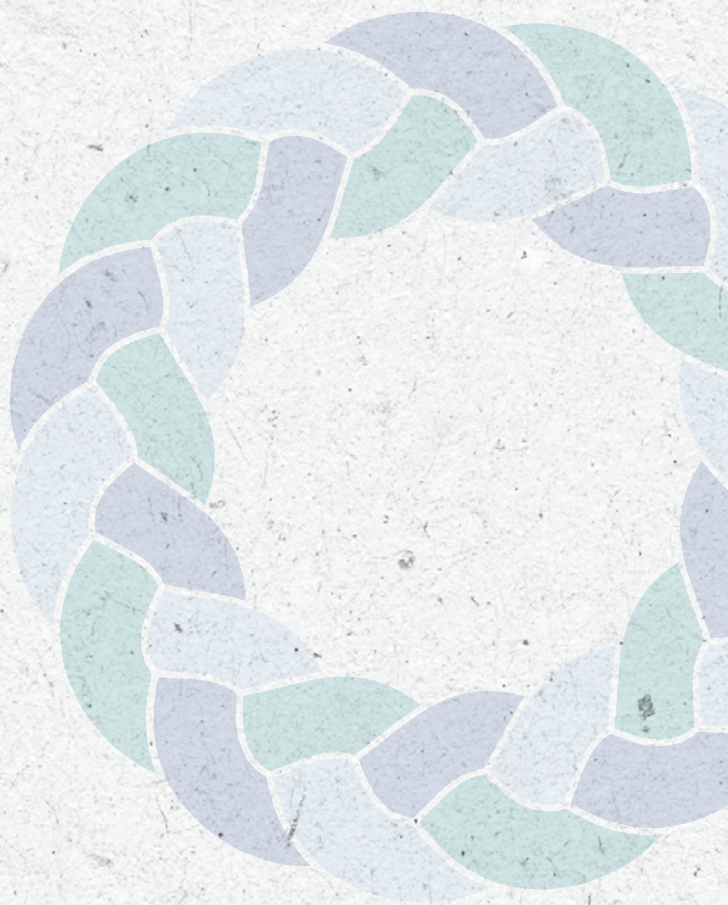
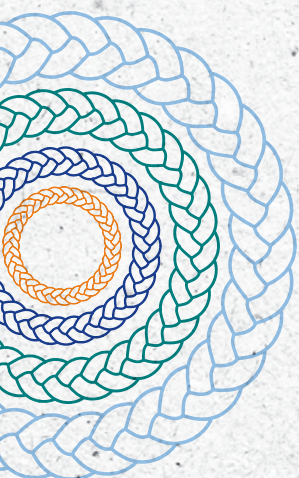
Rapid Evidence Assessment (REA)



THE EFFECTIVENESS OF LEADERSHIP TRAINING

a summary of research literature

January 2020



Culture Review **Implementation**
our journey of positive change



ACT
Government

This REA was produced by the Center for Evidence Based Management (CEBMA). The ACT Government acknowledges and thanks the CEBMA for allowing ACT Health to reproduce and redesign the content of their REA.

Any enquiries in relation to the content of this REA should be directed to CEBMA through their website: www.cebma.org

Acknowledgement of Country

ACT Health Directorate acknowledges the Traditional Custodians of the land, the Ngunnawal people. The Directorate respects their continuing culture and connections to the land and the unique contributions they make to the life of this area. It also acknowledges and welcomes Aboriginal and Torres Strait Islander peoples who are part of the community we serve.

Accessibility

The ACT Government is committed to making its information, services, events and venues as accessible as possible.

If you have difficulty reading a standard printed document and would like to receive this publication in an alternative format such as large print, please phone 13 22 81 or email HealthACT@act.gov.au



If English is not your first language and you require a translating and interpreting service, please phone Access Canberra on 13 22 81.

If you are deaf, or have a speech or hearing impairment and need the teletypewriter service, please phone 13 36 77 and ask for 13 22 81.

For speak and listen users, please phone 1300 555 727 and ask for 13 22 81. For more information on these services visit www.relayservice.com.au

© Australian Capital Territory, Canberra, July 2020.

This work is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without written permission from the Territory Records Office, ACT Government, GPO Box 158, Canberra City ACT 2601.

Enquiries about this publication should be directed to the ACT Health Directorate, Communications and Government Relations, GPO Box 825, Canberra City ACT 2601.

www.health.act.gov.au | www.act.gov.au

Enquiries: Canberra 13ACT1 or 13 22 81

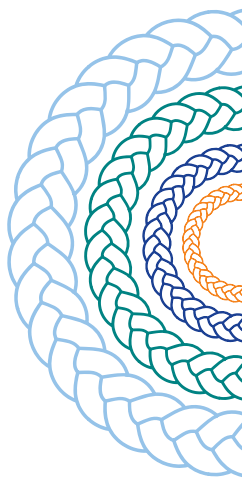
Contents

Background	4
.....	
What is a Rapid Evidence Assessment (REA)?	5
.....	
Main question: What does this REA answer?	6
.....	
Search strategy: How was the research evidence sought?	6
.....	
Selection process: How were the studies selected?	7
.....	
Critical appraisal: How were the quality of the included studies judged?	8
.....	
Outcome of the critical appraisal	9
.....	
Main findings	10
.....	
Conclusion	15
.....	
Limitations	15
.....	
References	16
.....	
Appendices	18
.....	



Background

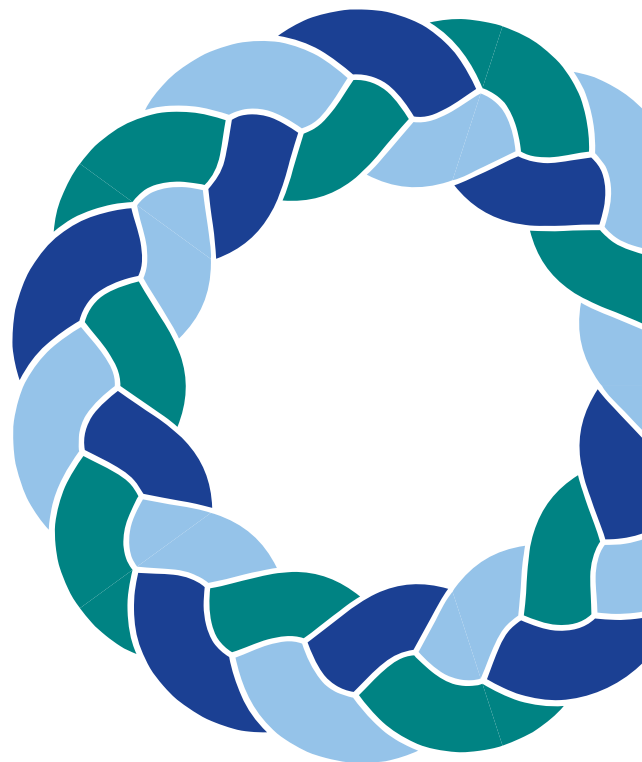
In 2013, a group of Dutch hospitals sought to review the effectiveness of their leadership training programs that were offered to their managers and physician executives. For this reason, the hospitals commissioned the Center for Evidence Based Management (CEBMa) to conduct a Rapid Evidence Assessment (REA) of the scientific literature on the attributes of effective leadership training programs. In 2019, CEBMa was approached by the ACT public health system to conduct an update of this review.

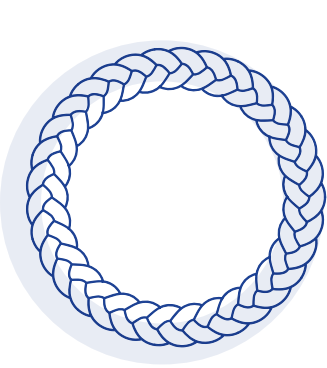


What is a Rapid Evidence Assessment (REA)?

Evidence reviews come in many forms. One of the best-known types is the conventional literature review, which provides an overview of the relevant scientific literature published on a topic. However, a conventional literature review's trustworthiness is often low: clear criteria for inclusion are often lacking and studies are selected based on the researcher's individual preferences. As a result conventional literature reviews are prone to severe bias. This is why 'rapid evidence assessments' (REAs) are being used.

This type of review is a specific research methodology that aims to identify the most relevant studies on a specific topic as comprehensively as possible, and to select appropriate studies based on explicit criteria. In addition, the methodological quality of the studies included is assessed by two independent reviewers on the basis of explicit criteria. In contrast to a conventional literature review, a REA is transparent, verifiable, and reproducible, and, as a result, the likelihood of bias is considerably smaller.



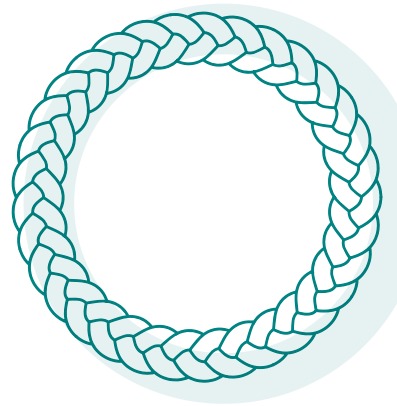


Main question: What does this REA answer?

What is known in the scientific literature about the effectiveness of leadership training programs? In addition, what are the characteristics of effective leadership training programs?

Supplementary questions

1. **What is leadership?**
2. **How can leadership be measured?**
3. **What is the effect of leadership training?**
4. **What are the characteristics of effective leadership training programs?**

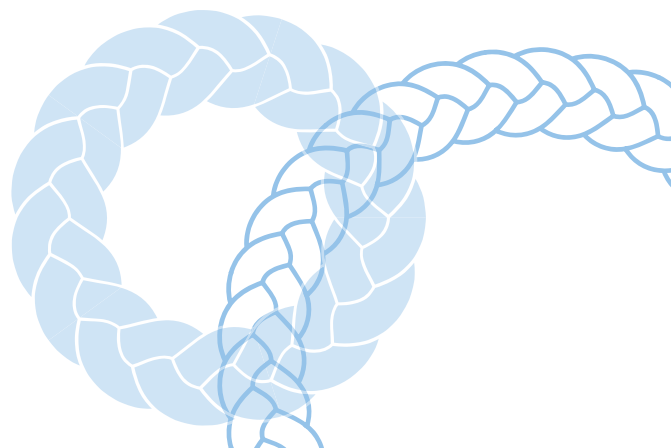


Search strategy: How was the research evidence sought?

The following four databases were used to identify studies: ABI/INFORM Global from ProQuest; Business Source Premier from EBSCO; PsycINFO from Ovid; and, Web of Science from Thomson Reuters. The following generic search filters were applied to all databases during the search:

1. **Scholarly journals, peer-reviewed.**
2. **Published in the period 1995 to 2019.**
3. **Articles in English.**

A search was conducted using combinations of different search terms, such as 'leadership', 'training', and 'course'. We conducted 53 different search queries and screened the titles and abstracts of more than 700 studies. All queries, criteria and search results are described in detail in Appendix 1.



Selection process: How were the studies selected?

Study selection took place in two phases. First, the titles and abstracts of the 700+ studies were screened for their relevance to this REA. In case of doubt, lack of information, or disagreement, the study was included. Duplicate publications were removed. This first phase yielded 25 (2013) and 21 (2019) studies. Secondly, studies were selected based on the full text of the article according to the following inclusion criteria:

1. Only meta-analyses, systematic reviews, or controlled/longitudinal studies.
2. Only studies in which the association with (knowledge worker) performance were measured.
3. Only studies related to workplace settings.
4. Only studies that were graded level C or above (see below).

Two reviewers worked independently to identify which studies would be included. Where the reviewers disagreed on inclusion, a third reviewer assessed whether the study was appropriate for inclusion with no prior knowledge of the initial reviewers' assessments. The decision of the third reviewer was final. This second phase yielded 17 (2013) and 14 (2019) studies, making a total of 31. An overview of the selection procedure is provided in Appendix II.





Critical appraisal: How were the quality of the included studies judged?

Methodological Appropriateness

In almost any situation it is possible to find a scientific study to support or refute a theory or a claim, and sometimes to quite a large degree. It is therefore important to determine which studies are trustworthy (i.e. valid and reliable) and which are not. The trustworthiness of a scientific study is first determined by its methodological appropriateness. For cause-and-effect claims (i.e. if we do A, will it result in B?) a study has a high methodological appropriateness when it fulfils the three conditions required for causal inference: co-variation, time-order relationship, and elimination of plausible alternative causes (Shaughnessy & Zechmeister, 2006).

A study that uses a control group, random assignment and a before-and-after measurement is therefore regarded as the 'gold standard'. Non-randomised studies and before-after studies come next in terms of appropriateness. Cross-sectional studies (surveys) and case studies are regarded as having the greatest chance of showing bias in the outcome and therefore sit lower down in the ranking in terms of appropriateness. Meta-analyses in which statistical analysis techniques are used to pool the results of controlled studies are therefore regarded as the most appropriate design.

To determine the methodological appropriateness of the included studies' research design, the classification system of Shadish, Cook and Campbell (2002), and Petticrew and Roberts (2006) was used. The following four levels of appropriateness were used for the classification:

Design	Level
Systematic review or meta-analysis of randomised controlled studies	AA
Systematic review or meta-analysis of controlled and/or before-after studies	A
Randomised controlled studies	
Systematic review or meta-analysis of cross-sectional studies	B
Non-randomised controlled before-after studies	
Interrupted time series	
Controlled studies without a pretest or uncontrolled studies with a pretest	C
Cross-sectional studies	D

It should be noted, however, that the level of methodological appropriateness as explained above is only relevant in assessing the validity of a cause-and-effect relationship that might exist between a predictor/driver (organisational culture) and its outcomes (performance), which is the purpose of this review.

Methodological Quality

In addition, a study's trustworthiness is determined by its methodological quality (its strengths and weaknesses). For instance, was the sample size large enough and were reliable measurement methods used? To determine methodological quality, all the studies included were systematically assessed on explicit quality criteria. Based on a tally of the number of weaknesses, the trustworthiness was downgraded and the final level was determined as follows: a downgrade of one level if two weaknesses were identified; a downgrade of two levels if four weaknesses were identified, etc.

Effect sizes

Finally, the effect sizes were identified. An effect (e.g. a correlation, Cohen's d or ω) can be statistically significant but may not necessarily be of practical relevance: even a trivial effect can be statistically significant if the sample size is big enough. For this reason, the effect size – a standard measure of the magnitude of the effect – of the studies included was assessed. To determine the magnitude of an effect, Cohen's rules of thumb (Cohen, 1988) were applied. According to Cohen 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 one that anybody can easily see because it is substantial.

Outcome of the critical appraisal

The overall quality of the 31 studies included in this review is moderate to high: 18 studies were graded level B or higher. An overview of all studies is provided in Appendix III.

Main findings

1. What is leadership?

In the domain of management and organisations, the term, 'leadership' is a popular term. In management books as well as in the scientific literature the effect of leadership is widely discussed and studied. Amazon offers thousands of management books with 'leadership' in the title, and the research database ABI/INFORM contains more than 15,000 peer reviewed papers on this topic. In fact, both scholars and practitioners, view leadership as an important if not fundamental driver for organisational performance (Pfeffer, 2015).

Despite the large number of publications, however, there is no consensus of what 'leadership' entails. As a result, there are many definitions of leadership available. For example, the Oxford dictionary defines leadership as 'the ability to be a leader or the qualities a good leader should have', whereas Wikipedia refers to leadership as 'a process of social influence in which a person can enlist the aid and support of others in the accomplishment of a common task'. Sometimes the term 'leadership' is limited to 'personal influence resulting in enthusiastic commitment of followers' (Schyns, 2013). For example, some management thinkers simply define leadership as 'the capacity to translate vision into reality' (Warren Bennis), or 'empowering others' (Bill Gates).

In addition, there is an ongoing controversy regarding the distinction between leadership and management. Conventional wisdom has it that managers are concerned with how things get done, whereas leaders build commitment and vision (Kotter, 1990). However, as Yukl (1989) stated, 'Nobody has proposed that managing and leading are equivalent, but the degree of overlap is a point of sharp disagreement' (Collins, 2002). In this review, however, the terms 'leadership' and 'management' are used interchangeably.

The lack of a widely-accepted definition of leadership and what it entails complicates research on the effect of leadership training programs. Leadership training programs use a plethora of definitions and core attributes of leadership, which makes a robust and comparable evaluation of their impact difficult.



2. How can leadership be measured?

There are many assessment tools and questionnaires available that claim to measure leadership. Most of these tools focus on a specific element or type of leadership, such as the Leadership Style Questionnaire (Northouse, 2011) and the Servant Leadership Behaviour Scale (Sendjaya, 2008). A recent systematic review that evaluates the psychometric properties of 17 leadership scales found that the majority lack some degree of rigour. Partly for this reason, most studies included in this review don't measure leadership as such, but rather evaluate the impact (transfer) of leadership training programs by using Kirkpatrick's model (Collins, 2001; Fritch, 2015). Kirkpatrick's assessment model comprises four levels, presented as a sequence:

1. **Reaction:** what participants think and feel about the training program.
2. **Learning:** the resulting increase in knowledge, skills, and change attitudes.
3. **Behaviour:** the resulting change in practice and behaviour.
4. **Results:** the resulting organisational benefits (e.g. performance, service delivery, clinical outcomes).

3. What is the effectiveness of leadership training programs?

- Finding 1: Overall, leadership training has a moderate effect (Level AA)**
- Finding 2: The effectiveness of leadership training has somewhat improved over the past decades (Level A)**
- Finding 3: The effects on some outcomes seem to be stable over time (Level A)**

Overall, leadership training and development programs have a moderate positive effect on a wide range of outcomes (Au, 2005; Avolio, 2009; Collins, 2004; Eden, 2000; Lacerenza, 2017; Mesmer, 2010; Taylor, 2009). This result is largely replicated across studies irrespective of the source of evaluation (self-, superior-, peer-, or subordinate ratings) or study designs (meta-analysis and controlled studies). The effects on learning outcomes, however, tend to be larger than the effects on performance outcomes (Fritch, 2015; Taylor, 2005).

Surprisingly, although there have been numerous innovations in leadership theory and training techniques, the effectiveness of leadership training seems to have only slightly improved over the past 20 years (Lacerenza, 2017; Powel, 2010). Finally, a meta-analysis of randomised controlled studies showed that, although effects on declarative knowledge decayed over time, training effects on skills and job behaviour remained stable or even increased (Taylor, 2005).

Finding 4: The evidence on the economic return on investment of leadership training is unclear (Level B)

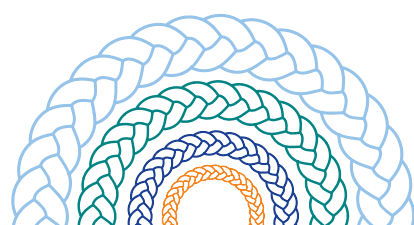
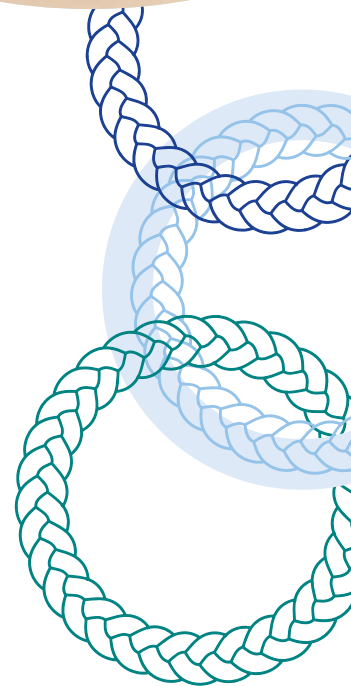
The studies included in this REA provide limited information on the economic utility or Return on Investment (ROI) of leadership training. A meta-analysis that attempted to calculate the ROI of leadership training, caution that the financial returns are unclear, especially for higher-level managers and leaders (Morrow, 1997). It is therefore recommended by some authors to limit the expenses through, for example, controlling additional expenses such as travel and accommodation, and avoiding expensive 'high-end' training.


Finding 5: The effectiveness of leadership training for nurses and physicians seems to be somewhat lower (Level C)

In relation to the target group for this review, being the health sector, meta-analyses that examined the effect of leadership training and development programs that aim to support nurses and physicians in leadership positions (e.g. Kiesewetter, 2013; Webb, 2014) showed lower effect sizes than meta-analyses that examined the effect of leadership training in general (e.g. Taylor, 2015). A possible explanation for this finding is provided by the outcome of a recent meta-analysis, indicating that leadership training programs for physicians and nurses show considerable heterogeneity concerning conceptual frameworks, teaching and learning methods, educational content, evaluation design, and outcomes measured (Frich, 2015; Lyons, 2018).

In addition, it was found the leadership programs described in the medical literature focus more on the 'know' and 'do' elements of leadership than the 'be' component, which some argue is fundamental in attaining the capacity to lead.

Finally, it was found that the concepts underlying leadership skills in medical training show little agreement as to what content should be emphasised in the development of leadership skill (Kiesewetter, 2013; Lyons, 2018), and is centred on imparting conceptual knowledge to physicians as individuals, for which lectures and seminars may be suitable, and has directed fewer resources to efforts in building self-awareness, for which action-based learning, feedback, and self-development activities may be more appropriate (see below).





4. What are the characteristics of effective leadership training programs?

The studies included in this review show a large variation of effectiveness of the training programs. These findings suggest that the effect of leadership training is moderated and/or mediated by several factors, such as delivery and implementation characteristics (Lacerenza, 2017; Morrow, 1997). Below is an overview of moderators and mediators.

Finding 6: Leadership training programs with the following characteristics tend to be more effective

6.1. Start with a 'training needs analysis' (Level B)

With regard to the training or learning content of leadership training programs, the studies included in this review indicate that a careful training needs analysis (e.g. through a survey, interview, focus group, or critical incidents) before the start of the program tends to increase its effectiveness (Lacerenza, 2017; Leskiw, 2017; Mesmer, 2010; Taylor, 2005). More specifically:

- a. Programs designed on the basis of an analysis of tasks and skill requirements and skills gaps are more effective than generic untailored programs, and
- b. Training need analyses have a positive impact on the motivation of trainees to learn, and
- c. Learning transfer is greater when trainees set clear and specific learning goals.

6.2. Focus on general management skills and interpersonal skills (Level B)

A meta-analysis of controlled studies (Taylor, 2009) indicates that skills that seem to transfer best to leader behaviour seem to be general management skills (e.g. goal setting, performance appraisal, time management) and, to a lesser extent, interpersonal skills (e.g. listening, questioning, negotiating, mentoring).

6.3. Duration and repetition (Level A)

In relation to design, several meta-analyses and high quality studies indicate that leadership training and development programs should be of reasonable length (at least 3 days or longer) and repeated periodically, to be effective (e.g. Lacerenza, 2017; Taylor, 2005). In addition, leadership training programs with spaced distribution are more effective than one-off trainings (Au, 2005). However, long and complex courses may not be required to achieve positive change; short (several days), punchy courses with clear objectives may well be as effective (Lyons, 2018).

6.4. Use multiple learning and instruction methods (Level A)

Several studies report that effective training programs are characterised by the use of a combination of didactic learning, tutorials and reflective learning (Lyons, 2018). In addition, these programs tend to use multiple instruction methods, such as lectures, group work, and action learning projects (Frich, 2015; Lacerenza, 2015; Steinert, 2012).

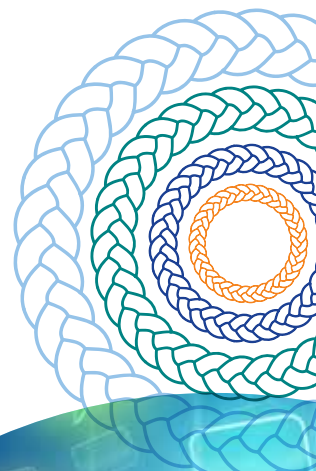
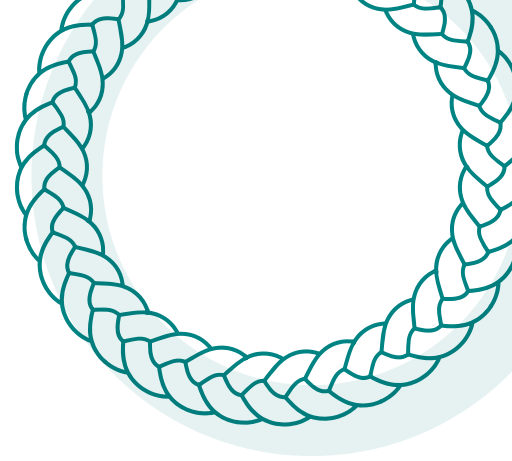
6.5. Provide opportunity to practice (Level A)

Although the use of multiple methods tend to be effective, the training program should clearly include opportunities for practice, linked to real world situations or trainee-generated scenarios (Frich, 2015; Lacerenza, 2015; Mianda, 2018; Steinert, 2012). This type of training is often referred to as 'experiential' or 'problem-based' learning.

In fact, in the domain of evidence-based education, there are a large number of studies that provide strong evidence that, in general, experiential learning tends to be more effective than traditional learning, especially when it concerns the teaching of practical skills and vocational knowledge. For example, several systematic reviews indicate that courses that include (simulations of) real-life situations yield better outcomes and have a larger (positive) effect on students' performance when compared with traditional teaching methods (Boet, 2014; Fung, 2015; Norman, 2012; Zelechowski, 2017).

6.6. Recognition and support (Level A)

In order for leadership training and development programs to be effective, it seems important, as is the case for other forms of training, to create and maintain a working climate that supports the transfer to the workplace of what was learned. Recognition, feedback, mentorship, and support from the organisation and the supervisor or peers for acquiring and applying new skills, can motivate trainees to transfer what they learned to their day-to-day work behaviour (Lacerenza, 2017; Leskiw, 2007; Steinert, 2012; Taylor, 2005). As such, leadership training should not be treated as a one-shot event. On the contrary, leadership development should be strongly and systematically integrated and communicated in the organisational culture and processes.



Conclusion

Leadership training programs are effective, but this effect is contingent on various design, delivery, and implementation characteristics. This REA supports the use of needs analysis, a focus on general management skills and interpersonal skills, multiple delivery methods, opportunity to practice, spaced training sessions, and organisational support.

Limitations

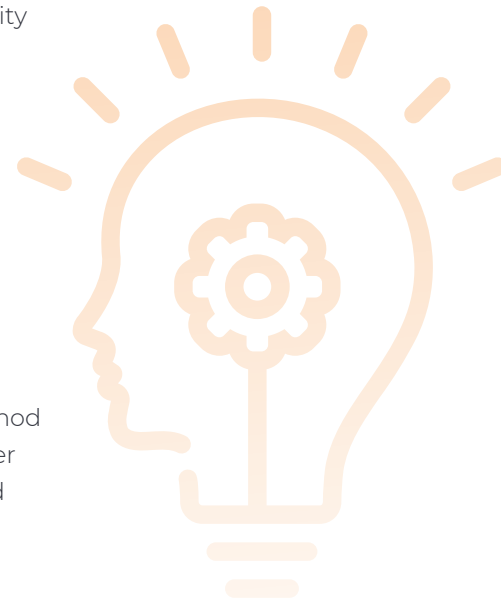
This REA aims to provide a balanced assessment of what is known in the scientific literature about the characteristics of effective leadership training programs, by using the systematic review method to search and critically appraise empirical studies. However, in order to be 'rapid', concessions were made in relation to the breadth and depth of the search process, such as the exclusion of unpublished studies, the use of a limited number of databases and a focus on meta-analyses and controlled studies published in the period 1995 to 2019. 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 REA focused only on quantitative studies, i.e. studies in which the link between leadership programs and performance/organisational outcomes was quantitatively measured. For this reason, findings from qualitative studies were not reported.

As a consequence, qualitative findings that are relevant for practice may have been missed.

Given these limitations, care must be taken not to present the findings presented in this REA as conclusive.



References

- Abdullah, D., & Suring, J. C. (2010, Dec 28-30). *The Relationship Between Motivation To Transfer, Training Design, Transfer Climate and Transfer of Training*. Paper presented at the International Conference on E-business, Management and Economics (ICEME 2010), Hong Kong
- Amagoh, F. (2009). Leadership development and leadership effectiveness. *Management Decision*, 47(5-6), 989-999.
- Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *Leadership Quarterly*, 20(5), 764-784.
- Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of Training: A Meta-Analytic Review. *Journal of Management*, 36(4),
- Ciucur, D., & Pirvut, A. F. (2012). The effects of a transactional analysis training programme on team leadership factors in automotive industry. *Psiworld* 2011, 33, 667-671.
- Cummings, G., Lee, H., MacGregor, T., Davey, M., Wong, C., Paul, L., & Stafford, E. (2008). Factors contributing to nursing leadership: a systematic review. *Journal of Health Services Research & Policy*, 13(4), 240-248.
- Day, D. V. (2000). Leadership development: A review in context. *Leadership Quarterly*, 11(4), 581-613.
- DeRue, D. S., Nahrgang, J. D., Hollenbeck, J. R., & Workman, K. (2012). A Quasi-Experimental Study of After-Event Reviews and Leadership Development. *Journal of Applied Psychology*, 97(5), 997-1015.
- Duygulu, S., & Kublay, G. (2011). Transformational leadership training programme for charge nurses. *Journal of Advanced Nursing*, 67(3), 633-642.
- Eden, D., Geller, D., Gewirtz, A., Gordon-Terner, R., Inbar, I., Liberman, M., . . . Shalit, M. (2000). Implanting Pygmalion Leadership Style through workshop training: Seven field experiments. *Leadership Quarterly*, 11(2), 171-210.
- Edwards, G., Elliott, C., Iszatt-White, M., & Schedlitzki, D. (2013). Critical and alternative approaches to leadership learning and development Introduction. *Management Learning*, 44(1), 3-10.
- Ely, K., Boyce, L. A., Nelson, J. K., Zaccaro, S. J., Hernez-Broome, G., & Whyman, W. (2010). Evaluating leadership coaching: A review and integrated framework. *Leadership Quarterly*, 21(4), 585-599.
- Gilbert, A., Hockey, P., Vaithianathan, R., Curzen, N., & Lees, P. (2012). Perceptions of junior doctors in the NHS about their training: results of a regional questionnaire. *Bmj Quality & Safety*, 21(3), 234-238.
- Gordon, M., Darbyshire, D., & Baker, P. (2012). Non-technical skills training to enhance patient safety: a systematic review. *Medical Education*, 46(11), 1042-1054.
- Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology*, 93(1), 59-69.
- Kelloway, E. K., & Barling, J. (2010). Leadership development as an intervention in occupational health psychology. *Work and Stress*, 24(3), 260-279.
- Krejci, J. W., & Malin, S. (1997). Impact of leadership development on competencies. *Nursing Economics*, 15(5), 235-8.
- Leslie, L. K., Miotto, M. B., Liu, G. C., Ziemnik, S., Cabrera, A. G., Calma, S., . . . Slaw, K. (2005). Training young pediatricians as leaders for the 21st century. *Pediatrics*, 115(3), 765-773.
- Mabey, C., & Gooderham, P. N. (2005). The impact of management development on perceptions of organisational performance in European firms. *European Management Review*, 2(2), 131-142.
- Malling, B., Mortensen, L., Bonderup, T., Scherpbier, A., & Ringsted, C. (2009). Combining a leadership course and multi- source feedback has no effect on leadership skills of leaders in postgraduate medical education. An intervention study with a control group. *Bmc Medical Education*, 9, 7.
- Markuns, J. F., Fraser, B., & Orlander, J. D. (2010). The Path to Physician Leadership in Community Health Centers: Implications for Training. *Family Medicine*, 42(6), 403-407.
- Morahan, P. S., Kasperbauer, D., McDade, S. A., Aschenbrener, C. A., Triolo, P. K., Monteleone, P. L., . . . Meyer, M. J. (1998). Training future leaders of academic medicine: Internal programs at three academic health centers. *Academic Medicine*, 73(11), 1159-1168.
- Morrow, C. C., Jarrett, M. Q., & Rupinski, M. T. (1997). An investigation of the effect and economic utility of corporate-wide training. *Personnel Psychology*, 50(1), 91-119.
- Powell, K. S., & Yalcin, S. (2010). Managerial training effectiveness A meta-analysis 1952-2002. *Personnel Review*, 39(1-2), 227-241.
- Serembus, J. F., Solecki, S., Meloy, F., & Olszewski, J. (2011). Preparing Tomorrow's Leaders A Leadership Course for Real-World Challenges. *Nurse Educator*, 36(3), 91-92.
- Solansky, S. T. (2010). The evaluation of two key leadership development program components: Leadership skills assessment and leadership mentoring. *Leadership Quarterly*, 21(4), 675-681.

Steinert, Y., Naismith, L., & Mann, K. (2012). Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review: BEME Guide No. 19. *Medical Teacher*, 34(6), 483-503.

Stoller, J. K. (2013). Commentary: Recommendations and Remaining Questions for Health Care Leadership Training Programs. *Academic Medicine*, 88(1), 12-15.

Storey, D. J. (2004). Exploring the link, among small firms, between management training and firm performance: a comparison between the UK and other OECD countries. *International Journal of Human Resource Management*, 15(1), 112-130.

Straus, S. E., Soobiah, C., & Levinson, W. (2013). The Impact of Leadership Training Programs on Physicians in Academic Medical Centers: A Systematic Review. *Academic Medicine*, 88(5), 710-723. doi: 10.1097/ACM.0b013e31828af493

Taylor, P. J., Russ-Eft, D. F., & Chan, D. W. L. (2005). A meta-analytic review of behaviour modeling training. *Journal of Applied Psychology*, 90(4), 692-709. doi: 10.1037/0021-9010.90.4.692

Taylor, P. J., Taylor, H., & Russ-Eft, D. F. (2009). Transfer of Management Training From Alternative Perspectives. *Journal of Applied Psychology*, 94(1), 104-121. doi: 10.1037/a0013006

Williams, A. P. (1973). Evaluation of management training - review - Whitelaw, M. *Occupational Psychology*, 47(1-2), 98-99.

Yeung, J. H. Y., Ong, G. J., Davies, R. P., Gao, F., & Perkins, G. D. (2012). Factors affecting team leadership skills and their relationship with quality of cardiopulmonary resuscitation. *Critical Care Medicine*, 40(9), 2617-2621.



Appendix I

Search terms & hits

Web of Science, ABI/Inform Global, Business Source Elite, Business Source Elite peer reviewed, scholarly journals, June 2013

Search terms	WSc	ABI	BSP
S1. ab(leader*) AND ab(training)	5,664	1,675	2,416
S2. ab(meta-anal*) AND S1	30	6	6
S3. ab("systematic review") AND S1	37	4	2
S4. ab("leader* training")	438	181	319
S5. ab("meta-anal*") AND S4	1	1	1
S6. ab(review) AND S4	36	12	17
S7. ab(leader*) AND ab(develop*)	23,453	9,734	12,515
S8. ab("meta-anal*") AND S7	90	16	17
S9. ab(review) AND S7	2,600	709	972
S10. ab("systematic review") AND S7	100	13	9
S11. ab("leader* develop*")	1,510	672	1,131
S12. ab("meta-anal*") AND S11	5	3	2
S13. ab(review) AND S11	108	52	122
S14: S11 AND filter WSc: management OR psychology applied OR health care sciences services OR nursing	53	-	-
S15. ab("systematic review") AND S11	5	1	0
S16. ab(leader*) AND ab(course*)	2,130	794	1,046
S17. ab("meta-anal*") AND S16	0	0	0
S18. ab(review) AND S16	215	45	186
S19. ab("systematic review") AND S16	4	0	0
S20. ab("leader* course")	39	14	25
S21. ab(manag*) AND ab(training)	44,064	8920	12,703
S22. ("meta-anal*") AND S21	280	20	23
S23. ab(review) AND S21	5,578	608	969

S24. ab("systematic review") AND 21	323	12	8
S25. ab("manag* training")	1,963	790	1,321
S26. ab("meta-anal*") AND S25	20	6	6
S27. ab(review) AND S25	156	48	120
S28. S25 AND filter: management OR psychology applied OR health care sciences services OR nursing	48	-	-
S29. ab("systematic review") AND S25	11	1	0
S30. ab(manag*) AND ab(develop*)	349,722	6,4872	4,639
S31. ab("meta-anal*") AND S30	1,395	92	38
S32. ab(review) AND S30	50,492	5,119	3,744
S33. ab("systematic review") AND S30	1,652	63	48
S34. ab("manag* develop*")	1,694	1,481	2,639
S35. ab("meta-anal*") AND S34	5	0	0
S36. ab(review) AND S34	120	128	744
S37. S34 AND filter: management OR psychology applied OR health care sciences services OR nursing	35	-	-
S38. ab("systematic review") AND S34	2	0	0
S39. ab(manag*) AND ab(course*)	38,501	4,319	6,830
S40. ab("meta-anal*") AND S39	178	2	5
S41. ab(review) AND S39	7,384	273	455
S42. ab("systematic review") AND S39	204	2	1
S43. ab("manag* course*")	652	193	285
S44. ab("meta-anal*") AND S43	6	0	0
S45. ab(review) AND S43	69	10	25
S46. ab("systematic review") AND S43	5	0	0
S47. ab("leader* training") AND ("health care" OR doctor* OR nurs* OR physician OR hospital)	77	13	35
S48. ab("manag* training") AND ab("health care" OR doctor* OR nurs* OR physician OR hospital)	271	173	191
S49. ab("meta-anal*") AND S48	3	0	0
Total	1,161	126	202

**ABI/Inform Global, Business Source Elite, PsycINFO,
ERIC and MEDLINE peer reviewed, scholarly journals, Nov 2019**

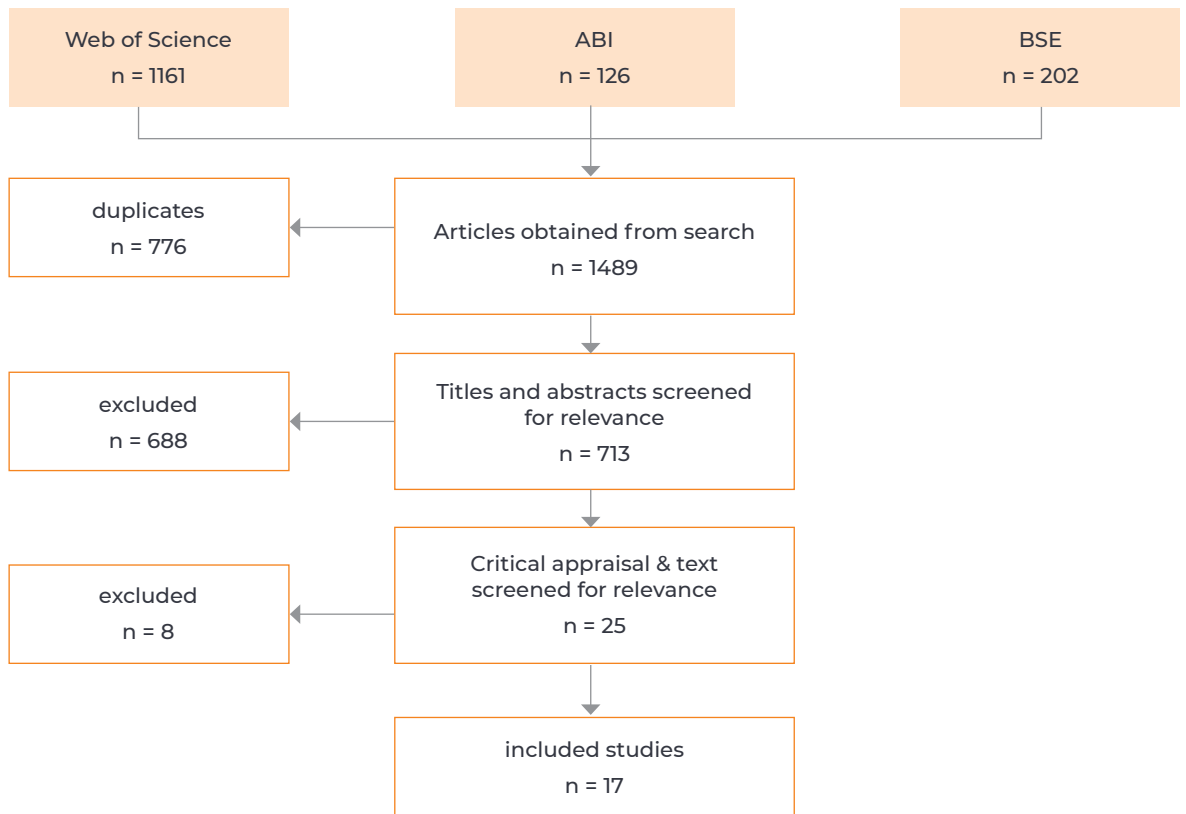
Search terms	ABI	BSP	PSY	ERIC	Medline
S1: ti(leader*) OR ti(manage*)	121,639	141,529	62,223	22,143	415,236
S2: ti(train*) OR ti(develop*)	95,185	123,695	172,020	76,214	707,490
S3: S1 AND S2	7,753	8,494	5,010	2,805	10,077
S4: S3 AND filter MAs, SRs or reviews	37	32	64	12	112*

* additional (database specific) filters were used

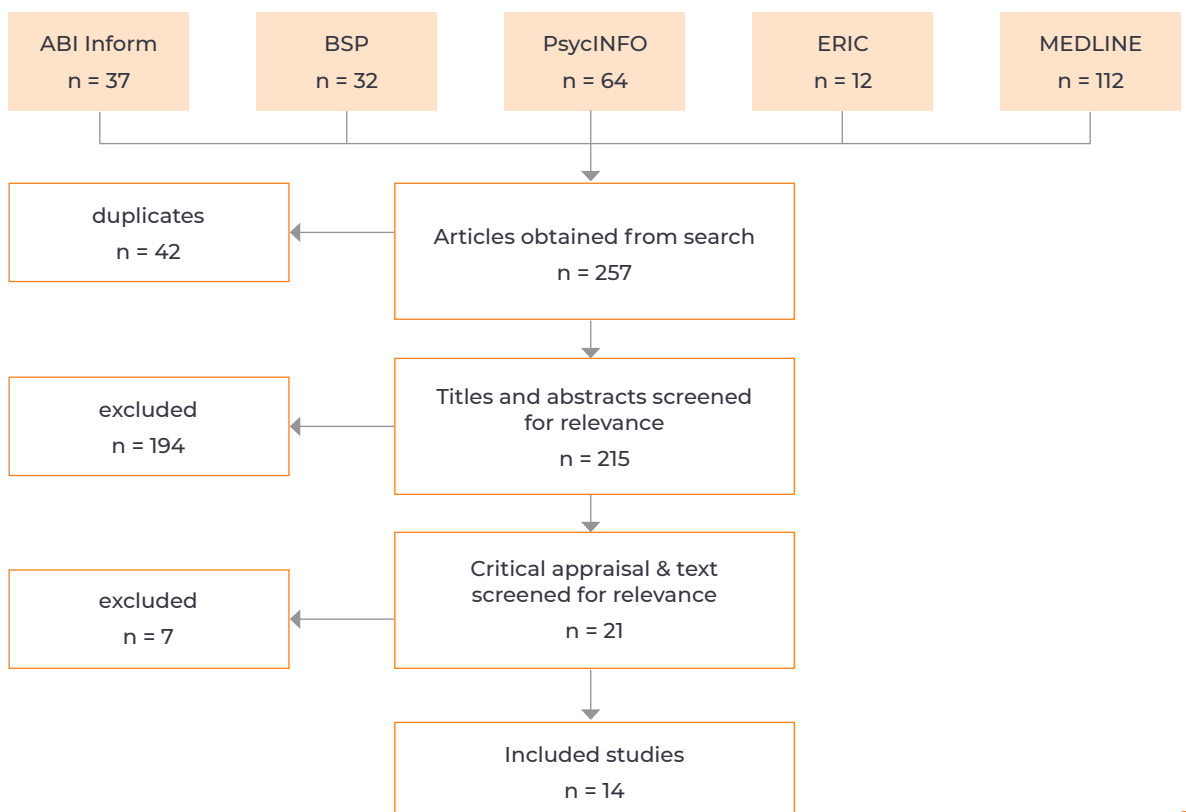
Appendix II

Selection of studies

Study selection, 2013



Study selection, 2020



Appendix III

Data extraction table - Leadership Training

Author & year	Design & sample size	Sector / Population	Main findings	Effect sizes	Limitations	Level
Abdullah, 2010	Cross-sectional study n = 120,	various, public, Malaysia	All of the transfer factors have a positive and significant relationship.	r = .53 to .59	Self-report, small sample, questionable validity of measures	D-
AU, 2005	Meta-analysis k = 58	Working population and vocational trainees	<ol style="list-style-type: none"> 1. There is a moderate effect size of overall leadership training programs. 2. Leadership training programs with a theoretical basis (a) were more effective, especially the application of Fiedler's contingency theory (b). 3. Leadership training programs with spaced distribution are more effective (a) than one-off trainings (b). 4. Leadership training programs using lower levels of a) evaluation criteria, b) subjective criteria, c) self-reported measurements and d) multiple-source evaluation criteria, as well as e) weak experimental designs, upwardly bias the results. 	1) Overall d = .64 2a) d = .74 2b) d = 1.10 3a) d = .65 3b) d = .50 4a) Reaction d = .82; Learning d = .71; Behaviour d = .40; Results d = .50 4b) Subjective only d = .68; Objective only d = .42. With objective d = .44 4c) Self d = .70; Others d = .65; Both d = .69 4d) Single d = .51; Multiple d = .96 4e) SGPP d = .80 POWC d = .68 PPWC d = .46 Random d = .37; Non-random d = .72	Quality of the included studies not assessed	A

Avolio, 2009	Meta- analysis, including experiments and quasi- experiments k = 13/28,	various, profit/ military	Training or development interventions have a moderate effect on 1) affective, 2) behavioural and 3) cognitive outcomes.	1. d = .39 2. d = .43 3. d = .62	Conclusions are not drawn on development interventions in particular, but on experimental leadership interventions in general.	AA
Ayeleke, 2019	systematic review including various low quality designs k = 19	healthcare	Quantitative studies (n = 3) showed that participation of health managers in a leadership development programme resulted in (some) improvement in leadership practices.	No standardised effect sizes were reported (only unstandardised MD's).	Criteria for quality appraisal not provided Limited nr of quantitative studies	C
Blume, 2010	Meta- analysis including field and lab studies k = 89 studies, mainly USA and Canada	students/ managerial/ non- managerial/ mainly USA and Canada	<p>1. Learning transfer is related to trainee characteristics (e.g., cognitive ability, experience, personality, motivation), work environment factors (i.e., support, climate, constraints/opportunity), training interventions, learning outcomes (i.e., knowledge, self-efficacy), and trainee reactions.</p> <p>2. The predictor–transfer relationships after removing same-source and same- measurement-context (SS/ SMC) bias is weaker.</p> <p>3. Regarding transfer measurement, predictor–transfer relationship is stronger in the case of (a) no time lag between training and the transfer measure, (b) self-measures compared to non-self (i.e., peer or objective) measures, and (c) use as measurement of transfer compared to effectiveness.</p> <p>4. In general predictor–transfer relationships are stronger for open than for closed skills.</p> <p>5. Predictor–transfer relationships are stronger in the laboratory context for cognitive ability and posttraining self-efficacy; for pretraining self-efficacy, motivation, posttraining knowledge and goal-setting the relationship is stronger in the field context.</p>	<p>1) See table 1 for complete overview. Examples: Cognitive ability $p = .37$ Neuroticism $p = -.19$ Posttraining self-efficacy $p = .22$</p> <p>2) See table 1 Example: Posttraining self-efficacy SS/SMC $p = .46$ Not SS/SMS $p = .20$</p> <p>3) See table 2. a) Example posttraining self-efficacy Time lag $p = .11$ no time lag $p = .38$ b) Example Motivation Self $p = .33$ other/objective $p = .11$ c) Example Motivation Use $p = .36$ Effectiveness $p = .10$</p> <p>4) See table 3. 5) See table 4.</p>	Concrete and elaborate information on the variables is lacking. Large number of variables	A

			<p>6. For all predictors except posttraining knowledge the predictor–transfer relationships were stronger in a published than an unpublished study.</p> <p>7. The longer the length of time between training and the transfer measure the weaker the predictor–transfer relationships for a) posttraining knowledge and b) posttraining self-efficacy but not for c) pretraining self-efficacy d) motivation and e) work environment (ns).</p> <p>8. There is a moderate relationship between trainees' rating versus a) supervisor's and b) peers' assessment of transfer and c) a large relationship between measures at different times by the same source.</p>	<p>6) See table 5.</p> <p>7) See table 6. a) $\beta = -.25$ b) $\beta = -.64$ c) $\beta = -.08$ (ns) d) $\beta = -.29$ (ns) e) $\beta = .07$ (ns)</p> <p>8) See table 7 a) $\rho = .28$ b) $\rho = .26$ c) $\rho = .57$</p>		
Ciur, 2012	Quasi-experiment n = 30	managerial, automotive industry, Romania	Significant differences for emotional stability and social boldness.	not reported	Small sample	B
Collins, 2001	Systematic review k = 54	Managers, leaders and/or executives	<p>1. Thirty percent of the studies reported organisational performance improvement as outcome. Seventy percent measured learning.</p> <p>2. 81% of studies with performance-level outcomes measured system performance while 19% addressed financial-level performance as the outcome variable.</p> <p>3. Strategic leadership is most frequently researched leadership development content area (33%), then employee development (20%) and supportive environment (15%).</p> <p>4. Formal training is primary leadership development intervention (41%), then job assignments (32%).</p>	Not reported	<p>Quality of the included studies unclear</p> <p>Qualitative review</p>	C

Collins, 2004	Meta-analysis k = 83 Posttest only control group (POWC), pretest posttest with control group (PPWC), single group pretest posttest (SGPP), correlational	Managers, leaders, executives, officers, supervisors, and/or foremen Various sectors	<ol style="list-style-type: none"> Leadership development programs are associated with knowledge objective outcomes. Leadership development programs are associated with expertise objective outcomes. Leadership development programs are associated with expertise subjective outcomes. Leadership development programs are associated with system objective outcomes. 	<p>1) POWC .96 CI95 .82 1.12 SGPP 1.36 CI95 1.18 1.56</p> <p>2) POWC .54 CI95 .14 .95 PPWC .35 CI95 .20 .50 SGPP 1.01 CI95 .87 1.15</p> <p>3) POWC .41 CI95 .25 .58 PPWC .40 CI95 .20 .61 SGPP .38 CI95 .30 .46</p> <p>4) POWC .39 CI95 .19 .59</p>	Quality of the included studies unclear Wide CI's	B
Cummings, 2008	Systematic review k = 9?	Nurses in leadership positions	All nine studies found positive results for (mostly) self-rated leadership and observed leadership.	not reported	Design of included studies unclear	C
DeRue, 2012	Cohort study (9 months) n = 173	MBA students, various, USA	Small positive effect on leadership development.	not reported	-Same source bias -Low incremental validity	B
Duygulu, 2011	Before – after study, repeated measures n = 30	nurses, health sector, Turkey	Significant effect on leadership practices (?).	not reported	Effect unclear Small sample size	C
Eden, 2000	Meta analysis of RCT's k = 7	various	Small to moderate overall effect.	not reported	no serious limitations	A

Frich, 2015	systematic review of longitudinal and controlled studies k = 45	physicians	<ol style="list-style-type: none">1. The studies showed considerable heterogeneity concerning conceptual frameworks, teaching and learning methods, educational content, evaluation design, and outcomes measured.2. All 45 studies reported positive outcomes, but few studies reported system-level (4) effects, such as improved performance on quality indicators for disease management or increased customer satisfaction.3. Our findings suggest that the leadership programs described in the medical literature focus more on the 'know' and 'do' elements of leadership than the 'be' component, which some argue is fundamental in attaining the capacity to lead.4. It was found that the literature on physician leadership development has been centred on imparting conceptual knowledge to physicians as individuals, for which lectures and seminars may be suitable, and has directed fewer resources to efforts in building self-awareness, for which action-based learning, feedback, and self-development activities may be more appropriate.5. Importantly, the few studies that documented favourable organisational outcomes, such as improvement in quality indicators for disease management, were characterised by the use of multiple learning methods, including lectures, seminars, and group work, and involved action learning projects in multidisciplinary teams. The implication of this finding is that greater investment in programs using teamwork and multiple learning methods is likely to have the largest impact in the area of leadership development for physicians.	Not reported	Limited search Most studies included evaluated the effect on KP level 1 or 2 (only 6 out of 45 reported on level 3 or 4)	B
----------------	--	------------	--	--------------	--	---

Kiesewetter, 2013	systematic review, k = 8	healthcare, physicians	<ol style="list-style-type: none"> 1. The small number of studies included in the review shows that the systematic and evidence-based development of leadership skills does not currently play a prominent role in the training of physicians. 2. The range of training programs is very broad and leadership skill components are diverse. 3. The diverse concepts underlying leadership skills in medical training show little agreement as to what content should be emphasised in the development of leadership skills. 4. Reactions of participants to trainings were positive (KP level 1), yet no behavioural changes (KP level 3 and 4) through training were examined. 	not reported	Limited nr of studies included Design & quality of the studies unclear	C
Krejci, 1997	Before- after study n = 87	nurses, health care, USA	Significant difference in the perception of the understanding and the ability of the leadership competencies.	No effect sizes reported	Self report Halo effect	C
Lacerenza, 2017	Meta-analysis k = 335 Repeated measure design, independent groups design, independent groups design with repeated measures	various	<ol style="list-style-type: none"> 1. Leadership training programs have a positive effect. 2. Leadership training programs have a positive effect on trainee reactions. 3. Leadership training programs have a positive effect on affective-, cognitive-, and skill-based learning outcomes. 4. Leadership training programs lead to the transfer of trained affective-, cognitive-, and skill-based concepts. 5. Leadership training programs positively influence organisational and subordinate outcomes. <p>The strength of these effects differs based on various design, delivery, and implementation characteristics. Moderator analyses support the use of needs analysis, feedback, multiple delivery methods (especially practice), spaced training sessions, a location that is on-site, and face-to-face delivery that is not self-administered. Results also suggest that the content of training, attendance policy, and duration influence the effectiveness of the training program.</p>	<ol style="list-style-type: none"> 1) $\delta = .76$ CI95 .64 .89 2) $\delta = .63$ CI95 .12 1.15 3) $\delta = .73$ CI95 .62 .85 4) $\delta = .82$ CI95 .58 1.06 5) $\delta = .72$ CI95 .60 .84 6) See tables 1 - 7 	Many hypotheses and moderators tested Studies from 1951 included Quality of the included studies unclear	B

Leskiw, 2007	Systematic review k = ?	Best practice organisa- tions	Six key factors were found to be vital for effective leadership development: a thorough needs assessment, the selection of a suitable audience, the design of an appropriate infrastructure to support the initiative, the design and implementation of an entire learning system, an evaluation system, and corresponding actions to reward success and improve on deficiencies.	Not reported	Very limited search Number, design and quality of included studies unclear	D
Leslie, 2005	Before - after study n = 56	paediatricians, health care, USA	<ul style="list-style-type: none"> » Participants were confident in many of their leadership qualities but desired increased training, particularly in areas of time and priority management and leading 'from the middle'. » Participants positively evaluated the training program and improved in self- reported basic competencies; 87% also reported fully or partially achieving a leadership-related goal. 	not reported	Self-report Small sample	C
Lyons, 2018	systematic review, k = 11 including RCTs	medical students	<ol style="list-style-type: none"> 1. Leadership curricula evaluated were markedly heterogeneous in their duration and composition. 2. The lack of a widely-accepted definition of clinical leadership and what it entails further complicates training, assessment and comparison of approaches. 3. A wide range of leadership curricula have shown subjective effectiveness. There is limited objective evidence however, and few studies have measured effectiveness at the system and patient levels (KP 4). 4. Effective programs tended to utilise a combination of didactic learning, tutorials and reflective learning. 5. Because all durations of intervention showed positive results, long and complex courses may not be required to achieve positive change; short, punchy courses with clear objectives may well be as effective. 	not reported	no serious limitations	A
Malling, 2009	Controlled before- after study n = 56	consultants responsible for postgraduate medical education, health sector	<ol style="list-style-type: none"> 1. No differences in multi-source feedback scores at one year follow up compared to baseline measurements, either in the intervention or in the control group. 2. Leadership course following a MSF procedure compared to MSF alone does not improve leadership skills. 	0 / ns	Prone to selection bias High dropout in control group	B

Mesmer, 2010	Meta-analysis k = 159	Adults	<p>Meta-analysis on impact of pre-training interventions on learning.</p> <ol style="list-style-type: none"> 1. Trainees provided with attentional advice prior to training will have higher average learning scores on measures of cognitive (H1a) and skill-based (H1b) learning than trainees not provided with such advice. 2. Trainees provided with general attentional advice prior to training did NOT have higher average learning scores on measures of cognitive (H2a) but did on skill-based (H2b) learning than trainees provided with specific attentional advice. 3. Trainees provided with meta-cognitive strategies prior to training have higher average learning scores on measures of cognitive (H3a), skill-based (H3b), and partially on affective (H3c) learning than trainees not provided with meta-cognitive strategies. 4. Trainees provided with 'why-based' meta-cognitive strategies may perform better on cognitive measures of learning than trainees provided with "think aloud" strategies, but NOT on skill-based learning. 5. Trainees provided with advance organizers performed better on indicators of cognitive (H5a) and skill-based (H5b) learning than trainees not provided with advance organizers. Stronger results were found for graphic rather than textual organizers for skill-based learning outcomes. 6. Trainees provided with a pre-training goal orientation (whether mastery- or performance-oriented) performed better on indicators of cognitive (H6a), skill-based (H6b), and affective (H6c) learning than trainees not provided with a pre - training goal. 7. Trainees provided with a mastery goal orientation will perform better on indicators of cognitive (H7a), skill-based (H7b), and affective (H7c) learning than trainees provided with a performance goal orientation. 8. Trainees provided with preparatory information will perform better on indicators of cognitive (H8a), skill-based (H8b), and affective (H8c) learning than trainees not provided with preparatory information. 	<p>1) a δ = .67 b δ = .80</p> <p>2) a δ = .71 vs .66 b δ = .88 vs .65</p> <p>3) a δ = .61 b δ = .51 c δ = .40</p> <p>4) Cognitive average Think δ = .62 (ns) Why δ = .60</p> <p>Skills average Think δ = .57 Why δ = .40</p> <p>5) a δ = .54 b δ = .71</p> <p>6) a δ = .71 b δ = .71 c δ = .85</p> <p>7) a not tested b δ = .89 vs δ = .60 c δ = .47 (direct comparison)</p> <p>8) n not tested b δ = .48 c δ = .45</p> <p>See tables 2 - 7 for more detailed effect sizes</p>	Design and quality of included studies unclear	B
--------------	--------------------------	--------	---	--	--	---

Morrow, 1997	Meta analysis of controlled and uncontrolled studies k = 18	various managerial levels, pharmaceutical sector, USA	Large variation of effectiveness of the programs.	1. Range of d-.09 to 1.11 Managerial training d.31, utility (ROI) 45%, with less than d.64 and ROI 418% for sales/technical training	Rate bias Small sample sizes	B
Mianda, 2018	systematic review of pre-post and cross- sectional studies k = 24	frontline healthcare providers	<ol style="list-style-type: none"> 1. Clinical leadership development is an on-going process and must target both novice and veteran frontline health care providers. 2. The content of clinical leadership development interventions must encompass a holistic conceptualisation of clinical leadership. 3. Interventions for clinical leadership development should use work-based learning approaches, and experiential and practice-based learning, as these are reported as the most effective. 	not reported	merely descriptive & anecdotal review	B
Powell, 2010	Meta- analysis of various designs k = 85 n = 4.779	entry level and middle management	<ol style="list-style-type: none"> 1. The overall effect of managerial training interventions is small. 2. The effects did not improve over time (50 years). 3. When the outcome being sought, or at least tested for, is a learning outcome, the effect size magnitudes are consistently larger than the other outcome groups and these effects are more consistently significant. 	<ol style="list-style-type: none"> 1. $r = .25$ 2. Between .18 and .38 3. Between .17 and .55 	Limited information on the measures or characteristics of the interventions.	A/B
Solansky, 2010	Survey with posttest n = 303,	administrators & leadership mentors, education, USA	<ol style="list-style-type: none"> 1. Note: Differences between self-report and observer report. 2. Coaching time (a) and number of contacts (b) are both significantly, positively and moderately related to mentees willingness to share information with mentors and the mentee group regarding their leadership. 	<ol style="list-style-type: none"> 2a. $\beta .18$ 2b. $\beta .02$ 	Observers were chosen by the participants	C
Steinert, 2012	systematic review (mostly pre-post) k = 41	clinical faculty members (US and Canada)	<ol style="list-style-type: none"> 1. Participants value leadership development activities and report changes in attitudes, knowledge, skills and behaviour. 2. Despite methodological limitations, certain program characteristics seem to be associated with positive outcomes: 1) the use of multiple instructional methods; 2) experiential learning and reflective practice; 3) individual and group projects; 4) peer support; 5) mentorship; and 6) institutional support. 	not reported	no serious limitations	B

Storey, 2004	Review of cross-sectional studies k = ?	various sectors, SME's, various countries	Mixed evidence for the effect of formal training.	not reported	sample unclear	C
Taylor, 2005	Meta analysis of randomised and non-randomised studies k = 117	various sectors, employees and countries	<ol style="list-style-type: none"> Effects were largest for learning outcomes, smaller for job behaviour and results outcomes. Although effects on declarative knowledge decayed over time, training effects on skills and job behaviour remained stable or even increased. Skill development was greatest when learning points were used and presented as rule codes and when training time was longest. Transfer was greatest when a) mixed (negative and positive) models were presented; b) practice included trainee-generated scenarios; c) trainees were instructed to set goals; d) when trainees' superiors were also trained; e) when rewards and sanctions were instituted in trainees' work environments. 	<p>1. declarative knowledge $d = 1.20$;</p> <p>procedural knowledge $d = 1.18$;</p> <p>attitudes $d = .33$;</p> <p>job behaviour $d = .27$;</p> <p>workgroup productivity $d = .13$;</p> <p>workgroup climate $d = .11$</p>	Contains 0 in some of the confidence intervals	A
Taylor, 2009	Meta-analysis of randomised and non-randomised studies k = 107	Various samples, managers, various countries	<ol style="list-style-type: none"> Managerial training is positively related to transfer of learning for all rating sources. Population effect size estimates for studies without control groups were larger than studies with control groups when based on a) superior and b) peer ratings, but smaller when based on c) self- and d) subordinate ratings. For self-ratings sources, superior ratings and subordinate ratings the standardised betas were relatively small, positive, and nonsignificant with regard to time lag between training and posttest. Transfer effect sizes for interpersonal skills management training programs, followed a similar pattern across the four rating sources as when all studies were compared. The topic with the greatest training transfer perceived by subordinates was for a) general management skills training compared to b) interpersonal skills and c) goal setting or performance appraisal skills. 	<p>1) Self $\delta = .64$ Superior $\delta = .53$ Peer $\delta = .26$ Subordinate $\delta = .13$</p> <p>2) a $\delta = .55$ vs .45 b $\delta = .35$ vs .21 c $\delta = .60$ vs .90 d $\delta = .10$ vs .18</p> <p>4) Table 5</p> <p>5) a $\delta = .50$ b $\delta = .11$ c $\delta = .34$</p> <p>More effect sizes in tables 2 - 6</p>	Sometimes conclusions are drawn on a rather small amount of studies (and small total n)	B

Van der Loch, 2013	Cross-sectional N = 595	Managers	<ol style="list-style-type: none"> 1. The use of identical elements* is positively related to training transfer. 2. The relationship between the use of identical elements and training transfer is partially mediated by motivation to transfer. 3. Motivation to learn is positively related to training transfer. 4. The relationship between motivation to learn and training transfer is mediated by motivation to transfer. 5. Expected utility is positively related to training transfer. 6. The relationship between expected utility and training transfer is partially mediated by motivation to transfer. 7. Identical elements will predict training transfer over and above trainees' motivation to learn and expected utility. <p>* Identical elements refers to the extent to which the stimuli and responses in the training setting are identical to those in the actual performance environment.</p>	<p>1) $r = 0.53$</p> <p>3) $r = 0.42$</p> <p>5) $r = 0.57$</p>	No serious limitations	D
Webb, 2014	systematic review k = 24	medical students	Overall, evaluation of effectiveness and quality of evidence showed that most curricula did not demonstrate changes in student behaviour or quantifiable results.	not reported	design and quality of the included studies unclear	C

Excluded studies

Author & year	Reason for exclusion
Amagoh, 2009	Narrative literature review / expert opinion.
Avolio, 2010	Study calculates Return on leadership development investment (RODI) for training at different organisational levels for on-site, off-site and on-line training. Effect sizes are taken from an earlier meta-analysis by Avolio (2009).
Burke, 1986	Publication date out of scope (1986), often cited study though. Results of 70 managerial training studies to empirically integrate the findings of the studies. The meta- analysis results for 34 distributions of managerial training effects representing six training content areas, seven training methods, and four types of criteria (subjective learning, objective learning, subjective behaviour, and objective results) indicated that managerial training is, on the average, moderately effective.
Day, 2000	Narrative literature review / expert opinion.
Jeyaraman, 2018	Scoping review, descriptive.
Kelloway, 2010	Narrative literature review / expert opinion.
Lopes, 2013	Review into characteristics of the use of business games in leadership development. in SCOPUS, ISI, and BKL (Bernie Keys Library) databases. The review identified five games that met the criteria and objectives of this research. This study shows that using business games for leadership development is still a hard task. It points up some problems and difficulties in this task and suggests ways to develop more effective methods for leadership development with business games.
Markuns, 2010	Expert opinion.
Morahan, 1998	Expert opinion.
Niemiec, 1992	Publication date out of scope (1992). The meta-analysis synthesized the results of 22 studies of management education in institutional settings. The results indicate that the treatment's median effect size is .7 -a fairly substantial effect. Several differential effects were noticed, including the experience level of the managers and the types of instrumentation and institutional setting.
Pearson, 2007	Does not concern leadership training.
Salas, 2012	Unsystematic literature review on effectiveness of training. Does provide an overview of previous meta-analysis and checklists for different phases of the training process.
Straus, 2013	Not relevant given the REA question: findings suggest that (some) leadership training programs affected participants advancement in academic rank and hospital leadership position and that participants were more successful in publishing papers. Other outcomes were based on too limited nr of studies.
Stoller, 2013	Expert opinion.
Yeung, 2012	Outcome measure (association between team leadership skills and cardiac arrest simulation test score, pre-shock pause, and hands off ratio) not relevant.



Australian
National
University

A partnership between the ACT Government through the ACT public health system and the ANU Research School of Management.

Any enquiries in relation to the content of this REA should be directed to CEBMa through their website: www.cebma.org



ACT Health



Canberra Health
Services