ARTICLE

Check for updates

Evaluations of training programs to improve capacity in K*: a systematic scoping review of methods applied and outcomes assessed

Samantha Shewchuk [1✉], James Wallace[1] & Mia Seibold[1]

This paper examines how frequently K* training programs have been evaluated, synthesizes information on the methods and outcome indicators used, and identifies potential future approaches for evaluation. We conducted a systematic scoping review of publications evaluating K* training programs, including formal and informal training programs targeted toward knowledge brokers, researchers, policymakers, practitioners, and community members. Using broad inclusion criteria, eight electronic databases and Google Scholar were systematically searched using Boolean queries. After independent screening, scientometric and content analysis was conducted to map the literature and provide in-depth insights related to the methodological characteristics, outcomes assessed, and future evaluation approaches proposed by the authors of the included studies. The Kirkpatrick four-level training evaluation model was used to categorize training outcomes. Of the 824 unique resources identified, 47 were eligible for inclusion in the analysis. The number of published articles increased after 2014, with most conducted in the United States and Canada. Many training evaluations were designed to capture process and outcome variables. We found that surveys and interviews of trainees were the most used data collection techniques. Downstream organizational impacts that occurred because of the training were evaluated less frequently. Authors of the included studies cited limitations such as the use of simple evaluative designs, small cohorts/sample sizes, lack of long-term follow-up, and an absence of curriculum evaluation activities. This study found that many evaluations of K* training programs were weak, even though the number of training programs (and the evaluations thereof) have increased steadily since 2014. We found a limited number of studies on K* training outside of the field of health and few studies that assessed the long-term impacts of training. More evidence from well-designed K* training evaluations are needed and we encourage future evaluators and program staff to carefully consider their evaluation design and outcomes to pursue.

[1] University of Delaware, Newark, USA. ✉email: shewchuk@udel.edu

Introduction

The generation and utilization of research knowledge plays a vital role in addressing inequities within the education system (Farley-Ripple et al., 2018; Honig & Coburn, 2008; Denaro et al., 2022), however it often does not play this role for many reasons (Malin et al., 2020). This realization has resulted in an expanding "knowledge field" which seeks to better understand how research evidence could have a greater bearing on policy and practice decisions (Lockton et al., 2022; Rycroft-Smith, 2022). Researchers in the "knowledge field" have used many terms to describe the set of functions and processes in which research evidence is produced, shared, and used by members of the research, practice, and policy communities. For this project, we draw from Shaxson et al. (2012) concept paper, which introduced the term K* to describe the "set of functions and processes at the interfaces between knowledge, practice, and policy" (p. 2). In other words, K* is focused on connecting researchers and their work to organizations and communities outside of academia so that research is useful, useable, and utilized. K* is part of a broader semantic cluster that includes the ideas of 'knowledge brokering,' 'boundary spanning,' 'knowledge mobilization,' 'knowledge translation,' 'knowledge exchange,' 'knowledge extension,' 'engaged scholarship,' and 'dissemination and implementation.' Definitions of these terms are provided in Supplemental File A. Shaxson and colleagues described K* as a 'catch-all' term, intended to represent the wide range of concepts within the larger cluster of terms.

Work is being done in multiple areas to create an environment conducive to K*, including funding research on the topic, supporting interactions between researchers and research users, developing policies that mandate open-access publishing of research findings, developing networks to serve the research needs of practice-based organizations, and building the capacity of individuals to promote and enable K*. Although it is beyond the scope of this paper to provide a detailed summation of the approaches being taken to improve the production, sharing, and use of research knowledge, we encourage readers to read the reviews by Fahim et al. (2023) and Walter et al. (2005) which provide a more thorough description of the topic.

Findings from Fahim et al. (2023) and Walter et al. (2005) suggest multi-pronged approaches are needed to promote research use. One important strategy is the development of individuals' knowledge, skills, and confidence to promote and enable K* (Holmes et al., 2014; Mishra et al., 2011; Tabak et al., 2017; Halsall et al., 2022), as capacity building may play a crucial role in predisposing change more broadly (Davis & Davis, 2009; Golhasany & Harvey, 2023). As such, more funding has been called for (Cooper et al., 2018; Phipps et al., 2016; Georgalakis & Rose, 2021) and invested to strengthen individual capacity in this area (e.g., Holmes et al., 2012; Garritzmann et al., 2023). CRE-ATEd—Collaboration, Research Equity, and Action Together—is a program designed to prepare individuals to promote strong, equitable relationships among the research and practice communities. In part, our work consists of offering a year-long fellowship centered around developing individual capacity to facilitate the exchange of knowledge among researchers and research users to support evidence-informed and equity-centered policy and practice. The fellowship consists of online modules, live workshops, and opportunities for fellows to apply their learning.

Along with investment in K* arises a need for evaluations to provide information for funders and program managers to determine if a program should continue, improve, end, or scale up, thereby ensuring the efficient and effective allocation of resources (Rycroft-Smith, 2022; Hartling et al., 2021). As such, CREATEd has embedded evaluation into all our activities with the goal of capturing lessons to continuously improve our work and to document and measure progress toward achieving our goals. Using evaluation methods (Matthews & Simpson, 2020) and metrics (Barwick et al., 2020) can indicate whether training goals were achieved. However, there is no consensus regarding which outcome indicators or evaluation methods to use. To help inform our evaluation, as well as contribute to the literature, we conducted a scoping review to synthesize information on the method and outcome indicators used to evaluate related training programs and identify areas for improvement in current training evaluation approaches.

Method

This review was based on Arksey and O'Malley's (2005) systematic scoping review methodological framework, which consists of five stages: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarizing, and reporting the results.

Identifying the research question. The population of interest for this review was K* professionals who work to connect people and ideas across research, practice, and policy communities. We included practitioners, policymakers, researchers (and current graduate students), community members, and knowledge brokers in our definition of K* professionals. The intervention of interest was any training or capacity building activity related to K*. As the purpose of this study is to identify the methods and outcomes used for training evaluations, study design and outcomes of the included studies were left intentionally broad.

Identifying relevant studies. We conducted a search on articles published before August 2022, in eight multidisciplinary electronic databases: ProQuest, ScienceDirect, JSTOR, EBSCO, PubMed, Web of Science, Academic OneFile, and Scopus. In addition, we searched for relevant gray literature in Google Scholar. Finally, we reviewed reference lists of included studies and relevant reviews to identify additional articles. The search terms are those defined in Supplemental File B.

Study selection. All citations were imported into Excel and duplicate citations were removed manually. A two-stage screening process for eligibility was conducted. Articles were eligible for inclusion if they met each component of the inclusion criteria and did not have any criteria for exclusion (Table 1). Studies were not excluded based on year of publication, country of publication, type of publication, field of publication, or quality of publication.

We retrieved a total of 1297 citations from the systematic searches of the eight multidisciplinary databases, the Google Scholar search engine, and review of reference lists. After removal of the duplicates, 824 articles were independently scanned by three researchers based on their title and abstracts. Results from the researchers were compared and discrepancies were discussed to arrive at agreement. In instances where the researchers had varied opinions of the inclusion of a specific article for the next stage, the majority decision prevailed. After this first round of screening, a total of 127 resources were processed for full-text review. Disagreements among reviewers in the full-text screening phase were reconciled by discussion and consensus. Resources that could not be obtained for full-text review through online databases, library searches, or through direct contact with the study's first author were excluded from the final analysis. In addition, as described in Table 1 above, literature reviews were excluded from inclusion in the scoping review, however the reference lists of reviews were scanned to identify eligible studies.

Table 1 Inclusion and exclusion criteria for scoping review.

| Criteria	Description
Inclusion	Study describes evaluations of K* training programs
	Study contains descriptions of the training program, methods used to evaluate the program, and outcomes assessed in the evaluation.
	Only English studies are included
	Country of publication, year of publication, type of publication, field of publication, or quality of publication are not restricted
Exclusion	Literature reviews with no primary data collection
	Study describes training program, without discussion of its evaluation
	Document is not accessible through online databases, library searches, or through direct contact with the study's first author

Fig. 1 PRISMA diagram. The PRISMA diagram depicts the flow of information through the distinct phases of the scoping review. It maps out the number of records identified, included and excluded, and the reasons for exclusions.

Furthermore, the reference lists of all eligible studies were reviewed to identify additional sources for inclusion. In total, 47 documents met our criteria for inclusion in the scoping review (Fig. 1). For a full list of the included publications, see Supplemental File C.

Extracting and analyzing data. We employed scientometric and content analysis to examine the included resources. The scientometric analysis provides a comprehensive overview of the included studies by visualizing the relationships among articles, journals, keywords, citations, and co-citation networks (Chen & Song, 2019). VOSviewer version [1.6.19] was used to conduct the scientometric analysis (Van Eck & Waltman (2010)). Different parameters, including publication evolution over time, citation analysis for core publications, co-authorship analysis, bibliographic coupling analysis, and finally, co-occurrence analysis were used to map the K* evaluation literature included for this

review (Chen & Song, 2019). To view the settings for each of the analyses run in VOSviewer, see Supplemental File D.

We also conducted content analysis to provide more in-depth insights related to the methodological characteristics, outcomes assessed, and future evaluation approaches proposed by the authors of the included studies. We extracted relevant information from resources included in the final analysis using a standardized data extraction template. Supplemental File E summarizes data extracted and definitions used for categorizing data. Data extraction was performed by two researchers, whereas a third researcher checked the workflow for completeness and accuracy. Disagreements were resolved by consensus. Methodological information extracted included the type of evaluation conducted, methodological approach used, type of design (one-versus two-group), sample size, data collection techniques, and timeline of data collection. We categorized training outcomes based on the Kirkpatrick model. The Kirkpatrick model was first developed in the 1950's to evaluate the effectiveness of training
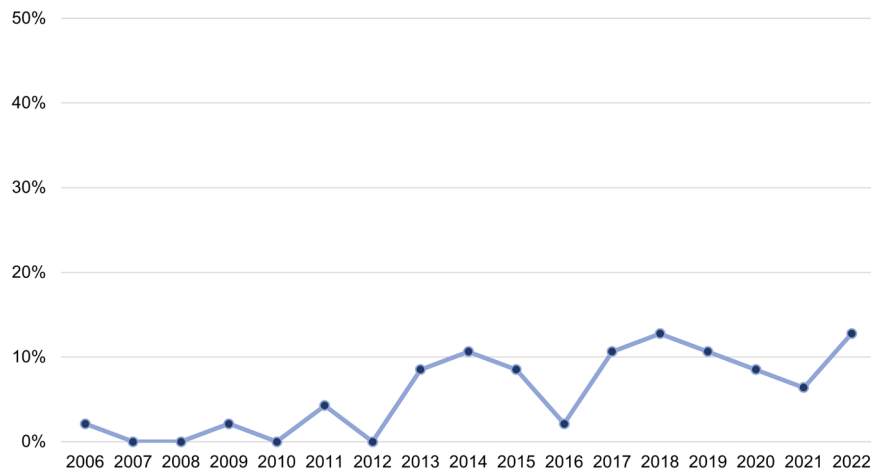
Fig. 2 Publication evolution: percentage of articles over time. Figure 2 depicts the percentage of articles published over time. The *y*-axis has a percentage scale (0–50%), while the *x*-axis shows the years in which articles were published (2006–2022).

and educational programs and is still the most applied model to date (Alsalamah & Callinan, 2021b). It can be applied to any style of training, both formal and informal, to determine the efficacy of a training program based on four levels. The levels, in order, are reaction, learning, behaviors, and results, respectively. Each successive level of the model represents a more precise measure of the effectiveness of a training program. In the first level, the *reactions* of trainees are explored. This includes measuring the extent to which learners found the training to be relevant, engaging, useful, and enjoyable. In level two, the *learning* of participants is examined to understand whether learners acquired the intended knowledge, skills, attitude, confidence, and commitment because of the training. Level three focuses on the *behavior* of trainees after completion of the training. In this level, evaluators focus on measuring whether learners change their behaviors because of the training. Level four of the model looks at the downstream *results* or impacts that occur because of the training (Kirkpatrick & Kirkpatrick, 2006). Finally, future evaluation approaches proposed by the authors of the included studies were extracted and inductively analyzed to identify themes and characteristics.

Collating, summarizing, and reporting the results. It is important to note that "scoping reviews do not aim to produce a critically appraised and synthesized answer to a particular question, rather they aim to provide an overview or map of the evidence" (Munn et al., 2018). Therefore, the results of the included sources are described in the context of the overall aim of the review. Also, the aggregated findings provide an overview of the research rather than an assessment of the quality of individual studies.

Limitations. To conduct a broad search of the published literature, we included eight multi-disciplinary databases, Google scholar, and scanned the reference lists of included articles. We recognize that we may have missed some K* training evaluations if the studies were not published or accessible online. In addition, as Shaxson et al. (2012) and other researchers have noted, there are many different terms used to refer to the processes and functions of connecting research and practice; this study may have omitted some terms from the search criteria and may therefore have excluded relevant studies. However, other reviews of K* (and related terms) have used similar search criteria as we have employed (e.g., Golhasany & Harvey, 2023; Mallidou et al., 2018; Murunga et al., 2020; Tait & Williamson, 2019). Therefore,

while this review may not be exhaustive, it provides a comprehensive overview of the literature on K* training evaluations. As described above, 47 documents were included in the review. Due to the small sample size, caution must be exercised when drawing generalizations and inferences from the data. A further limitation in our study is that we did not analyze information about the structure of the training programs themselves or how the evaluations were used to improve the programs under study. We note that other scoping reviews have been conducted to investigate these aspects (e.g., Golhasany & Harvey, 2023; Mallidou et al., 2018). While such studies are useful to identify strategies to increase the evidence base in this area, the field currently lacks consensus on what outcome indicators or evaluation methods to use. Our review addresses this issue and provides a roadmap for methodological improvement of K* training evaluations.

Results
Scientometric mapping of included studies. In this section, we report on the findings from the scientometric analysis. Different parameters, including publication evolution over time, citation analysis for core publications, co-authorship analysis, bibliographic coupling of documents, and keyword co-occurrence analysis are presented to map the bibliographic information from the included studies.

Publication evolution. Figures 2 and 3 illustrate the publication trends of the included studies. Most articles (i.e., 43 of 47) were published after 2012, accounting for 91% of the data sample. The top five journals contain 25 out of 47 items, representing 53% of the included publications. Implementation Science, a journal devoted to publishing articles on the implementation of evidence-based practices and programs in healthcare, has published the most articles focused on evaluating K* training programs, constituting approximately 36% of the publications. Worldviews on Evidence-Based Nursing, the Journal of Continuing Education in the Health Professions, The Pan African Medical Journal, and the International Development Research Centre round out the top five journals, each with two published articles.

Citation analysis for core publications. To identify the most influential publications of K* training program evaluations, we examined the total citation counts (as of August 2023) for each article. The top five highly cited articles within our dataset are shown in Table 2. The most cited paper is Meissner and colleagues' article The US Training Institute for Dissemination and

Fig. 3 Publication evolution: leading journals. Figure 3 depicts the percentage of articles that were published in different journals. The y-axis shows the journals in which articles were published, while the x-axis has a percentage scale (0–50%).

Implementation Research in Health, with 147 citations. As shown in the table, the top five cited articles were all published within Implementation Science.

Co-authorship analysis: authors, institutions, and countries. Co-authorship analysis is used as a proxy for collaboration (Newman, 2004). The most collaborative countries, organizations, and authors on K* training evaluations are illustrated in Figs. 4–6. In Figs. 4–6, the larger each node (circle) is the higher the number of documents the corresponding country, institution, and author have. In addition, the thicker the link between the nodes, the more collaboration has occurred between them. The number of documents of an author, organization, and country was set to two (see Supplemental File D for more information on VOSviewer settings). Of the 241 authors in the sample, 30 met the thresholds set in VOSviewer. The author's co-authorship map has 30 nodes, five clusters, 117 links, and a total link strength of 220. Ross Brownson from Washington University collaborated on the highest number of publications ($n = 7$), followed by Enola Proctor from Washington University ($n = 5$), Karen Emmons from Harvard University ($n = 4$), and Sharon Straus from the University of Toronto ($n = 4$). Figure 4 shows the time-based overlay visualization of collaborative relationships amongst authors based on the number of author publications and average-publication-year. Figure 4a (left) shows the visualization for the entire collaboration network ($n = 30$), while the Fig. 4b (right) figure shows a 'zoomed-in' view of the largest set ($n = 18$) of connected items.

Of the 100 organizations in the sample, 15 met the thresholds set in VOSviewer. The institution co-authorship map has seven nodes, six clusters, 21 links, and a total link strength of 33. Washington University collaborated on the highest number of publications ($n = 8$), followed by the National Cancer Institute ($n = 4$), McMaster University ($n = 4$), and St. Michaels Hospital ($n = 4$). Figure 5 shows the time-based overlay visualization of collaborative relationships amongst organizations based on the number of publications by organizations and average-publication-year. Figure 5a (left) shows the visualization for the entire collaboration network ($n = 15$), while Fig. 5b (right) shows a 'zoomed-in' view of the largest set ($n = 12$) of connected items.

Of the 27 countries in the sample, seven met the thresholds set in VOSviewer. The country co-authorship map has seven nodes, five clusters, three links, and a total link strength of five. The United States collaborated on the highest number of publications ($n = 18$), followed by Canada (13). Figure 6 shows the time-based overlay visualization of collaborative relationships amongst countries based on the number of publications by country and average-publication-year. Figure 6a (left) shows the visualization for the entire collaboration network ($n = 7$), while Fig. 6b (right) shows a 'zoomed-in' view of the largest set ($n = 4$) of connected items.

Bibliographic coupling of documents. For a better understanding of the extent to which the 47 documents in our sample shared references in common (Van Eck & Waltman, 2014), we constructed a bibliographic coupling network. The bibliographic coupling map has 47 nodes, 17 clusters, 247 links, and a total link strength of 569. The three studies with highest link strength (i.e., articles with the highest number of references in common with other articles) are Moore et al. (2018), Padek et al. (2018), and

Table 2 The top five highly cited articles included in the review.

Rank	Article title	Author(s)	TC[a]	Source
1	The US Training Institute for Dissemination and Implementation Research in Health	(Meissner et al., 2013)	147	Implementation Science
2	Core competencies in the science and practice of knowledge translation: Description of a Canadian strategic training initiative	(Straus et al., 2011)	124	Implementation Science
3	Building capacity for dissemination and implementation research: One university's experience	(Brownson et al., 2017)	56	Implementation Science
4	Developing the next generation of dissemination and implementation researchers: Insights from initial trainees	(Stamatakis et al., 2013)	52	Implementation Science
5	Training scholars in dissemination and implementation research for cancer prevention and control: A mentored approach	(Padek et al., 2018)	51	Implementation Science

[a]TC total citations.

Brownsen et al. (2017), with a total link strength of 99, 90, and 89, respectively. While the three studies with the highest number of citations were Meissner et al. (2013), Straus et al. (2011), and Stamatakis et al. (2013), with total link strengths of 34, 20, and 24. The time-based overlay visualization of the bibliographic coupling analysis is presented in Fig. 7. The visualization reveals that a major cluster of coupling strength exists, which is predominantly composed of articles published between 2009 and 2015. Figure 7a shows the visualization for the entire bibliographic coupling network ($n = 47$) while Fig. 7b shows a 'zoomed-in' view of the largest set ($n = 35$) of connected items.

Keyword co-occurrence analysis. Keyword co-occurrence analysis provides a description of the principal areas of focus in the research field (Chen & Song, 2019). In total, 227 keywords were used to describe the documents included in the review. Keywords that were used more than once were selected to map the network. The co-occurrence network of keywords is mapped in Fig. 8. In total, 56 keywords were mapped. The top ten keywords with the highest total link strength were knowledge translation, implementation, dissemination, capacity building, science, training, evaluation, education, mixed methods, and implementation science.

Findings from qualitative content analysis. In this section, we report the methodological characteristics and outcomes assessed in the included evaluation studies. In addition, we summarize the future evaluation approaches proposed by the authors of the included studies.

Methodological characteristics of evaluations. We were interested in capturing the methodological details of the evaluation studies included in the review (Table 3). Most evaluations were designed to examine process and outcome variables. Authors also used other terminology to describe their evaluations, including comparative evaluation, narrative evaluation, pluralistic evaluation, participatory evaluation, collaborative autoethnography, and environmental scan. When considering the benefits of the training programs, costs of training were not assessed. Evaluations were most likely to include one-group designs over two-group designs. Most of the included studies either did not report their sample size, or their sample size was small, which lowered the study's statistical power to detect true treatment effects. A wide range of data collection techniques were used, including surveys, interviews, document analysis, focus groups, participant feedback, and critical reflection. Researchers used several types of surveys to obtain data from trainees, including reaction surveys, pre- and post-competency surveys, and network surveys. Data collection most often occurred before the intervention and within 1-month immediately after the intervention. Long-term data collection (+1-month post-training) was less likely to occur across the included studies.

Outcomes assessed in the included studies. Authors were most likely to report measuring trainees' learning ($n = 38$, 80.9%) and reactions ($n = 37$, 78.7%). Many studies also measured the behavior change ($n = 32$, 68.1%) of trainees after they finished the training and returned to their jobs. Less common were studies that examined the downstream results ($n = 20$, 42.6%) that occurred because of the training. In the following sub-sections, we summarize the types of data collection techniques and outcome indicators used in relation to each level of the Kirkpatrick Framework.

Reaction: While evaluators typically used self-report surveys (utilizing both Likert and open-ended questions) to collect data on trainees' reactions to the training, qualitative approaches such
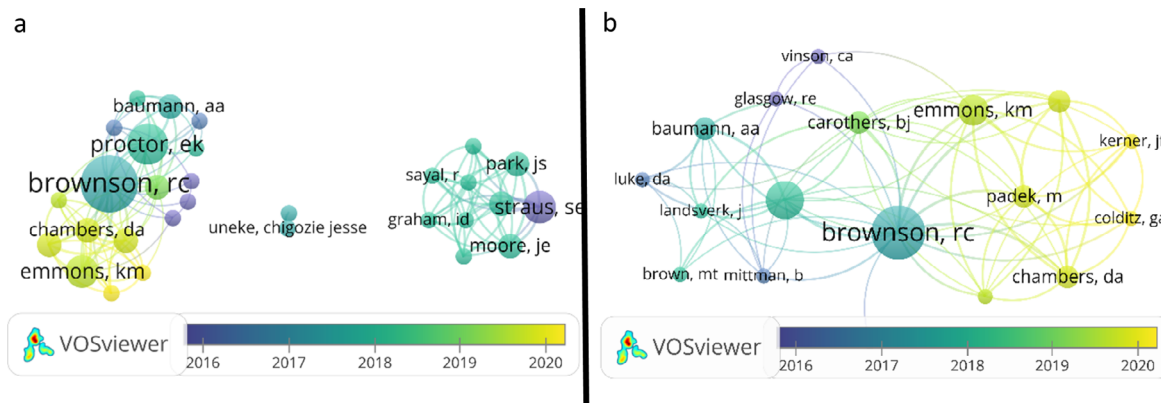
Fig. 4 Collaboration network between authors. Figure 4 shows the time-based overlay visualization of collaborative relationships amongst authors based on the number of author publications and average-publication-year. **a** (left) shows the visualization for the entire collaboration network, while **b** (right) zooms in to show the largest set of connected items.
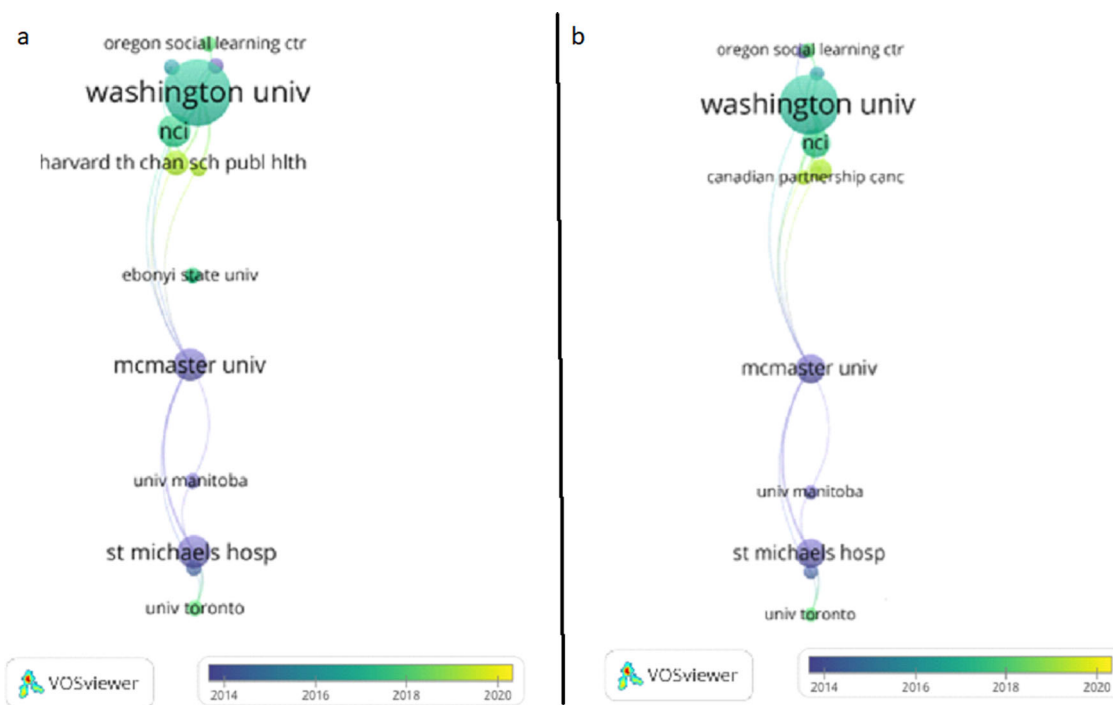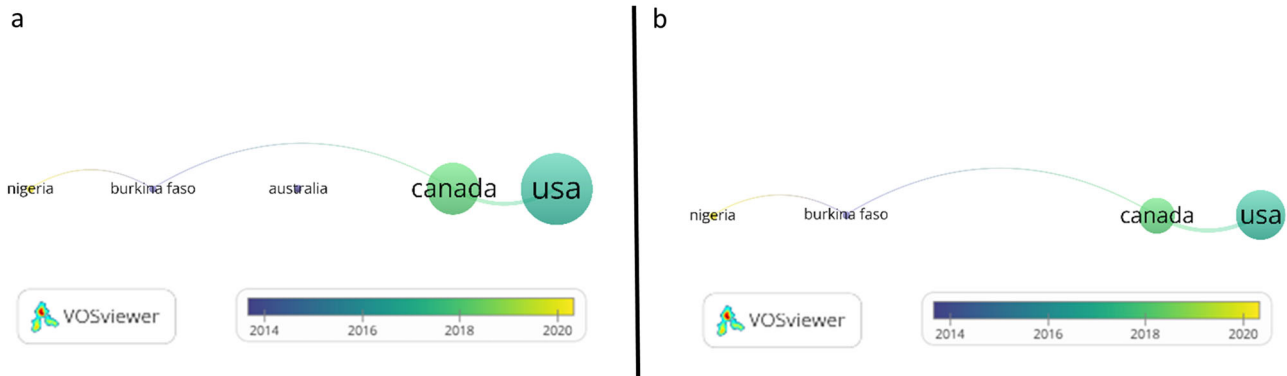

Fig. 5 Collaboration network between organizations. Figure 5 shows the time-based overlay visualization of collaborative relationships between organizations based on the number of publications per organization and average-publication-year. **a** (left) shows the visualization for the entire collaboration network, while **b** (right) zooms in to show the largest set of connected items.

as interviews and focus groups were also utilized. Authors used several approaches for reporting this data, including use of descriptive statistics such as mean, mode, and range to analyze survey responses, while qualitative data was grouped and analyzed thematically. To gauge trainees' reactions to the training, evaluators examined the level of satisfaction with the training program and its specific components (Oronje et al., 2022; Salloum et al., 2022), the extent to which the curriculum is clear and well organized (Astle et al., 2020; Lo Hog Tian et al., 2022; Morrato et al., 2015), the format of the training program (Froese & Montgomery, 2014; Gaid et al., 2022; Greenhalgh & Russell, 2006; Hess et al., (2013)), the competence of trainers (Brownson et al., 2021; Cunningham-Erves et al., 2021; Dagenais et al., 2015), the value of cohort-based learning (Brownson et al., 2017), trainees' level of engagement (Brownson et al., 2017), and the usefulness and relevance of the training to the trainees' actual job performance (Jones et al., 2015; Meissner et al., 2013; Olejniczak, 2017;

Provvidenza et al., 2020; Vinson et al., 2019). In addition, trainees were also asked to describe the key strengths of the program and provide suggestions for improvement (Moore et al., 2018; Rakhra et al., 2022).

Learning: When examining 'learning,' evaluators explored trainees' changes in knowledge or skills and/or changes in confidence and commitment to perform new K* knowledge and skills. Data collection approaches included the use of interviews, focus groups, observation, student data, and self-report surveys. Survey data were analyzed using mean ratings and tests of significance (e.g., Mbuagbaw et al., 2014). Depending on the evaluation, items were analyzed individually or were grouped into subscales to assess the underlying constructs (e.g., Proctor et al., 2019). While pre- and post-competency surveys were typically used to measure the change in participants' knowledge and skills, some evaluations only used post-competency assessments. Post-competency

a

b

Fig. 6 Collaboration network between countries. Figure 6 shows the time-based overlay visualization of collaborative relationships between countries based on the number of publications per country and average-publication-year. **a** (left) shows the visualization for the entire collaboration network, while **b** (right) zooms in to show the largest set of connected items.

a

b
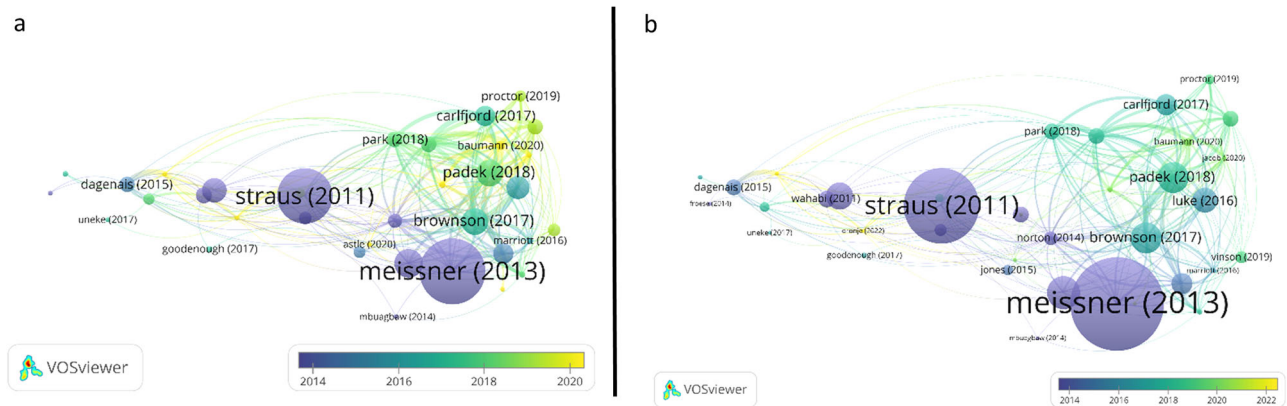
Fig. 7 Bibliographic coupling network. Figure 7 shows the time-based overlay visualization of the bibliographic coupling network based on the number of documents with shared references in common and average-publication-year. **a** (left) shows the visualization for the entire bibliographic coupling network, while **b** (right) zooms in to show the largest set of connected items.
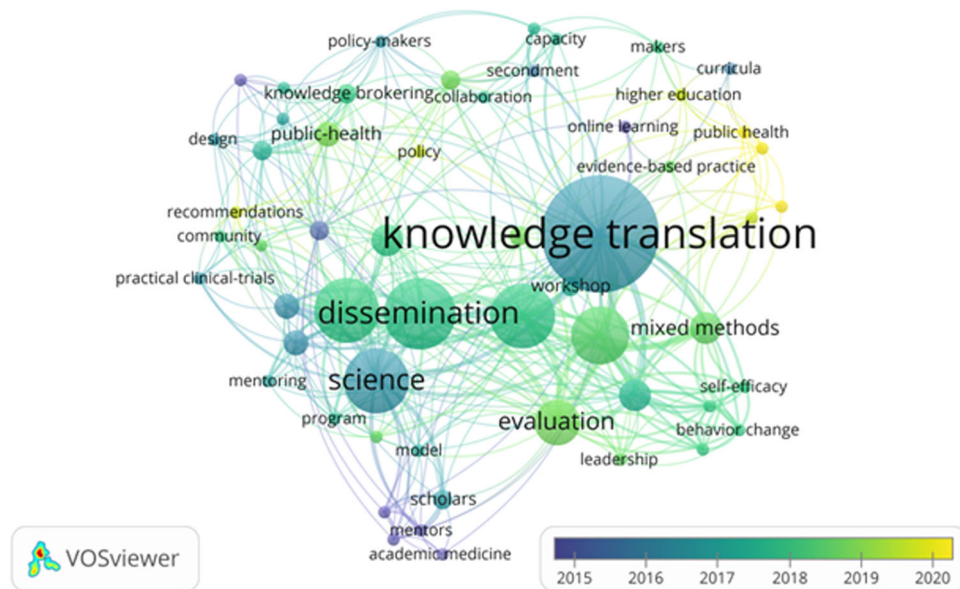
Fig. 8 Keyword co-occurrence network. Figure 8 shows the time-based overlay visualization of the keyword co-occurrence network based on the keywords used and average-publication-year.

Table 3 Methodological characteristics of studies.

Methodological details	Frequency (n = 47)	Percentage
Type of evaluation[a]		
Formative	2	4.3
Process	32	68.1
Outcome	29	61.7
Impact	6	12.8
Cost-benefit	0	0.0
Other term used	7	14.9
Method		
Mixed methods	42	89.4
Qualitative	4	8.5
Quantitative	1	2.1
One group or two group design		
One group	43	91.5
Two group	3	6.4
N/A	1	2.1
Sample size[b]		
Sample size not available	9	19.1
29 or fewer participants (large effect size, $\rho = 0.5$ or greater)	17	36.2
Between 30 and 84 (medium effect size, $\rho = 0.3$–0.5)	14	29.8
Between 85 and 781 participants (small effect size, $\rho = 0.3$–0.1)	7	14.9
782 participants or greater (very small effect size, $\rho = 0.1$ or less)	0	0.0
Data collection techniques		
Survey[c]	36	76.6
Reaction	31	86.1
Pre- and post-competency	19	52.8
Post-competency only	11	30.6
Network	3	8.3
Interviews	14	29.8
Document analysis	7	14.9
Participant feedback	5	10.6
Critical reflection	5	10.6
Focus group	5	10.6
Student data	4	8.5
Bibliometrics	2	4.3
Observation	2	4.3
Program data	1	2.1
Timeline of data collection[d]		
Before intervention	13	27.7
During intervention	8	17.0
Immediately following intervention (within 1 month)	20	42.6
1–5 months post intervention	2	4.3
6–11 months post intervention	10	21.3
After 12 months post intervention	10	21.3
Unclear	17	36.2

[a]Authors could report on one or more types of evaluation. As such, percentages add up to more than 100%.
[b]Sample size calculated for small ($\rho = 0.1$), medium ($\rho = 0.3$), and large ($\rho = 0.5$) effect sizes (with $\alpha = 0.05$, and power = 0.80) using a two-tailed bivariate normal model in G*Power.
[c]Percentages for the different types of survey are based on denominator ($n = 36$).
[d]The timeline of the data collection had to be clearly described (e.g., pre-competency survey was administered a week before the intervention and post-competency survey was administered a week after the intervention). In cases where the timeline is not clear (e.g., a pre-competency and post-competency survey was completed), resources have been marked as 'unclear'.

assessments typically occurred immediately following the completion of the training program. However, to measure trainees' sustained competence, some evaluations re-surveyed trainees at a later point in time. For example, Park et al. (2018) conducted interviews, focus groups, and surveyed individuals at baseline (pre-training), during training, and 6- and 12-months post-training to capture participants sustained knowledge in K* and self-efficacy in performing new K* skills.

Behavior: Several different techniques were used to collect data on trainees' changes in behavior. For example, Santacroce et al. (2017) used student data; Marriott et al. (2015), Morrato et al. (2015), Ndalameta-Theo et al. (2021), Vinson et al. (2019), and Meissner et al. (2013) used self-report surveys, and Hilbig et al. (2013) used interviews to gather data on participants' activities after taking part in their training program. Evaluators examined the extent to which trainees accessed resources, engaged in K* focused activities, and/or influenced the thinking of colleagues. Some (Brownson et al., 2021; Luke et al., 2016; Morrato et al., 2015) evaluators were also interested in examining the development of collaborations and partnerships between trainees. In these instances, evaluators used social network surveys to collect data on different types of relationships.

Results. As we previously noted, evaluators were less likely to report on downstream outcomes and impacts that occurred because of the training. Of those who did, many continued to rely on self-report surveys to capture result data (e.g., Carlfjord et al., 2017). However, other approaches were also utilized. For example, Baumann et al. (2020) used bibliometric analysis to understand the extent to which trainees had increased publications and grant funding compared to a control group. Kho et al. (2009) utilized participant feedback to understand how training affected participants' employment. In addition, Luke et al. (2016) used social network analysis to examine the extent to which post-training collaborations were sustained over time. Finally, evaluators also used qualitative approaches to capture perceived changes to organizational processes, structure, culture, and obtainment of organizational goals (Clark et al., 2022; Provvidenza et al., 2020; Vinson et al., 2019).

Future evaluation approaches proposed by authors. Common limitations noted by authors included the use of simple evaluative designs, small cohorts/sample sizes, only evaluating short-term outcomes, and lack of curriculum evaluation activities. Of the included studies, 33 (70.2%) proposed future evaluation approaches for overcoming the current challenges associated with evaluating K* training programs.

Several authors (Breen et al., 2018; Brownson et al., 2017; Clark et al., 2022; Dagenais et al., 2015; Goodenough et al., 2017; Jacob et al., 2020; Jessani et al., 2019; Morrato et al., 2015; Norton, 2014; Rakhra et al., 2022; Straus et al., 2011; Uneke et al., 2018; Wahabi & Al-Ansary, 2011) reported that stronger evaluative designs are needed. Goodenough et al. (2017) call for multi-variate repeated measure designs, Clark et al. 2022 suggest the use of a control group, while Jacob et al. (2020) suggest the use of combined evaluation approaches to fully understand the impact of program activities. Norton (2014) and Jessani et al. (2019) also reported the need for the use of both pre- and post-measures to examine training outcomes. Relatedly, the need for more rigorous and standardized measures to evaluate the outcomes of training programs were highlighted by Jacob et al. (2020) and Wahabi and Al-Ansary (2011). Goodenough et al. (2017) and Stamatakis et al. (2013) also report that sufficiently large sample sizes are needed to ensure statistical power. Finally, Dagenais et al. (2015) argued that every component of a training program's theory of action and/or logic model should be evaluated to explain the effects obtained.

Authors noted that there was a need for future evaluative activities to examine the longer-term impact of training activities (Baumann et al., 2020; Clark et al., 2022; Froese & Montgomery, 2014; Gerrish & Piercy, 2014; Luke et al., 2016; Moore et al., 2018;

Murong & Nsangi, 2019; Padek et al., 2018; Park et al., 2018; Provvidenza et al., 2020; Ramaswamy et al., 2019; Salloum et al., 2022; Uneke et al., 2017; Uneke et al., 2018). It was suggested that longitudinal (Moore et al., 2018; Park et al., 2018; Provvidenza et al., 2020), time series (Clark et al., 2022), or stepped wedge (Clark et al., 2022) designs may be appropriate approaches for measuring long-term impact and behavior change. Park et al. (2018) recommended that future evaluations expand outcome assessments to consider 'spillover' effects of participants engaging in additional training opportunities outside of the training program being studied. Similarly, Baumann et al., 2020 suggested that evaluators consider opportunity for behavior change within participants' local contexts. Other authors suggested conducting longer-term evaluation activities that examine training outcomes by participants' discipline/field, changes in collaboration with stakeholders (through conducting social network analysis), and the effects training had on participants' employment or position. The use of case studies and qualitative analysis was suggested by Padek et al. (2018) as a potential way to provide more robust feedback on the overall impact of the training program on individual participants.

The authors also highlighted the fact that their current evaluations did not measure the extent to which the various components of the training program produced desired results (Baumann et al., 2020; Goodenough et al., 2017; Olejniczak, 2017). As such, it was suggested that future evaluations assess the relative effectiveness of different training components. Further, Goodenough et al. (2017) suggested that future evaluations examine which individuals might be the best target of training.

Discussion

An increasing number of institutions offer K* training programs to researchers, practitioners, and other stakeholders, thereby potentially providing them the opportunity to ensure findings from research are useful, useable, and utilized. Given the investment in these programs, evaluations have been conducted to identify the effectiveness of K* training programs. To inform our own evaluation of a K* training program, we aimed to understand how other K* training programs were being evaluated. In this section, we provide a summary of the scientometric and content analyses findings, followed by practical implications for evaluators and staff of K* training programs.

Summary of scientometric analysis findings. The findings from the scientometric analysis suggests the concept of K* training is still quite young, and literature regarding the evaluation of K* training programs started to appear in the mid to late 2000's. As such, the number of documents included in this review is small. We note that the literature has grown after 2012 and given the development of the K* field and increasing calls for capacity development in this area, it is reasonable to expect that the growth of the literature will continue. At present, most publications come from the fields of health and implementation science. However, as the field begins to mature, we expect to see researchers from different disciplines (e.g., education and other social science and humanities disciplines) contribute to building the literature base. As the concept of K* training is relatively new, it is understandable that collaboration amongst authors was not widespread. However, limited collaboration may result in a lack of sharing knowledge and resources, thereby resulting in K* training program developers and evaluators risking "re-creating the wheel." As the field continues to develop, we encourage K* program staff and evaluators to connect and collaborate with others engaging in similar types of initiatives.

Summary of content analysis findings. Findings from the content analysis revealed that process and outcome evaluations were the most applied evaluation designs, while commonly used data collection techniques included surveys and interviews. Many of the authors of the included studies recognized the inherent limitations of their evaluations and pointed out issues with small sample sizes, lack of long-term follow-up, and difficulties in measuring long-term impact.

Most studies assessed the 'reactions' of trainees, with evaluators using surveys to obtain feedback from participants. This is likely because surveys can be conducted easily after training sessions. Surveys typically consisted of Likert-style questions, coupled with open-ended items to better understand why trainees might value different program dimensions. The reaction level was typically measured using multiple dimensions (Alsalamah & Callinan, 2021a), such as quality of training content, delivery methods, cohort development opportunities, quality of the trainer, and flexibility and accessibility of training approach. Measuring trainees' reactions are important as "both positive and negative comments can be used to modify the program and to ensure… support for the training program" (Reio et al., 2017). In addition, understanding data captured at Level 1 can form the basis for analyzing subsequent levels of training evaluation. For example, Level 1 reaction data may reveal barriers that impede trainees' learning (Level 2). However, Reio et al., 2017 go on to explain that "favorable reactions to the training do not, by themselves, guarantee that learning (Level 2) or improved performance (Level 3) has occurred," and as such, evaluators must also capture data on trainees' learning and behavior, as well as the downstream impacts that occurred because of the training.

A substantial proportion of studies also assessed the 'learning' of trainees, with many evaluations using non-experimental (i.e., no control group) pre- and post-intervention designs. Pre- and post-intervention questionnaires were often self-report instruments instead of direct measures for assessing capacity. However, we also found several evaluations in which only post-program surveys were conducted to assess participants' skill and knowledge development. Our findings revealed that over two-thirds of studies assessed behavior change. Behavior change was often measured by evaluators between 6- to 12-months post intervention through self-report methods, such as through surveys and interviews. Assessing downstream results were less likely to be evaluated in K* training programs. As a part of these evaluations, evaluators assessed changes in trainees' outputs (e.g., increased publications or grant funding) or broader organizational changes that occurred because of trainees participating in the training intervention.

Recommendations for future K* training evaluations. Based on our experience reviewing the evaluation of K* training programs, we offer several recommendations to future evaluators and program staff who choose to pursue this line of work.

Increase overall rigor of evaluations. To maximize the rigor of evaluation studies, we recommend the use of a logic model or theory of action to guide the development and conduct of studies (Dagenais et al. 2015). Relatedly, evaluators and program staff should clearly identify and evaluate the linkages between the training program components and its outcomes. Evaluators and program staff should also include both process as well as outcome (i.e., impact evaluations) components in the evaluation design. Additional evaluative components that include cost-benefit analyses and curriculum evaluations can also be included to justify programmatic action.

Improve the soundness of research designs. Future evaluations of K* training programs can improve on their research designs by

utilizing pre- and post-intervention designs. We note that evaluators can use pre- and post-tests in two ways – through self-report surveys or through competence assessments. Self-report surveys require participants to state their perceived level of competence in a domain. Self-report surveys can accurately assess self-efficacy and aptitude and are usually inexpensive and easy to use. However, we note that they may not be the best method to assess dynamic processes such as comprehension, and they may suffer from self-reporting and recall bias (Paulhus & Vazire, 2007). Competence assessments require, as a minimum, that trainees can show evidence of competence in K* (e.g., by answering questions on a test). Competence assessments provide a more objective and relevant measure of performance, however, there are potential concerns regarding the possibility that participants in varying settings may have different time available to study for the test, and there may be differences in professional exposure to concepts covered on the test (Grissom et al., 2019). We recommend that evaluators and program staff weigh the benefits and challenges related to self-report and competence assessments and use the type of pre/post-test that works best for their own evaluations. Future evaluations of K* training programs can also improve on their research designs by triangulating data (i.e., collect multiple indicators of the same outcomes). In the current review, self-report measurements of learning and behavior changes were predominately used by evaluators. While this method is simple, it also presents issues related to self-reporting and recall bias. Additionally, if trainees perceive the self-assessment as being linked to performance management, results may be skewed. As such, self-report measures are best used in conjunction with other methods to reliably measure behavior and learning change (Hagger et al., 2020). For example, evaluators can use 360° feedback, whereby a small number of behaviors are assessed by trainees, direct supervisors, and other stakeholders to examine trainees' performance (Kanaslan & Iyem, 2016). We also recommend that future evaluations choose a suitable sample size to detect true treatment effects; however, we recognize that obtaining an adequate sample size can be challenging for training programs due to resource limits and potential issues that may arise related to recruiting, enrolling, and retaining participants (Avellar et al., 2017). Finally, where possible, we recommend that future evaluations include two-group designs (i.e., treatment and control groups).

Evaluate impact using contribution analysis. First, future studies are needed that include and evaluate Level 4 Kirkpatrick evaluation criteria (i.e., downstream results). However, it should be noted that attributing changes in downstream results, such as improved relationships between the research and practice communities, are challenging to assess as they are multifactorial and complex. Other factors, such as national and state legislation on research use, additional training opportunities attended by trainees, and trainees' organizational contexts may also contribute to better K* outcomes. We point to contribution style approaches (e.g., Kok & Schuit, 2012; Morton, 2015) as a potential way in which evaluators can address issues of attribution in future evaluation studies. Contribution analysis is a theory-based evaluation approach that provides a systematic way to arrive at credible causal claims about a program's contribution to change (Mayne, 2008; 2012). The approach involves developing and assessing the evidence for a logic model to explore the program's contribution to observed outcomes. The approach is particularly useful in situations where an experimental (i.e., two-group) design is not feasible (Mayne, 2008; 2012). The findings from a contributions analysis do not provide definitive proof that a program attributed to outcomes but allows evaluators to draw a plausible conclusion that the program has contributed to documented results (Mayne, 2008; 2012).

If possible, longitudinal data should be collected. We recommend that longitudinal data be collected on K* trainees to examine the effects of K* training programs over time. For example, evaluators can follow-up with trainees at pre-determined time periods (e.g., 6-, 12-, 18, and 24-months post training) to examine Level 3 (behavior) and Level 4 (results) outcomes.

How the scoping review has informed CREATEd's evaluation. As argued by Dagenais et al. (2015), a key component of good evaluation planning is the use of a theory of action (ToA) to anchor the implementation of the program to its objectives and intended outcomes, and to provide a basis to formulate the questions and evaluation needs to address. As such, over the period of several meetings, the CREATEd team collaboratively developed a ToA that we use to shape the work of CREATEd and guide evaluation activities. We use an evaluation framework to organize and link relevant ToA outcomes to evaluation measures, evaluation questions, data collection tools, data sources, data analysis procedures, and the year(s) in which data collection will occur. This framework is updated and refined on a yearly basis to ensure that the evaluation team continues to gather data that reflects' CREATEd's priorities. To evaluate CREATEd initiatives, the evaluation team collects data corresponding to its' initiatives, and its' short, intermediate, and long-term outcomes. We provide a summary of each of the tools and/or approaches that we use to collect data in relation to the Fellowship.

We use multiple tools and/or approaches to collect data on the Fellowship. For example:

- **Project Data Records:** we collect data on the number of applications, number of fellows that participate, and fellows' demographic information.
- **Module/Workshop Reaction Surveys:** this survey asks about fellows' thoughts and opinions about the online modules and workshops they complete during the Fellowship.
- **Pre- and Post-Training Surveys:** this survey asks about the extent to which fellows have key knowledge and skills across key competency areas. Fellows complete the survey both pre- and post-engagement in the curriculum.
- **Interviews:** in a series of interviews, the CREATEd evaluation team asks about fellows' experiences and opinions on engaging in the fellowship.
- **Observations:** during the fellowship program, fellows are asked to facilitate event(s) that include diverse stakeholders in the education community. A member of the CREATEd evaluation team observes the event(s) and record notes of what they see and hear.
- **Annual Alumni Follow-Up Survey:** this survey asks fellows about their professional experiences after completing the fellowship program and includes social network items to assess the development and sustainment of fellows' social ties. Fellows complete the survey once a year for up to three years following their graduation from the Fellowship.

CREATEd program staff are committed to using evaluation data to foster program improvement. As such, we regularly engage in reflection exercises to map the evidence we have collected onto the components of the ToA to assess the extent to which our program has contributed to outcomes following the intervention.

Conclusion
The scoping review presents a comprehensive assessment of the K* training evaluation literature. The review was conducted with the required rigor and transparency advocated by Arksey and

O'Malley (2005). The review is comprised of two parts. First, a thorough scientometric analysis was conducted to present the scholarly networks and the research trends. The scientometric analysis identified influential articles, authors, and collaboration networks. The scientometric analysis was followed by a qualitative content analysis to synthesize the methods applied and outcomes assessed within the respective articles. The evidence presented in this review contributes to discussions about how the K* training evaluation literature has grown and is changing over time. We believe the findings from this scoping review will be of interest to evaluators and program designers and will help inform the design of future evaluations of K* training programs.

Data availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Alsalamah A, Callinan C (2021a) Adaptation of Kirkpatrick's four-level model of training criteria to evaluate training programmes for head teachers. Educ Sci 11(3):1–25. https://doi.org/10.3390/educsci11030116

Alsalamah A, Callinan C (2021b) The Kirkpatrick model for training evaluation: Bibliometric analysis after 60 years (1959–2020). Ind Commer Train 54(1): 36–63. https://doi.org/10.1108/ICT-12-2020-0115

Arksey H, O'Malley L (2005) Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 8(1): 19–32. https://doi.org/10.1080/1364557032000119616

Astle B, Reimer-Kirkham S, Theron MJ, Lee JWK (2020) An innovative online knowledge translation curriculum in graduate education. Worldviews Evid Based Nurs 17(3): 229–238. https://doi.org/10.1111/wvn.12440

Avellar S, Borradaile K, Strong D (2017) Tips for enrolling and retaining evaluation participants (Evaluation Technical Assistance Brief Number 4). Mathematica. https://www.mathematica.org/publications/tips-for-enrolling-and-retaining-evaluation-participants-evaluation-technical-assistance-brief. Accessed 18 Aug 2022

Barwick M, Dubrowski R, Petricca K (2020) Knowledge translation: The rise of implementation. Washington, DC: American Institutes for Research. https://ktdrr.org/products/kt-implementation/KT-Implementation-508.pdf. Accessed 18 Aug 2022

Baumann AA, Carothers BJ, Landsverk J, Kryzer E, Aarons GA, Brownson RC, Glisson C, Mittman B, Proctor EK (2020) Evaluation of the Implementation Research Institute: Trainees' publications and grant productivity. Adm Policy Ment Health Ment Health Serv Res. 47: 254–264. https://doi.org/10.1007/s10488-019-00977-4

Breen AV, Twigger K, Duvieusart-Déry C, Boulé J, Borgo A, Fernandes R, Lychek M, Ranby S, Scott C, Whitehouse E (2018) We learn by doing": teaching and learning knowledge translation skills at the graduate level. Can J Scholarsh of Teach Learn 9(1):1–20. https://doi.org/10.5206/cjsotl-rcacea.2018.1.7

Brownson RC, Jacob RR, Carothers BJ, Chambers DA, Colditz GA, Emmons KM, Haire-Joshu D, Kerner JF, Padek M, Pfund C, Sales A (2021) Building the next generation of researchers: Mentored training in dissemination and implementation science. Acad Med 96(1): 86–92. https://doi.org/10.1097/acm.0000000000003750

Brownson RC, Proctor EK, Luke DA, Baumann AA, Staub M, Brown MT, Johnson M (2017) Building capacity for dissemination and implementation research: One university's experience. Implement Sci 12(104):1–12. https://doi.org/10.1186/s13012-017-0634-4

Carlfjord S, Roback K, Nilsen P (2017) Five years' experience of an annual course on implementation science: An evaluation among course participants. Implement Sci 12(101):1–8. https://doi.org/10.1186/s13012-017-0618-4

Chen C, Song M (2019) Visualizing a field of research: A methodology of systematic scientometric reviews. PLoS ONE 14(10):e0223994. https://doi.org/10.1371/journal.pone.0223994

Clark EC, Dhaliwal B, Ciliska D, Neil-Sztramko SE, Steinberg M, Dobbins M (2022) A pragmatic evaluation of a public health knowledge broker mentoring education program: a convergent mixed methods study. Implement Sci Commun 3(18):1–13. https://doi.org/10.1186/s43058-022-00267-5

Cooper A, Rodway J, Read R (2018) Knowledge mobilization practices of educational researchers across Canada. Can J High Educ 48(1):1–21. https://doi.org/10.7202/1050839ar

Cunningham-Erves J, Stewart E, Duke J, Akohoue SA, Rowen N, Lee O, Miller ST (2021) Training researchers in dissemination of study results to research participants and communities. Transl Behav Med 11(7): 1411–1419. https://doi.org/10.1093/tbm/ibab023

Dagenais C, Somé TD, Boileau-Falardeau M, McSween-Cadieux E, Ridde V (2015) Collaborative development and implementation of a knowledge brokering program to promote research use in Burkina Faso, West Africa. Global Health Action 8(1):1–11. https://doi.org/10.3402/gha.v8.26004

Davis D, Davis D (2009) Formal educational interventions. In Straus S, Tetroe J, & Graham G (ed) Knowledge translation in healthcare: Moving from evidence to practice. Wiley-Blackwell, p.113-122

Denaro K, Dennin K, Dennin M, Sato B (2022) Identifying systemic inequity in higher education and opportunities for improvement. PLoS ONE 17(4):e0264059. https://doi.org/10.1371/journal.pone.0264059

Fahim C, Kasperavicius D, Beckett R, Quinn de Launay K, Chandraraj A, Crupi A, Theivendrampillai S, Straus SE (2023) Funding change: an environmental scan of research funders' knowledge translation strategic plans and initiatives across 10 high-income countries/regions. FACETS 8(1): 1–26. https://doi.org/10.1139/facets-2022-0124

Farley-Ripple E, May H, Karpyn A, Tilley K, McDonough K (2018) Rethinking connections between research and practice in education: A conceptual framework. Educ Res 47(4): 235–245. https://doi.org/10.3102/0013189X18761042

Froese KA, Montgomery J (2014) From research to practice: The process of training school psychologists as knowledge transfer professionals. Proc- Soc Behav Sci 141: 375–381. https://doi.org/10.1016/j.sbspro.2014.05.066

Gaid D, Mate K, Ahmed S, Thomas A, Bussières A (2022) Nationwide environmental scan of knowledge brokers training. J Continuing Educ Health Professions 42(1): e3–e11. https://doi.org/10.1097/ceh.0000000000000355

Garritzmann JL, Häusermann S, Palier B (2023) Social investments in the knowledge economy: The politics of inclusive, stratified, and targeted reforms across the globe. Soc Policy Adm 57(1): 87–101. https://doi.org/10.1111/spol.12874

Georgalakis J, Rose P (eds) (2021) Maximising the impact of global development research – a new approach to knowledge brokering. Brighton: Institute of Development Studies. https://doi.org/10.35648/20.500.12413/11781/ii365

Gerrish K, Piercy H (2014) Capacity development for knowledge translation: Evaluation of an experiential approach through secondment opportunities. Worldviews Evidence-Based Nurs 11(3): 209–216. https://doi.org/10.1111/wvn.12038

Golhasany H, Harvey B (2023) Capacity development for knowledge mobilization: A scoping review of the concepts and practices. Humanit Soc Sci Commun 10(235):1–12. https://doi.org/10.1057/s41599-023-01733-8

Goodenough B, Fleming R, Young M, Burns K, Jones C, Forbes F (2017) Raising awareness of research evidence among health professionals delivering dementia care: Are knowledge translation workshops useful? Gerontol Geriatr Educ 38(4): 392–406. https://doi.org/10.1080/02701960.2016.1247064

Greenhalgh T, Russell J (2006) Promoting the skills of knowledge translation in an online master of science course in primary health care. J Continuing Educ Health Professions 26(2): 100–108. https://doi.org/10.1002/chp.58

Grissom JA, Mitani H, Woo DS (2019) Principal preparation programs and principal outcomes. Educ Admin Quart 55(1): 73–115. https://doi.org/10.1177/0013161X18785865

Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T (eds) (2020) The handbook of behavior change. Cambridge University Press. https://doi.org/10.1017/9781108677318

Halsall T, McCann E, Armstrong J (2022) Engaging young people within a collaborative knowledge mobilization network: Development and evaluation. Health Expect 25(2):617–627. https://doi.org/10.1111/hex.13409

Hartling L, Elliott SA, Buckreus K, Leung J (2021) Development and evaluation of a parent advisory group to inform a research program for knowledge translation in child health. Res Involv Engagem 7(38):1–13. https://doi.org/10.1186/s40900-021-00280-3

Hess J, Siegelman J, Lamm R, Moll J (2013) An innovative adult-learning curriculum merging evidence-based medicine, knowledge translation, and research design. Ann Emerg Med 62(4):S159. https://doi.org/10.1016/j.annemergmed.2013.07.267

Hilbig A, Proske A, Damnik G, Faselt F, Körndle H (2013) Designing workplace learning and knowledge exchange - a postgraduate training program for professionals in SME. In: Proceedings of the 5th International Conference on Computer Supported Education - CSEDU, Dresden University of Technology, Germany, p. 635-638. https://doi.org/10.5220/0004345706350638

Holmes BJ, Scarrow G, Schellenberg M (2012) Translating evidence into practice: The role of health research funders. Implement Sci 7(39):1–10. https://doi.org/10.1186/1748-5908-7-39

Holmes BJ, Schellenberg M, Schell K, Scarrow G (2014) How funding agencies can support research use in healthcare: An online province-wide survey to determine knowledge translation training needs. Implement Sci 9(71):1–10. 10.1186%2F1748-5908-9-71

Honig MI, Coburn C (2008) Evidence-based decision making in school district central offices: Toward a policy and research agenda. Educ Policy 22(4):578–608. https://doi.org/10.1177/0895904807307067

Jacob RR, Gacad A, Padek M, Colditz GA, Emmons KM, Kerner JF, Chambers DA, Brownson RC (2020) Mentored training and its association with dissemination and implementation research output: A quasi-experimental evaluation. Implement Sci 15(30):1–8. https://doi.org/10.1186/s13012-020-00994-0

Jessani NS, Hendrick L, Nicol L, Young T (2019) University curricula in evidence-informed decision making and knowledge translation: Integrating best practice, innovation, and experience for effective teaching and learning. Front Public Health 7(313):1–13. https://doi.org/10.3389/fpubh.2019.00313

Jones K, Armstrong R, Pettman T, Waters E (2015) Knowledge translation for researchers: Developing training to support public health researchers KTE efforts. J Public Health 37(2):364–366. https://doi.org/10.1093/pubmed/fdv076

Kanaslan EK, Iyem C (2016) Is 360-degree feedback appraisal an effective way of performance evaluation? Int J Acad Res Bus Soc Sci 6(5):172–182. https://doi.org/10.6007/IJARBSS/v6-i5/2124

Kho ME, Estey EA, DeForge RT, Mak L, Bell BL (2009) Riding the knowledge translation roundabout: Lessons learned from the Canadian Institutes of Health Research Summer Institute in knowledge translation. Implement Sci 4(33):1–7. https://doi.org/10.1186/1748-5908-4-33

Kirkpatrick D, Kirkpatrick J (2006) Evaluating training programs: The four levels. Berrett-Koehler Publishers

Kok MO, Schuit AJ (2012) Contribution mapping: A method for mapping the contribution of research to enhance its impact. Health Res Policy Syst 10(21):1–16. https://doi.org/10.1186/1478-4505-10-21

Lockton M, Caduff A, Rehm M, Daly AJ (2022) Refocusing the lens on knowledge mobilization: An exploration of knowledge brokers in practice and policy. Educ Policy Manag 7:1–24. https://doi.org/10.53106/251889252022060007001

Lo Hog Tian JM, Watson JR, Deyman M, Tran B, Kerber P, Nanami K, Norris D, Samson K, Cioppa L, Murphy M, Mcgee A, Ajiboye M, Chambers LA, Worthington C, Rourke SB (2022) Building capacity in quantitative research and data storytelling to enhance knowledge translation: A training curriculum for peer researchers. Res Involv Engagem 8(1):69. https://doi.org/10.21203/rs.3.rs-1420986/v1

Luke DA, Baumann AA, Carothers BJ, Landsverk J, Proctor EK (2016) Forging a link between mentoring and collaboration: A new training model for implementation science. Implement Sci 11(137):1–12. https://doi.org/10.1186/s13012-016-0499-y

Malin JR, Brown C, Ion G, van Ackeren I, Bremm N, Luzmore R, Flood J, Rind GM (2020) World-wide barriers and enablers to achieving evidence-informed practice in education: What can be learnt from Spain, England, the United States, and Germany? Humanit Soc Sci Commun 7(99):1–14. https://doi.org/10.1057/s41599-020-00587-8

Mallidou AA, Atherton P, Chan L, Frisch N, Glegg S, Scarrow G (2018) Core knowledge translation competencies: A scoping review. BMC Health Serv Res 18(502):1–15. https://doi.org/10.1186/s12913-018-3314-4

Marriott BR, Rodriguez AL, Landes SJ, Lewis CC, Comtois KA (2015) A methodology for enhancing implementation science proposals: Comparison of face-to-face versus virtual workshops. Implement Sci 11(62):1–11. https://doi.org/10.1186/s13012-016-0429-z

Matthews L, Simpson SA (2020) Evaluation of behavior change interventions. In: Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T (ed) The handbook of behavior change. Cambridge University Press, p. 318-332. https://doi.org/10.1017/9781108677318

Mayne J (2012) Contribution analysis: Coming of age. Evaluation 18(3):270–280. https://doi.org/10.1177/1356389012451663

Mayne J (2008) Contribution analysis: An approach to exploring cause and effect. International Learning and Change Brief 16

Mbuagbaw L, Thabane L, Ongolo-Zogo P (2014) Training Cameroonian researchers on pragmatic knowledge translation trials: A workshop report. Pan Afr Med J 19(190):1–6. https://doi.org/10.11604/pamj.2014.19.190.5492

Meissner HI, Glasgow RE, Vinson CA, Chambers D, Brownson RC, Green LW, Ammerman AS, Weiner BJ, Mittman B (2013) The U.S. training institute for dissemination and implementation research in health. Implement Sci 8(12):1–9. https://doi.org/10.1186/1748-5908-8-12

Mishra L, Banerjee AT, MacLennan ME, Gorczynski PF, Zinszer KA (2011) Wanted: Interdisciplinary, multidisciplinary, and knowledge translation and exchange training for students of public health. Can J Public Health 102(6):424–426. https://doi.org/10.1007/bf03404192

Moore JE, Shusmita R, Park JS, Khan S, Straus SE (2018) Longitudinal evaluation of a course to build core competencies in implementation practice. Implement Sci 13(106):1–13. https://doi.org/10.1186/s13012-018-0800-3

Morrato EH, Rabin B, Proctor J, Cicutto LC, Battaglia CT, Lambert-Kerzner A, Leeman-Castillo B, Prahl-Wretling M, Nuechterlein B, Glasgow RE, Kempe A (2015) Bringing it home: Expanding the local reach of dissemination and implementation training via a university-based workshop. Implement Sci 10(94):1–12. https://doi.org/10.1186/s13012-015-0281-6

Morton S (2015) Progressing research impact assessment: a 'contributions' approach. Res Eval 24(4):405–419. https://doi.org/10.1093/reseval/rvv016

Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 18(143):1–7. https://doi.org/10.1186/s12874-018-0611-x

Murong M, Nsangi A (2019) Reflections on experiences in doctoral training and its contribution to knowledge translation in an African environment. Makerere University. Kampala, Uganda. https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57991/IDL-57991.pdf?sequence=2&isAllowed=y

Murunga VI, Oronje RN, Bates I, Tagoe N, Pulford J (2020) Review of published evidence on knowledge translation capacity, practice and support among researchers and research institutions in low- and middle-income countries. Health Res Policy Sys 18(16):1–21. https://doi.org/10.1186/s12961-019-0524-0

Ndalameta-Theo EM, Monde MW, Wales A, Moyo M, Vallis J, Mudenda C, Kanyengo CW, Mwafulilwa CM (2021) Building the capacity of African health librarians to become knowledge brokers through a knowledge broker learning programme. SCECSAL. https://www.scecsal.org/publications/papers2021/scecsal_papers2021_ndalameta-theo.pdf

Newman ME (2004) Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences 101:5200–5205. https://doi.org/10.1073/pnas.0307545100n

Norton WE (2014) On academics: Advancing the science and practice of dissemination and implementation health: A novel course for public health students and academic researchers. Public Health Rep. 129(6):536–542. https://doi.org/10.1177/003335491412900613

Olejniczak K (2017) The game of knowledge brokering: A new method for increasing evaluation use. Am J Eval 38(4):554–576. https://doi.org/10.1177/1098214017716326

Oronje RN, Mukiira C, Kahurani E, Murunga V (2022) Training and mentorship as a tool for building African researchers' capacity in knowledge translation. PLoS ONE 17(3):e0266106. https://doi.org/10.1371/journal.pone.0266106

Padek M, Mir N, Jacob RR, Chambers DA, Dobbins M, Emmons KM, Kerner J, Kumanyika S, Pfund C, Proctor EK, Stange KC, Brownson RC (2018) Training scholars in dissemination and implementation research for cancer prevention and control: A mentored approach. Implement Sci 13(18):1–13. https://doi.org/10.1186/s13012-018-0711-3

Park JS, Moore JE, Sayal R, Holmes BJ, Scarrow G, Graham ID, Jeffs L, Timmings C, Rashid S, Johnson AM, Straus SE (2018) Evaluation of the "Foundations in Knowledge Translation" training initiative: Preparing end users to practice KT. Implement Sci 13(63):1–13. https://doi.org/10.1186/s13012-018-0755-4

Paulhus DL, Vazire S (2007) The self-report method. In: Robins RW, Fraley RC, Krueger RF (ed) Handbook of research methods in personality psychology. The Guilford Press, p. 224-239

Phipps D, Jensen KE, Johnny M, Poetz A (2016) Supporting knowledge mobilization and research impact strategies in grant applications. J Res Admin 47(2):49–67. https://files.eric.ed.gov/fulltext/EJ1152268.pdf

Proctor EK, Ramsey AT, Brown MT, Malone S, Hooley C, McKay V (2019) Training in Implementation Practice Leadership (TRIPLE): Evaluation of a novel practice change strategy in behavioral health organizations. Implement Sci 14(66):1–11. https://doi.org/10.1186/s13012-019-0906-2

Provvidenza C, Townley A, Wincentak J, Peacocke S, Kingsnorth S (2020) Building knowledge translation competency in a community-based hospital: a practice-informed curriculum for healthcare providers, researchers, and leadership. Implement Sci 15(54):1–14. https://doi.org/10.1186/s13012-020-01013-y

Rakhra A, Hooley C, Fort M, Weber MB, Price L, Nguyen HL, Ramirez M, Muula AS, Hosseinipour M, Apusiga K, Davila-Roman V, Gyamfi J, Adjei KGA, Andesia J, Fitzpatrick A, Launois P, Baumann AA (2022) The WHO-TDR Dissemination and Implementation Massive Open Online Course (MOOC): evaluation and lessons learned from eight low- and middle-income countries. Research Square. https://doi.org/10.21203/rs.3.rs-1455034/v1

Ramaswamy R, Mosnier J, Reed K, Powell BJ, Schenck AP (2019) Building capacity for Public Health 3.0: Introducing implementation science into an MPH curriculum. Implement Sci 14(18):1–10. https://doi.org/10.1186/s13012-019-0866-6

Reio TG, Rocco TS, Smith DH, Chang E (2017) A critique of Kirkpatrick's evaluation model. New Horiz Adult Educ Human Resour Dev 29(2):35–53. https://doi.org/10.1002/nha3.20178

Rycroft-Smith L (2022) Knowledge brokering to bridge the research-practice gap in education: Where are we now? Rev Educ 10(1):e3341. https://doi.org/10.1002/rev3.3341

Salloum RG, LeLaurin JH, Nakkash R, Akl EA, Parascandola M, Ricciardone MD, Elbejjani M, Kabakian-Khasholian T, Lee J, El-Jardali F, Shelley D, Vinson

CA (2022) Developing capacity in dissemination and implementation research in the Eastern Mediterranean region: evaluation of a training workshop. Glob Implement Res Appl 2(4):340–349. https://doi.org/10.1007/s43477-022-00067-y

Santacroce SJ, Leeman J, Song M (2017) A training program for nurse scientists to promote intervention translation. Nursing Outlook 66(2):149–156. https://doi.org/10.1016/j.outlook.2017.09.003

Shaxson L, Bielak A, Ahmed I, Brien D, Conant B, Fisher C, Gwyn E, Klerkx L, Middleton A, Morton S, Pant L, Phipps D (2012) Expanding our understanding of K* (Kt, KE, Ktt, KMb, KB, KM, etc.). In: A concept paper emerging from the K* conference held in Hamilton, Ontario, Canada, April 2012, United Nations University. https://assets.publishing.service.gov.uk/media/57a08a6e40f0b649740005ba/KStar_ConceptPaper_FINAL_Oct29_WEBsmaller.pdf

Stamatakis KA, Norton WE, Stirman SW, Melvin C, Brownson RC (2013) Developing the next generation of dissemination and implementation researchers: insights from initial trainees. Implement Sci 8(29):1–6. https://doi.org/10.1186/1748-5908-8-29

Straus SE, Brouwers M, Johnson D, Lavis JN, Légaré F, Majumdar SR, McKibbon KA, Sales AE, Stacey D, Klein G, Grimshaw J (2011) Core competencies in the science and practice of knowledge translation: Description of a Canadian strategic training initiative. Implement Sci 6(127):1–7. https://doi.org/10.1186/1748-5908-6-127

Tabak RG, Padek MM, Kerner JF, Stange KC, Proctor EK, Dobbins MJ, Colditz GA, Chambers DA, Brownson RC (2017) Dissemination and implementation science training needs: Insights from practitioners and researchers. Am J Prev Med 52(3):S322–S329. https://doi.org/10.1016/j.amepre.2016.10.005

Tait H, Williamson A (2019) A literature review of knowledge translation and partnership research training programs for health researchers. Health Res Pol Syst 17(98):1–14. https://doi.org/10.1186/s12961-019-0497-z

Uneke CJ, Ezeoha AE, Uro-Chukwu HC, Ezeonu CT, Igboji J (2018) Promoting researchers and policy-makers collaboration in evidence-informed policy-making in Nigeria: Outcome of a two-way secondment model between university and health ministry. Int J Health Policy Manag 7(6):522–531. https://doi.org/10.15171/ijhpm.2017.123

Uneke CJ, Sombie I, Uro-Chukwu HC, Johnson E, Okonufua F (2017) Using Equitable Impact Sensitive Tool (EQUIST) and knowledge translation to promote evidence to policy link in maternal and child health: report of first EQUIST training workshop in Nigeria. Pan African Med J 28(37):1–10. https://doi.org/10.11604/pamj.2017.28.37.13269

Van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84(2):523–538. https://doi.org/10.1007/s11192-009-0146-3

Van Eck NJ, Waltman L (2014) Visualizing bibliometric networks. In: Ding Y, Rousseau R, Wolfram D (ed) Measuring scholarly impact. Springer, p. 285-320

Vinson CA, Clyne M, Cardoza N, Emmons KM (2019) Building capacity: A cross-sectional evaluation of the US Training Institute for Dissemination and Implementation Research in Health. Implement Sci 14(97):1–6. https://doi.org/10.1186/s13012-019-0947-6

Wahabi HA, Al-Ansary LA (2011) Innovative teaching methods for capacity building in knowledge translation. BMC Med Educ 11(85):1–10. https://doi.org/10.1186/1472-6920-11-85stylefix

Walter I, Nutley S, Davies H (2005) What works to promote evidence-based practice? A cross-sector review. Evid Policy 1(3):335–363. https://doi.org/10.1332/1744264054851612

Acknowledgements

Author contributions

Conception and design: S.S. Data collection and analysis: S.S., J.W. and M.S. Original draft: S.S. and J.W. Review and editing: S.S. and J.W.

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was not required as the study did not involve human participants.

Informed consent

Informed consent was not required as the study did not involve human participants.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-023-02403-5.

Correspondence and requests for materials should be addressed to Samantha Shewchuk.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.