Received: 5 April 2024 | Revised: 30 July 2024 | Accepted: 27 August 2024

DOI: 10.1002/hpja.922

RESEARCH ARTICLE

Health Promotion Journal of Australia | Australian Health Promotion Association | WILEY

Development and testing of the capacity of organisations for system practices scale

Melinda Craike[1,2] | Larissa Bartlett[3] | Amy Mowle[1] | Therese Riley[1] | Michelle Krahe[4,5] | Bojana Klepac[1]

[1]Mitchell Institute for Education and Health Policy, Victoria University, Melbourne, Victoria, Australia

[2]Institute for Health and Sport, Victoria University, Melbourne, Victoria, Australia

[3]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

[4]Health Group, Griffith University, Brisbane, Queensland, Australia

[5]College of Medicine and Dentistry, James Cook University, Cairns, Queensland, Australia

Correspondence
Melinda Craike, Institute for Health and Sport, Victoria University, Melbourne, Vic, Australia.
Email: melinda.craike@vu.edu.au

Funding information
Paul Ramsay Foundation, Grant/Award Number: 681

Handling editor: Mark Robinson

Abstract

Background: Systems change approaches are increasingly adopted in public health to address complex problems. It is important that measures of systems change be developed so that the effects of systems change on health outcomes can be evaluated. Organisational practices are potential levers for systems change. However, robust measures of organisational capacity to engage in these practices are lacking. Informed by the Theory of Systems Change, we developed and tested the Capacity of Organisations for System Practices (COSP) scale. The COSP scale comprises four inter-related system practices within organisations—adaptation, alignment, collaboration and evidence-driven action and learning.

Methods: We applied a three-stage process: (1) Item generation; (2) Scale pre-testing; and (3) Structural analyses. Item response theory tests and semantic review, together with factor analytic techniques, were applied to refine the item set and determine the scale structure.

Results: An initial pool of 97 items was generated and pre-tested with six content experts and four target audience representatives. Modifications resulted in 60 items. In total, 126 participants provided data for the structural analysis. A second-order hierarchical four-factor model fit the data better than the more basic correlated factor model ($\Delta\chi^2 = 1.758$, $p = .415$). The fit indices for the final 31-item model were acceptable (RMSEA = .084, TLI = .819).

Conclusions: The COSP scale is ready for further testing to ensure construct validity, stability and utility.

So What? Once validated, the Capacity of Organisations for System Practices (COSP) scale has the potential to advance the theory and practice of systems change approaches.

KEYWORDS

evaluation, measure, place-based approach, scale, systems approach, systems change, systems perspective

1 | BACKGROUND

Seminal works that inform public health, such as The Ottawa Charter for Health Promotion,[1] draw attention to addressing non-medical determinants of health through cross-sector collaboration and coordinated action, adapting and aligning strategies and programs with local needs, and the creation of supportive environments for health. These principles are echoed in systems change approaches, which have been increasingly adopted to address complex public health challenges over the past decade.[2–9] Given the increasing attention on systems change it is important to assess the effect of systems change on health outcomes. Testing the effect of systems change on health outcomes requires the development of robust tools to measure systems change.

Systems change occurs when 'levers for change are targeted that can trigger shifts across systems components',[10] (p. 194). There are a range of potential levers for systems change, including the system practices embedded in organisations. Organisations provide programs, services and other support and thus form essential infrastructure in a community and are one of the vehicles through which systems operate.[11,12] Further, organisations can influence the broader system[13,14] by, for instance, influencing public policy through advocacy or influencing other organisations through the diffusion of knowledge.[15] It follows that changing the capacity of organisations to engage in system practices could be a vehicle for system change. There are currently no validated tools to measure organisational capacity for system practices that could contribute to system change. The development of such a measure would provide several opportunities for future research, for instance, testing whether organisational capacity for system practices contributes to system change and, ultimately, health outcomes.[16]

Measurement is essential for empirical research as it enables knowledge to be attained about a particular phenomenon of interest.[17] Measurement scales are tools that are used to give scores to phenomena that cannot be measured directly.[17] To maximise utility, a scale that measures organisational capacity for system practices should be 'pragmatic' and feasible to use not only for research purposes but also in practice and policy.[18] Pragmatic measures are designed for actionability, low participant burden and to allow for the interpretation of findings without external facilitation.[18,19]

For guidance on identifying the components of organisational capacity for system practices, we look to the Theory of Systems Change.[20] The Theory of Systems Change is an initial middle-range theory that proposes four interrelated practices are necessary for systems change: (1) adaptation to external opportunities and challenges; (2) alignment with the strengths and needs of the target population; (3) collaboration within and across sectors and (4) engagement in evidence-driven action and learning. In this context, system practices are conceptualised as approaches, processes and/or behaviours that can influence systems change. Engaging in and sustaining these practices requires attention to be paid to three inter-related dimensions of capacity (1) individual-level capacities, such as confidence, skills, attitudes and knowledge[21,22]; (2) organisational level capacities, such as organisational culture and leadership, systems and structures (e.g., IT systems, procedures and policies)[23] and (3) an enabling environment, including public policy, legislation and broader political conditions.[24] Based on the Theory of Systems Change, organisational capacity to adapt, align, collaborate and engage in evidence-driven action and learning could comprise a measure of an organisation's capacity to engage in system practices to influence systems change efforts.

It is well established that addressing the determinants of health requires collaboration across health and non-health-related sectors and organisations.[14,25] As such, a scale that measures organisational capacity for system practices must be appropriate for application not only within organisations providing health-related programs and services but also for those providing, for example, education and employment-related programs and services. Importantly, developing this scale is the first step and once developed, it could be tested and refined in other components of a system, building a more comprehensive picture of the capacity for practices necessary for systems change.

1.1 | Aims

We aimed to develop a scale to measure organisational capacity for system practices. We call the scale the Capacity of Organisations for System Practices (COSP) scale. In this article, we report on the item development and initial testing of the COSP scale. The objectives of this study were to:

1. Generate items to measure organisational capacity for system practices;
2. Conduct pre-testing to provide an indication of clarity, acceptability, useability, length and face validity of the initial set of items and refine them before conducting a structural analysis; and
3. Perform structural and item performance analyses to test the scale's internal consistency and factor structure and determine the best-performing items to maximise its psychometric validity and performance.

1.2 | Research and community context

Pathways in Place aims to simultaneously develop the evidence base for place-based systems change initiatives and build capacity to support the design and implementation of effective place-based systems change initiatives to improve population health outcomes. Jointly led by research teams based at two Australian universities (Victoria University, Victoria and Griffith University, Queensland), we work closely with partner communities to build capacity for effective place-based systems change approaches. This article refers to the work of the Pathways in Place-Victoria University (www.pathwaysinplace.com.au/victoria-university) team, which focuses on optimising pathways from education to employment for young people (15–24). Being place-based, we predominantly work with partner communities in the City of Brimbank (referred to as 'Brimbank') located in the urban regions

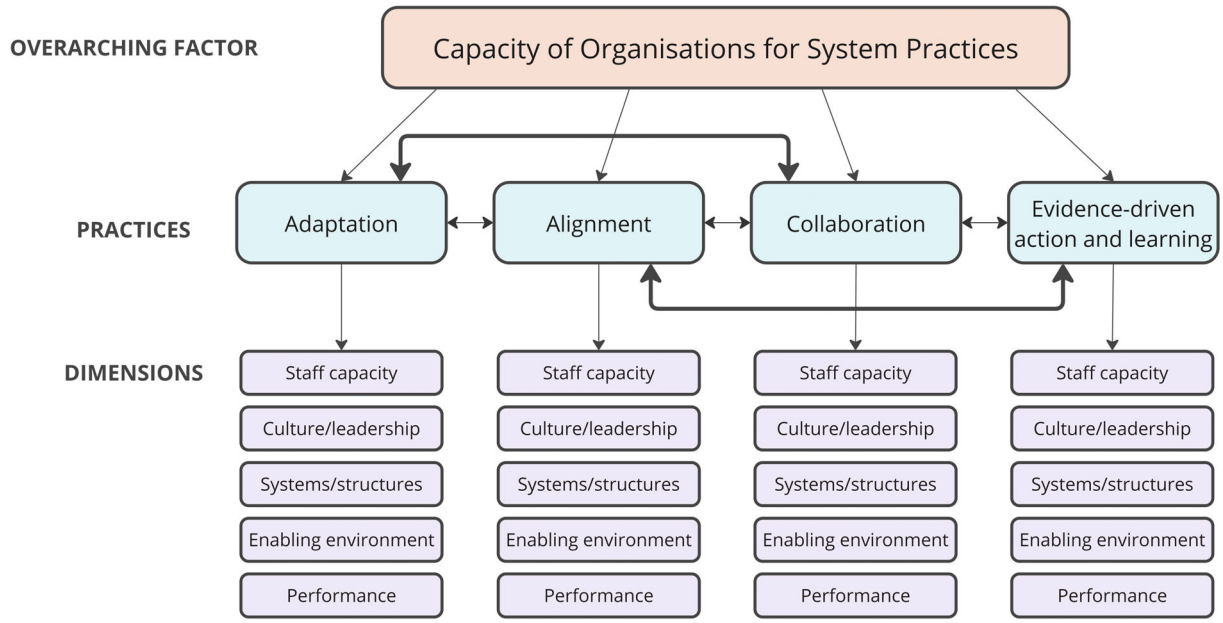

FIGURE 1 Structure of the capacity of organisations for system practices scale.

of Melbourne, Victoria. We also seek to work with and scale out our research innovations to other communities in Australia and internationally.

2 | METHODS

Based on recommendations[17,26] we developed the COSP scale using a three-stage process. Before commencement, this project received ethical approval from the Victoria University Human Research Ethics Committee. For ease of reading and interpretation, we structure this section by combining methods with results for each stage.

2.1 | Stage 1: Item generation

The initial pool of items was generated by a team of six researchers, including five co-authors (M.C., B.K., M.K., T.R., A.M.), who bring expertise in systems thinking, public health, and measurement. The items were based on the practices of adaptation, alignment, collaboration and evidence-driven action and learning (see definitions below) according to capacity and functioning dimensions. Capacity and functioning dimensions include staff capacity, organisational culture and leadership, systems and structures, enabling environment and performance. The items were deliberately designed to be applicable across sectors, rather than targeting health-related organisations specifically.

First, we extensively reviewed the literature from various disciplines and fields to articulate and define the concepts we were attempting to measure. Based on the literature, we expected that the practices would be interrelated[27–32] (see Figure 1). Next, we generated items using deductive methods, which included a literature review of existing scales[33–69] and open discussions between four researchers, three of whom are co-authors (M.C., B.K., M.K.). The generation of items were subject to several rounds of workshopping and revision.

During the item generation stage, we also considered the form, wording, and appropriate number of items, suitable scale instructions, and the types of responses the items were designed to elicit.[17,26] Some items were purposefully worded to relate to young people, who are the focus of Pathways in Place-Victoria University work. The initial scale had 97 items and included a range of 5-point Likert style response sets.

2.1.1 | Definitions

Practice items

Adaptation. Adaptation is the ability to respond to external changes, both opportunities and challenges.[43,70]

Alignment. Alignment is 'sharing the same or complementary perceived needs' of the target population, and 'how these needs will be met', across various system levels (e.g., practitioners, researchers, policymakers)[69] (pp. 8–9).

Collaboration. Collaboration is any joint activity by two or more parties to link or share information, resources, activities, and capabilities to achieve aims that no single party could have achieved separately.[71]

Evidence-driven action and learning. The initial item pool was developed to assess evidence-informed practice. This was changed to the

broader, more encompassing term of evidence-driven action and learning following preliminary testing and feedback.

Evidence-driven action and learning is the continual process that guides decision-making and action. Cycles of evidence-driven action and learning include: (1) *Situation analysis and problem framing*: using diverse sources of evidence to understand the current circumstances and to frame a problem. Attention is paid to the broader context in which the problem manifests (see Refs. 72,73); (2) *Coordinated action*: following the identification of problems and underlying causes, co-created, purposefully coordinated actions that target multiple levels of the system are designed and implemented.[72] (3) *Monitoring and evaluation*: progress is monitored to determine whether and how actions achieve desired outcome(s) (or not); findings are acted on. Attention is paid to unexpected outcomes; (4) *Communication and dissemination*: findings and knowledge are communicated and disseminated to facilitate application across the system.[74–76]

2.1.2 | Capacity and functioning dimensions

Capacity is defined as 'the ability of individuals, organisations, or broader systems to perform appropriate functions and address issues and concerns effectively, efficiently, and sustainably'[20,77] (p. 4). For our purposes, the capacity dimensions include staff capacity, organisational culture and leadership, systems and structures,[78,79] and the enabling environment. *Staff capacity*: staff confidence, skills, attitudes, and knowledge related to adaptation, alignment, collaboration and engaging in evidence-driven action and learning.[21,22] *Culture and leadership*: organisational leadership and culture related to adaptation, alignment, collaboration, and engaging in evidence-driven action and learning. *Systems and structures*: organisational IT systems, procedures and policies related to adaptation, alignment, collaboration and engaging in evidence-driven action and learning. The *enabling environment* includes a range of interrelated conditions, including public policy, legal, and political conditions related to the staff and the organisational capacity for adaptation, alignment, collaboration and evidence-driven action and learning (adapted from Ref. 24). One aspect of an enabling environment is funding processes, which we used as our indicator of the enabling environment. The *functioning* dimension is how well the organisation performs at each system practice. The proposed structure of the COSP scale can be seen in Figure 1.

2.2 | Stage 2: Pre-testing

We sought the opinions of content experts (experts in measurement or the target domains) and target audience representatives (potential scale users) to indicate the scale's clarity, acceptability, useability, length, and face validity.

2.2.1 | Data collection

During November and December 2021, eight content experts and six target audience representatives, identified from the co-authors' networks and through snowballing, were invited to participate in pre-testing of the scale.

All participants were provided with information about the study and were required to provide signed informed consent before participating.

Target audience representatives were provided with the draft scale and asked to provide comments on the scale itself and address a series of questions. The draft scale included four sections evidence-informed practice (49 items), adaptation (22 items), alignment (11 items) and collaboration (19 items). The questions related to the relevance of the items for their organisation; whether the items were clear and easy to understand; whether any items were difficult to understand; suggestions for improving the clarity and/or suitability of the items; and the length of time taken to complete the scale.

Content experts were provided with an expert briefing document containing information about project aims and background, the rationale for the development of a new scale, the design of the scale, and an Excel spreadsheet to provide comments about individual items and the sets of items (e.g., are the items consistent with the definition, are any items missing), their suggestions for rewording the items and to assess the overall scale in terms of its clarity, acceptability, useability, length and validity.

2.2.2 | Data analysis

We semantically examined participant responses and looked for recurring themes across participants. Four co-authors (M.C., B.K., T.R., A.M.) discussed the findings and determined the modifications to be made, balancing both the theory and rationale for items with participant feedback.

2.2.3 | Results and modifications

Six content experts and four target audience representatives consented to be involved in the study. The main feedback and changes to the initial scale are shown in Table 1. Along with general structural feedback, participants also commented on individual items and wording, and items were modified accordingly. Following modifications, there were 60 items in the scale—adaptation (12 items) alignment (13 items), collaboration (17 items), and evidence-driven action and learning (18 items). The full list of items and associated dimension is available as File S1. These 60 items were then tested in Stage 3.

2.3 | Stage 3: Structural analyses

This stage aimed to assess the scale's internal consistency, factor structure, and item performance to guide item modification.

2.3.1 | Data collection

Data were collected via online and paper surveys. Staff involved in the planning, implementation and/or evaluation of programs, services,

and policies related to education, employment, or mental wellbeing for young people (aged 15–24) were invited to participate in the study. At the time of this study, due to COVID-19 restrictions, the Victorian State Government had stopped all research in schools. Therefore, we had to exclude teachers and other relevant personnel working in secondary schools.

We had a two-step recruitment process. First, we recruited participants who worked in Brimbank as the focus of Pathways in Place-Victoria University initiatives. The co-authors compiled a database of organisations within Brimbank that provide education, employment, or mental well-being related services and programs.[80] The database was developed as part of a comprehensive community asset mapping project that was undertaken as part of Pathways in Place, in which we identified the range of services, programs and organisations that support young people's well-being, education and employment in Brimbank.[80] This approach helped to ensure that we invited participants from a range of organisations across sectors. Managers of the identified organisations were emailed and invited to nominate relevant staff to participate in the study. Those who completed the survey were invited to pass on the invitation to participate to their colleagues in Brimbank. We also disseminated information about the study through local and social media inviting people to participate. The second step involved opening the study up to a broader sample of participants working in Australia. We recruited these participants through local media and social media networks, Pathways in Place-Victoria University research team networks, and via an invitation on the Pathways in Place website. All participants were provided with a plain language statement and had to indicate their consent before their involvement in the study.

2.3.2 | Statistical methods

Items were assigned to the pre-defined, theory-based four-factor structure (i.e., the four practices) before analysis. All cases with missing responses were removed. The normality of item response distributions was inspected using histograms. As the data was ordinal and polytomous, polychoric item–item and item–scale correlations were assessed for each factor. The uni-dimensional structure of the prescribed factors was checked using serial maximum likelihood factor analyses. Standard indices of model fit are reported, following guidelines for latent variable analyses and factor analytic procedures.[81–84] The Root Mean Squared Error of Approximation (RMSEA) measures poor fit. Ideally, it should be <.06. The Tucker Lewis Index (TLI) and Comparative Fit Index (CFI) are measures of good fit. Ideally, they should be >.95. The χ^2 coefficient and *p*-value indicate the degree to which the covariance matrix derived from the model represents the population covariance. Ideally, χ^2 should not be significant. Cronbach's test of internal consistency was assessed for each factor (α).

Confirmatory Factor Analysis (CFA) was then run to establish the baseline fit for the prescribed four-factor model. To guide the refinement of the scale, parameters from Item Response Theory (IRT) modelling were inspected.[85,86] Set with the Generalised Partial Credit Model (GPCM), IRT was conducted on each uni-dimensional factor. The 'a' score from IRT is the discrimination parameter and ideally should be >.8. The 'b' scores are difficulty parameters that indicate the extent to which the item detects different levels of agreement. Modification Indices (M.I.) provide an estimated drop in the value of the test statistic when the model parameters are freed to include redundant or poorly performing items. Each item was then reviewed by the statistician (L.B.) and flagged for potential deletion based on the distribution of responses, high item-level intercorrelations, discrimination and difficulty parameters, and MIs. This information was used by co-authors (L.B., M.C. and B.K.) to determine items to be retained or excluded in three separate rounds of review (Round 1 in

TABLE 1 Participant feedback and implemented changes.

Participant feedback	Implemented changes
Survey too long and repetitive	• The number of items was reduced. • To further reduce *cognitive load*, the language used in each of the questions was simplified and key words in each of the questions differentiated by using bold text.
Too much focus on university research evidence, at the expense of other important types of evidence (e.g., practice knowledge and internal research)	• 'Evidence-informed practice' was changed to 'evidence-driven action and learning', which shifted the focus to a *process* of planning, action and learning, using a range of different types of evidence.
Confusion about whether the 'importance' items related to the individual completing the questionnaire or their organisation	• Items were reworded to make it clear that these questions refer to the organisation, not the person completing the survey—this change provided clarity on the focus of the items.
Likert response scales—unclear and some confusion about how to use these	• The number of different response sets from six (i.e., 'in your organisation to what extent …', 'in your organisation to what extent do staff engaged in policy/program/service development generally have the skills to…') was reduced to only one 'Do you agree or disagree with the following statements?'—this change makes the questions easier to answer and also consistency means that survey participants should be able to answer the questions more quickly with less cognitive load.
Layout and ordering of sections/questions confusing	• The sections were reordered according to the main practices we are measuring (e.g., collaboration, adaptation, alignment and evidence-driven action and learning). • The items were ordered consistently within each section—this change made the questions easier to answer with less cognitive load.

June 2023, Rounds 2 and 3 in July 2023). Items identified for deletion in each round of review were dropped and serial FA on independent factors followed by CFA on the full model were performed. The resulting fit indices and loadings from the McDonald's Omega test of factor saturation were used to assess model fit and identify cross-loading items for further review by co-authors. The item set with the best-fit indices, and which aligned closest with the semantically and theoretically prescribed structure, was selected based on co-author consensus. Finally, analysis of variance was applied to test whether a correlated four-factor model or a second-order hierarchical four-factor model fit the data better.[83]

Analyses were all conducted in the R computing environment,[87] with the lavaan[88] and mirt[89] packages to support structural and item response analyses respectively. Significance testing was set for all analyses with α =.05.

2.3.3 | Sample description

In total, 126 participants provided useable data; the mean age was 43.7 years (SD 13.2); 54.8% identified as a man. Of these participants 7.1% reported being employed as CEOs, general managers, or other similar roles, 15.9% held senior management roles, 29.4% held other management roles, and 37.3% were other staff members.

2.3.4 | Scale structure and refinement

Histograms showed no item where mean responses clustered around the neutral mid-point or at either extreme of the Likert scale. Correlation matrices for each prescribed factor showed consistent positive item-to-item associations, with several pairs higher than $r = .7$ (see File S2). There was strong internal consistency across the items within the four factors (Adaptation: $\alpha = .91$, Alignment: $\alpha = .92$, Collaboration: $\alpha = .94$, Evidence-driven action and learning: $\alpha = .94$).

The results from all structural models are presented in Table 2. The fit indices from FA on individual factors (RMSEA range 1.24–2.03) and CFA on the full planned set of 60-items arranged in the four-factor structure (RMSEA = .098) all showed considerable room for improvement.

Results from IRT modelling are reported for item discrimination (a) and difficulty (b) together with item-factor loadings, skewness, kurtosis, and alpha test results in File S3. Together with correlations and MIs, these items were used to guide item retention and deletion in three rounds of co-author review. Items included in each round of testing are listed (drop/retain) in File S3.

Items with identical response patterns, high inter-item correlations or MIs, poor discrimination or ability scores, or with flat or disordered trace curves were reviewed in Round 1. Fourteen items were dropped from the original scale. Unidimensional FA at factor level was followed by CFA for the correlated 4-factor model using the remaining 46 items. As shown in Table 2, fit indices showed a minor improvement on the original scale (RMSEA = .096, TLI = .716). McDonald's Omega test of factor saturation in this 46-item set showed that many items cross-loaded onto more than one factor if they were not restrained as is done in CFA. These items were inspected by co-authors (Round 2) and three alternative approaches to item reduction were tested. In Round 2 Model 1, six items pertaining to leadership and culture were dropped, which improved the fit slightly from the planned model (RMSEA = .096, TLI = .734). In Round 2 Model 2, the six items were returned and a separate set of nine items identified as potentially redundant were dropped. This model was also better than the planned model (RMSEA = .090; TLI = .779). For Round 2 Model 3, the six items pertaining to leadership and culture and the nine items flagged as potentially redundant were dropped. This 31-item model returned the best fit of the series (RMSEA = .084, TLI = .818). The four factors remaining in this model were adaptation, alignment, collaboration, and evidence-driven action and learning. These factors were highly correlated, with an average factor correlation of .76. Adaptation and Alignment had the weakest correlation ($r(124) = .67$, $p = .05$) while Adaptation and Collaboration had the strongest positive correlation ($r(124) = .86$, $p = .05$). All polychoric correlations were significant with $\alpha = 0.05$ and the fit indices were close to the best practice cut points for 'good' model fit in psychometrics.[82,90]

As the planned factorial structure was theoretically salient, and the four factors were expected to be nested within the over-arching construct of organisational capacity for system practices (general factor), refer to Figure 1, we tested the fit for the remaining 31 items in a higher-order model.[83] This hierarchical, second-order model (Round 2 Model 3-H) had almost identical fit indices to the more parsimonious correlated factors model (RMSEA = .084, TLI = .819). When compared statistically, the fit for the hierarchical model was not worse than the correlated factors model ($\Delta\chi^2 = 1.758$, $p = .415$) and the information criteria were almost the same (hierarchical: AIC = 8267.6, BIC = 8454.8; correlated; AIC = 8269.6, BIC = 8454.8). We favoured the hierarchical, second-order four-factor model over the correlated factors model on theoretical grounds (Figure 2). Internal consistency for the 31-item hierarchical COSP scale was good at the whole scale ($\alpha = .95$) and factor level (adaptation $\alpha = .84$, alignment $\alpha = .85$, collaboration $\alpha = .91$, learning $\alpha = .89$). Item-factor loadings and IRT coefficients for each of the 31 items in the final model are presented in Table 3.

2.3.5 | Sensitivity analysis

Three more structural analyses were conducted to see if a better fit could be achieved, without losing the theoretical and semantic objectives of the measure. The 31-item set was subjected to the unrestrained Omega test and nine items still cross-loaded on more than one factor. Following co-author review, four items were identified for potential deletion (Round 3 Model 1). The initial CFA test for the 27-item set (Round 3 Model 1) returned promising fit indices (RMSEA = .071, TLI = .893). This item set was then tested for fit with the theoretically informed second-order hierarchical structure (Round

TABLE 2 Fit indices and refinements.

Structural analyses	No. items	c^2	c^2 df	c^2 p	RMSEA	CFI	TLI
Planned 4-factor model*	60	3771.95	1704	.000	.098	.665	.652
Round 1, Model 1*	45	1954.97	896	.000	.096	.731	.716
Round 2, Model 1*	40	1590.26	734	.000	.096	.750	.734
Round 2, Model 2*	37	1185.94	588	.000	.090	.794	.779
Round 2, Model 3*	31	808.46	428	.000	.084	.833	.818
*Round 2, Model 3-H**	*31*	*810.21*	*430*	*.000*	*.084*	*.833*	*.819*

Note: Fit indices from factor analyses in uni-dimensional data and confirmatory factor analyses (*) in multi-dimensional data with maximum likelihood operator and polychoric correlations. Round 2, Model 3-H: second-order model with a single common latent variable and four sub-scales, using the items retained in the final correlated factors model (Round 2, Model 3).
Abbreviations: CFI, Comparative Fit Index; RMSEA, Root Mean Squared Error of Approximation; TLI, Tucker Lewis Index.

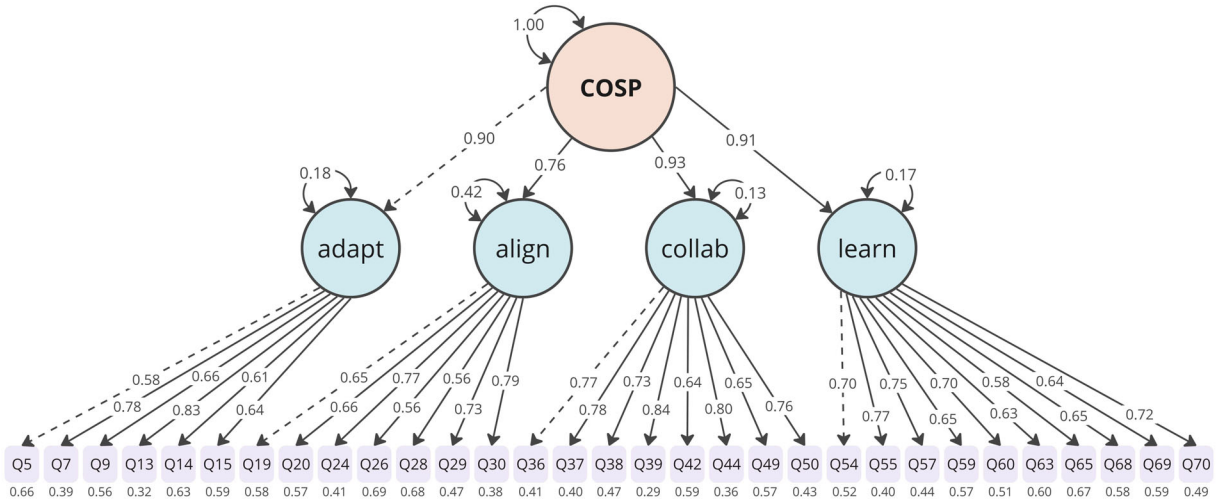

FIGURE 2 Final hierarchical second-order four factor model.

3 Model 1-H). Results showed a poorer fit than the correlated factors model (RMSEA = .082, TLI = .840). The test of difference between the two Round 3 models favoured the correlated factors model over the hierarchical one ($\Delta\chi^2 = 258.96$, $p < .001$). Despite having better psychometric properties, the Round 3 Model 1 structure was rejected in favour of the hierarchical structure supported in Round 2 Model 3, in which the superordinate general factor is more clearly comprised of four nested subordinate factors,[83] refer to Figure 1. Finally, a sensitivity analysis was conducted to test if a unifactorial structure would fit the data better now that the final item set had been determined. Results showed this was not the case (RMSEA = .168, TLI = .519) and the single-factor structure was rejected. The final scale items can be seen in Table 4.

2.3.6 | Scoring

The hierarchical second-order factor structure provides a defensible model in which the scores for the subordinate factors can be combined into an overall superordinate general factor score.[81,83] All the retained items have a positive orientation so there is no reverse-scoring requirement. The scale can be scored in two ways. First, by taking the mean score of responses across all participants to produce an overall score of organisational capacity for system practices (range 1–5) and sub-scale scores for capacity to adapt (range 1–5), capacity to align (range 1–5), capacity to collaborate (range 1–5) and capacity to engage in evidence-driven action and learning (range 1–5). Second, and recognising that additional testing will be needed, by taking the mean scores of responses from a single organisation, to produce a score of their capacity for system practices.

3 | DISCUSSION

Following a robust, three-stage development process, the Capacity of Organisations for System Practices (COSP) scale included 31 items to measure organisational capacity to adapt ($n = 6$), align ($n = 7$), collaborate ($n = 8$) and engage in evidence-driven action and learning ($n = 10$). Adaptation, alignment, collaboration and evidence-driven action and learning were conceptually and statistically different and

TABLE 3 Item characteristics for the final model.

Item no.	Factor loadings	a	b1	b2	b3	b4
Adaptation						
1	0.58	1.14	−3.50	−1.70	−0.97	1.87
2	0.79	2.08	−2.49	−1.08	−0.46	1.36
3	0.67	1.50	−3.31	−1.34	−0.59	1.29
4	0.83	3.09	−2.18	−1.22	−0.39	1.23
5	0.61	1.51	−1.97	−1.78	−0.82	1.51
6	0.64	1.30	−2.35	−1.98	−1.02	0.84
Alignment						
7	0.65	1.29	−3.87	−1.20	−1.05	2.06
8	0.66	1.10	−2.83	−1.30	−0.86	2.02
9	0.77	2.60	−3.13	−1.90	−0.63	0.91
10	0.56	0.77	−2.40	−2.10	0.27	2.06
11	0.56	1.16	−3.59	−2.01	−1.47	1.84
12	0.73	2.32	−3.17	−1.92	−0.79	1.18
13	0.79	2.27	−3.32	−1.26	−0.47	1.16
Collaboration						
14	0.77	2.16	−2.21	−0.90	−0.18	1.80
15	0.78	2.14	−2.32	−1.37	−0.36	1.25
16	0.73	1.79	−2.46	−1.30	−0.38	1.63
17	0.84	3.08	−2.25	−1.02	−0.15	1.36
18	0.64	1.31	−2.86	−1.94	−0.55	1.53
19	0.80	2.58	−2.06	−1.03	−0.60	1.40
20	0.65	1.42	−2.50	−2.41	−0.74	1.71
21	0.76	1.77	−2.87	−1.33	−0.63	1.17
Evidence-driven action and learning						
22	0.70	1.21	−2.34	−1.15	0.43	1.18
23	0.77	2.18	−2.02	−1.30	−0.58	1.20
24	0.74	1.82	−2.56	−1.34	−0.65	1.04
25	0.69	1.43	−2.19	−2.23	−0.88	1.31
26	0.69	1.77	−2.93	−1.56	−0.85	1.26
27	0.65	1.19	−2.07	−1.84	−0.45	1.52
28	0.58	1.15	−3.25	−2.03	−0.94	1.79
29	0.65	1.43	−3.07	−1.80	−0.77	1.73
30	0.64	1.27	−2.37	−2.09	−1.10	1.36
31	0.71	1.49	−2.07	−1.95	−0.60	0.67

Note: Item responsiveness tests conducted at factor (unidimensional scale) level using generalised partial credit model for ordinal, polytomous data. a: discrimination parameter showing item's ability to discriminate between different levels of the underlying trait relative to the threshold parameter. b1, b2, b3, b4: threshold parameters for the point on the scale of the latent trait with 0.5 positive response probability.

inter-correlated, as expected. The hierarchical structure supports the presence of an over-arching latent variable—organisational capacity for system practices—comprised of four factors: the capacity to adapt, align, collaborate and engage in evidence-driven action and learning. Although this suggests that each factor could be assessed separately, we believe the strength of the COSP scale is that it brings together the proposed practices necessary for systems change.

Adaptation, alignment, collaboration and engagement in evidence-driven action and learning were clearly interrelated, which supports the proposition of the initial Theory of Systems Change. Some of these relationships are supported by previous studies that have examined similar concepts. For example, collaboration was associated with the sustainment of evidence-based interventions,[28] adaptation[29] and alignment[30,31,91]; and adaptation was related to organisational learning.[32] The COSP scale data show that alignment

TABLE 4 Final capacity of organisations for system practices scale.

	Item	Dimension
	Adaptation	
1	My organisation prioritises adaptability when recruiting staff	Capacity (systems and structures)
2	Funding guidelines encourage my organisation to respond quickly to external opportunities and challenges	Capacity (enabling environment- funding)
3	My organisation has effective systems to respond to external opportunities and challenges	Capacity (systems and structures)
4	My organisation is good at responding to external opportunities and challenges	Functioning (Performance)
5	Staff in my organisation have the skills to respond to external opportunities and challenges	Capacity (staff capacity)
6	Leaders in my organisation value staff who can respond to external challenges and opportunities	Capacity (culture and leadership)
	Alignment	
7	My organisation has systems that allow us to identify young people's needs	Capacity (systems and structures)
8	Funding guidelines encourage my organisation to tailor our activities to the needs of young people	Capacity (enabling environment- funding)
9	My organisation is good at tailoring our activities to build on the strengths of young people	Functioning (Performance)
10	There is consistency between the vision of decision-makers, practitioners and young people in our community	Functioning (Performance)
11	Staff in my organisation have the skills to identify the needs of young people	Capacity (staff capacity)
12	Staff in my organisation have the skills to address the needs of young people	Capacity (staff capacity)
13	Leaders in my organisation prioritise activities that meet the needs of young people	Capacity (culture and leadership)
	Collaboration	
14	My organisation commits adequate budget and resources to collaborating with other organisations	Capacity (systems and structures)
15	My organisation is flexible and responsive to the requirements of other organisations	Functioning (Performance)
16	My organisation evaluates our collaborations	Capacity (systems and structures)
17	My organisation has systems that make it easy to collaborate with other organisations	Capacity (systems and structures)
18	Funding guidelines encourage us to collaborate with other organisations	Capacity (enabling environment- funding)
19	My organisation is good at collaborating with other organisations	Functioning (Performance)
20	Staff in my organisation are skilled at developing and maintaining relationships with other organisations	Capacity (staff capacity)
21	Leaders in my organisation prioritise collaboration with other organisations	Capacity (culture and leadership)
	Evidence-driven action and learning	
22	My organisation rewards risk taking and experimentation	Capacity (systems and structures)
23	My organisation has systems to support continuous cycles of planning, action and learning	Capacity (systems and structures)
24	My organisation uses the outcomes of our evaluations to inform future activities	Functioning (Performance)
25	In my organisation it is important to consider our activities as part of a broader system	Capacity (culture and leadership)
26	My organisation engages in continuous cycles of planning, action, and learning	Functioning (Performance)
27	Funding guidelines encourage my organisation to engage in a continuous process of planning, action and learning	Capacity (enabling environment- funding)
28	Staff in my organisation have the skills to monitor and evaluate our activities	Capacity (staff capacity)
29	Staff in my organisation have the skills to use diverse forms of evidence (e.g., professional experience and research) to inform our activities	Capacity (staff capacity)
30	Leaders in my organisation prioritise continuous improvement	Capacity (culture and leadership)
31	Leaders in my organisation encourage staff to learn from failures	Capacity (culture and leadership)

and evidence-driven action and learning are positively correlated. This makes theoretical sense because organisations that engage in cycles of action and learning are also likely to align their actions with the strengths and needs of their target population. To the best of our knowledge, there is no existing measure that can provide such valuable information about how system practices relate to each other.

Most other scales are developed to be sector-specific, which limits their ability to be applied across the range of organisations that

have a role in addressing the determinants of health. The COSP scale is unique in this regard, as it is designed for application across health-related and other organisations that provide services and programs to support different population groups. This makes it amenable to application in initiatives that bring together cross-sectoral organisations to address determinants of health and also across communities to examine the capacity of organisations within that initiative or community, respectively to engage in system practices. The COSP scale is also likely applicable across different target population groups. Although the items in the developed COSP scale focus on support for young people, the target population group of Pathways in Place-Victoria University, we anticipate that minor modifications, such as changing references to 'young people' in the items to better reflect the target group, can be made to the scale and it will retain its properties. However, the scale must be tested in other contexts to test this assumption.

Due to its flexibility and cross-sector application, the COSP scale could advance the evidence base for systems change and similar approaches in several ways. First, combined with other data, it could be administered across organisations to test the propositions of the initial Theory of Systems Change. For instance, to test the links between organisational capacity for system practices, systems change and population health outcomes or the proposition that embedding capacity for system practices within organisations has the potential to multiply health gains.[21,22,92,93] Second, it could be administered across organisations involved in systems change initiatives to assess the effectiveness of interventions that seek to develop organisational capacity for system practices. Finally, there is potential for the COSP scale to be administered *within* health-related and other types of organisations as a self-evaluation tool to assess their level of capacity for system practices.[63,94–97]

3.1 | Strengths and limitations

The strengths of this study are: (1) novelty, being the first attempt to develop a measure of organisational capacity for system practices; (2) a theory-driven approach, based on the initial Theory of Systems Change; (3) the use of an established and robust three-stage process for scale development and testing (based on Refs. 17,26); and (4) the use of established data analysis techniques to assess the structure of the measure and item qualities.

The relatively small sample size on which the analyses are based represents a limitation of this study. Further, while considered acceptable for this pragmatic measure, the fit indices did not meet best practice cut-offs therefore, further refinement and testing are needed. We deemed the hierarchical four-factor structure of the 31-item scale to be most acceptable. The sensitivity analysis showed it was feasible to achieve better psychometric fit if the theoretical model was not retained. However, this is a theory-driven instrument, and we opted to retain theoretical over psychometric dominance. Our sample of 126 complete cases was smaller than the recommended 10 participants per item, so we also suspect the small sample size used for structural analyses may be at least partly responsible for the suboptimal fit.[98] While we believe this measure will provide valuable insights into organisational capacity for system practices, we acknowledge that it does not capture the interrelationships between the components of a system that give rise to system behaviour. Finally, it is important to note that the Theory of Systems Change upon which the scale is based, is an initial middle-range theory that is yet to be fully tested. We anticipate that as the theory is refined, adjustments to the scale may be warranted.

3.2 | Future research

The COSP scale can now be refined and further tested. We suggest further development of items to measure evidence-driven action and learning. We believe adding items related to identifying underlying causes of problems and communicating and sharing knowledge across the system would strengthen the scale. Following modification, the scale should be tested for its construct validity, reliability, and utility. Following this testing, the COSP scale could then be further refined for use in other parts of the system (e.g., public policy).

3.3 | Conclusion

As public health continues to embrace systems change approaches, knowledge accumulates, new lines of enquiry emerge, and the development of novel measurement scales becomes crucial.[26] The nature of systems change requires scales that can cross traditional sector boundaries and be applied to health-related and other sectors. The COSP scale could provide valuable insights into organisational capacity for system practices. Although the scale shows promise, it requires further refinement and testing before it can be recommended for application. We believe that once validated, the COSP scale will ultimately help to advance systems change theory and practice.

AUTHOR CONTRIBUTIONS

M.C. and B.K. conceptualised the study. M.C., B.K., A.M., T.R., and M.K. contributed to item development and/or refinement. L.B. designed and conducted data analyses. M.C., A.M. and L.B. drafted the initial manuscript. T.R., B.K. and M.K. contributed to writing the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Mr Alex Molony, Ms Maria Cox, Ms Kate Garvey, Mx Georgia Gardner, Dr Enrique Garcia Bengoechea, Professor Dianne Finegood, Professor Lyndal Bond, Dr Vasileios Stavropoulos, Ms Michelle Morgan, Dr Liza Hopkins for participating in the initial pre-testing of the survey and providing valuable feedback that informed modifications for subsequent testing. We also thank Dr Annette Graham and Dr Jaimee Stuart for their contributions to this study. Open access publishing facilitated by Victoria University, as part of the Wiley -

Victoria University agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

This work was supported by the Paul Ramsay Foundation (grant number 681). Any opinions, findings, or conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the Foundation.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

This study was approved by the Victoria University Human Research Ethics Committee (HRE 21-163 and HRE22-021).

ORCID

Melinda Craike https://orcid.org/0000-0002-7374-1286
Larissa Bartlett https://orcid.org/0000-0003-0474-8107
Bojana Klepac https://orcid.org/0000-0002-9141-1862

REFERENCES

1. World Health Organization. Ottawa charter for health promotion: first international conference on health promotion Ottawa, 21 November 1986. Geneva: World Health Organization; 1986. https://www.euro.who.int/__data/assets/pdf_file/0004/129532/Ottawa_Charter.pdf
2. Allender S, Millar L, Hovmand P, Bell C, Moodie M, Carter R, et al. Whole of systems trial of prevention strategies for childhood obesity: WHO STOPS childhood obesity. Int J Environ Res Public Health. 2016;13(11):1143. https://doi.org/10.3390/ijerph13111143
3. Baugh Littlejohns L, Rasali D, McKee G, Naiman D, Faulkner G. Elusive boundaries: using an attribute framework to describe systems for population physical activity promotion. Health Promot Int. 2024;39:1–10. https://doi.org/10.1093/heapro/daae003
4. Bolton KA, Whelan J, Fraser P, Bell C, Allender S, Brown AD. The public health 12 framework: interpreting the 'meadows 12 places to act in a system' for use in public health. Arch Public Health. 2022;80(1):72. https://doi.org/10.1186/s13690-022-00835-0
5. Bransgrove NJ, Porter JE, Prokopiv V, Wildblood R, Prezioso M, Bradley V, et al. Local men local communities: a rural placed-based approach to increase men's engagement with mental health training and events. Health Promot J Austr. 2024;1–8. https://doi.org/10.1002/hpja.861
6. Koorts H, Rutter H. A systems approach to scale-up for population health improvement. Health Res Policy Syst. 2021;19(27):27. https://doi.org/10.1186/s12961-021-00679-0
7. McGill E, Er V, Penney T, Egan M, White M, Meier P, et al. Evaluation of public health interventions from a complex systems perspective: a research methods review. Soc Sci Med. 2021;272:113697. https://doi.org/10.1016/j.socscimed.2021.113697
8. Nau T, Bauman A, Smith BJ, Bellew W. A scoping review of systems approaches for increasing physical activity in populations. Health Res Policy Syst. 2022;20(104):104. https://doi.org/10.1186/s12961-022-00906-2
9. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, et al. The need for a complex systems model of evidence for public health. Lancet. 2017;390:2602–4. https://doi.org/10.1016/S0140-6736(17)31267-9
10. Foster-Fishman PG, Behrens TR. Systems change reborn: rethinking our theories, methods, and efforts in human services reform and community-based change. Am J Community Psychol. 2007;39(3-4):191–6.
11. Evans S, Raymond C, Perkins DD. Organizational and community capacity building. In: Scott V, Wolfe S, editors. Community psychology: foundations for practice. Los Angeles: SAGE Publications; 2014. p. 189–219.
12. Kegler MC, Norton BL, Aronson R. Achieving organizational change: findings from case studies of 20 California healthy cities and communities coalitions. Health Promot Int. 2008;23(2):109–18. https://doi.org/10.1093/heapro/dan009
13. Foster-Fishman PG, Long R. The challenges of place, capacity, and systems change: the story of yes we can! Foundation Rev. 2009;1(1):69–84. https://doi.org/10.4087/FOUNDATIONREVIEW-D-09-00005
14. Riley T, Hopkins L, Gomez M, Davidson S, Jacob J. System action learning: reorientating practice for system change in preventive health. Syst Pract Act Res. 2023;37(7):1–16. https://doi.org/10.1007/s11213-023-09638-y
15. Kasman M, Hammond RA, Heuberger B, Mack-Crane A, Purcell R, Economos C, et al. Activating a community: an agent-based model of Romp & Chomp, a whole-of-community childhood obesity intervention. Obesity. 2019;27(9):1494–502. https://doi.org/10.1002/oby.22553
16. Wilkins NJ, Kossover-Smith RA, Hogan S-A, Espinosa R, Wilson LF. Developing indicators to evaluate systems thinking and application in state injury and violence prevention programs. New Dir Eval. 2021;2021(170):67–80. https://doi.org/10.1002/ev.20456
17. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. Scale development: ten main limitations and recommendations to improve future research practices. Psicol: Reflex Crit. 2017;30(3):3. https://doi.org/10.1186/s41155-016-0057-1
18. Glasgow RE, Riley WT. Pragmatic eamsures: what they are and why we need them. Am J Prev Med. 2013;45(2):237–43. https://doi.org/10.1016/j.amepre.2013.03.010
19. Lewis CC, Mettert KD, Stanick CF, Halko HM, Nolen EA, Powell BJ, et al. The psychometric and pragmatic evidence rating scale (PAPERS) for measure development and evaluation. Implement Res Pract. 2021;2:1–6. https://doi.org/10.1177/26334895211037391
20. Craike M, Klepac B, Mowle A, Riley T. Theory of systems change: a middle-range theory of public health research impact. Res Eval. 2023;00:1–19. https://doi.org/10.1093/reseval/rvad030
21. Hanusaik N, Sabiston CM, Kishchuk N, Maximova K, O'Loughlin J. Association between organizational capacity and involvement in chronic disease prevention programming among Canadian public health organizations. Health Educ Res. 2015;30(2):206–22. https://doi.org/10.1093/her/cyu062
22. United Nations Development Group. UNDAF companion guide: capacity development. New York: United Nations; 2017. https://unsdg.un.org/sites/default/files/UNDG-UNDAF-Companion-Pieces-8-Capacity-Development.pdf
23. Preskill H, Torres RT. Building capacity for organizational learning through evaluative inquiry. Evaluation. 1999;5(1):42–60. https://doi.org/10.1177/135638909900500104
24. Thindwa J, Monico C, Reuben W. Enabling environments for civic engagement in PRSP countries (social development notes 82). Washington D.C: World Bank; 2003. https://openknowledge.worldbank.org/handle/10986/11319
25. Matheson A, Walton M, Gray R, Wehipeihana N, Wistow J. Strengthening prevention in communities through systems change: lessons

from the evaluation of healthy families NZ. Health Promot Int. 2020; 35(5):947–57. https://doi.org/10.1093/heapro/daz092
26. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health. 2018;6:149. https://doi.org/10.3389/fpubh.2018.00149
27. Contandriopoulos D, Hanusaik N, Maximova K, Paradis G, O'Loughlin JL. Mapping collaborative relations among Canada's chronic disease prevention organizations. Healthcare Policy. 2016; 12(1):101–15.
28. Green AE, Trott E, Willging CE, Finn NK, Ehrhart MG, Aarons GA. The role of collaborations in sustaining an evidence-based intervention to reduce child neglect. Child Abuse Negl. 2016;53:4–16. https://doi.org/10.1016/j.chiabu.2015.11.013
29. Loblay V, Garvey K, Shiell A, Kavanagh S, Hawe P. Can adaptation to 'extraordinary' times teach us about ways to strengthen community-based chronic disease prevention? Insights from the COVID-19 pandemic. Crit Public Health. 2022;32(1):127–38. https://doi.org/10.1080/09581596.2021.2006147
30. Nowell B. Out of sync and unaware? Exploring the effects of problem frame alignment and discordance in community collaboratives. J Publ Adm Res Theor. 2010;20:91–116. https://doi.org/10.1093/jopart/mup006
31. Timm I, Rapp S, Jeuter C, Bachert P, Reichert M, Woll A, et al. Interorganizational networks in physical activity promotion: a systematic review. Int J Environ Res Public Health. 2021;18(14):7306. https://doi.org/10.3390/ijerph18147306
32. van Herwerden LA, Reidlinger DP, Palermo C. The role of communication, building relationships, and adaptability in non-profit organisational capacity for health promotion. Health Promot Int. 2022;37(3): 1–12. https://doi.org/10.1093/heapro/daac074
33. Armstrong R, Waters E, Moore L, Dobbins M, Pettman T, Burns C, et al. Understanding evidence: a statewide survey to explore evidence-informed public health decision-making in a local government setting. Implement Sci. 2014;9(1):188. https://doi.org/10.1186/s13012-014-0188-7
34. Belita E, Squires JE, Yost J, Ganann R, Burnett T, Dobbins M. Measures of evidence-informed decision-making competence attributes: a psychometric systematic review. BMC Nurs. 2020;19(1):44. https://doi.org/10.1186/s12912-020-00436-8
35. Brennan SE, McKenzie JE, Turner T, Redman S, Makkar S, Williamson A, et al. Development and validation of SEER (seeking, engaging with and evaluating research): a measure of policymakers' capacity to engage with and use research. Health Res Policy Syst. 2017;15(1):1-19. https://doi.org/10.1186/s12961-016-0162-8
36. Brewster AL, Tan AX, Yuan CT. Development and application of a survey instrument to measure collaboration among health care and social services organizations. Health Serv Res. 2019;54(6):1246–54. https://doi.org/10.1111/1475-6773.13206
37. Canadian Foundation for Healthcare Improvement. Accelerating healthcare improvement—CFHI assessment tool. Ottawa, ON: Canadian Foundation for Healthcare Improvement; 2014. https://www.cfhi-fcass.ca/docs/default-source/itr/tools-and-resources/self-assessment-tool-2014-e.pdf
38. Charbonnier-Voirin A, El Akremi A, Vandenberghe C. A multilevel model of transformational leadership and adaptive performance and the moderating role of climate for innovation. Group Org Manag. 2010;35(6):699–726. https://doi.org/10.1177/1059601110390833
39. Charbonnier-Voirin A, Roussel P. Adaptive performance: a new scale to measure individual performance in organizations. Can J Admin Sci. 2012;29:280–93. https://doi.org/10.1002/cjas.232
40. Denison DR, Neale WS. Denison organizational culture survey: Facilitator guide. Ann Arbor: Denison Consulting LLC; 2018.
41. Dobbins M. Knowledge translation and exchange questionnaire. Ottawa, ON: Canadian Institute for Health Research; 2004.
42. Ehrhart MG, Aarons GA, Farahnak LR. Assessing the organizational context for EBP implementation: the development and validity testing of the implementation climate scale (ICS). Implement Sci. 2014; 9(157):157. https://doi.org/10.1186/s13012-014-0157-1
43. Engle NL. Adaptive capacity and its assessment. Glob Environ Change. 2011;21(2):647–56. https://doi.org/10.1016/j.gloenvcha.2011.01.019
44. Frey BB, Lohmeier JH, Lee SW, Tollefson N. Measuring collaboration among grant partners. Am J Eval. 2006;27(3):383–92. https://doi.org/10.1177/1098214006290356
45. Gedifew MT, Muluneh GS. Building a change adaptive university: a system of composite indicators to measure and facilitate pervasive changes. High Educ Q. 2021;75(2):227–46. https://doi.org/10.1111/hequ.12279
46. Gorley C. Adaptive capacity as a proactive approach [PhD Thesis]. Fielding Graduate University. 2012.
47. Greenbaum PE, Wang W, Boothroyd R, Kutash K, Friedman RM. Multilevel confirmatory factor analysis of the Systems of Care Implementation Survey (SOCIS). J Behav Health Serv Res. 2011;38(3):303–26. https://doi.org/10.1007/s11414-011-9240-4
48. Griffin MA, Parker SK, Mason CM. Leader vision and the development of adaptive and proactive performance: a longitudinal study. J Appl Psychol. 2010;95(1):174–82. https://doi.org/10.1037/a0017263
49. Held F, Hawe P, Roberts N, Conte K, Riley T. Core and peripheral organisations in prevention: insights from social network analysis. Health Promot J Austr. 2021;32(3):492–502. https://doi.org/10.1002/hpja.374
50. Hocevar SP, Jansen E, Thomas GF. Inter-organizational collaborative capacity (ICC) assessment. Monterey, CA: Calhoun; 2012. https://calhoun.nps.edu/handle/10945/38477
51. Hocevar SP, Thomas GF, Jansen E. Building collaborative capacity: an innovative strategy for homeland security preparedness. In: Beyerlein MM, Beyerlein ST, Kennedy FK, editors. Innovation through collaboration. Leeds: Emerald Group Publishing Limited; 2006. p. 255–74. https://doi.org/10.1016/S1572-0977(06)12010-5
52. Jacobs JA, Clayton PF, Dove C, Funchess T, Jones E, Perveen G, et al. A survey tool for measuring evidence-based decision making capacity in public health agencies. BMC Health Serv Res. 2012;12(1):57. https://doi.org/10.1186/1472-6963-12-57
53. Labin SN, Duffy JL, Meyers DC, Wandersman A, Lesesne CA. A research synthesis of the evaluation capacity building literature. Am J Eval. 2012;33(3):307–38. https://doi.org/10.1177/1098214011434608
54. Lockwood M, Raymond CM, Oczkowski E, Morrison M. Measuring the dimensions of adaptive capacity: a psychometric approach. Ecol Soc. 2015;20(1):37. https://doi.org/10.5751/ES-07203-200137
55. Makkar SR, Turner T, Williamson A, Louviere J, Redman S, Haynes A, et al. The development of ORACLe: a measure of an organisation's capacity to engage in evidence-informed health policy. Health Res Policy Syst. 2015;14(4):4. https://doi.org/10.1186/s12961-015-0069-9
56. Mazzucca S, Parks RG, Tabak RG, Allen P, Dobbins M, Stamatakis KA, et al. Assessing organizational supports for evidence-based decision making in local public health departments in the United States: development and psychometric properties of a new measure. J Public Health Manag Pract. 2019;25(5):454–63. https://doi.org/10.1097/PHH.0000000000000952
57. Melnyk BM, Fineout-Overholt E, Mays MZ. The evidence-based practice beliefs and implementation scales: psychometric properties of two new instruments. Worldviews Evid Based Nurs. 2008;5(4):208–16. https://doi.org/10.1111/j.1741-6787.2008.00126.x

58. Rice K, Hwang J, Abrefa-Gyan T, Powell K. Evidence-based practice questionnaire: a confirmatory factor analysis in a social work sample. Adv Soc Work. 2010;11(2):158–73. https://doi.org/10.18060/897
59. Ruzafa-Martinez M, Lopez-Iborra L, Moreno-Casbas T, Madrigal-Torres M. Development and validation of the competence in evidence based practice questionnaire (EBP-COQ) among nursing students. BMC Med Educ. 2013;13(19):1-10. https://doi.org/10.1186/1472-6920-13-19
60. Salehi V, Veitch B. Measuring and analyzing adaptive capacity at management levels of resilient systems. J Loss Prev Process Ind. 2020;63:104001. https://doi.org/10.1016/j.jlp.2019.104001
61. Schwarzman J, Bauman A, Gabbe BJ, Rissel C, Shilton T, Smith BJ. Understanding the factors that influence health promotion evaluation: the development and validation of the evaluation practice analysis survey. Eval Program Plann. 2019;74:76–83. https://doi.org/10.1016/j.evalprogplan.2019.03.002
62. Sussman C. Building adaptive capacity: the quest for improved organizational performance. Newton, MA: Sussman Associates; 2004. http://wikiciv.org.rs/images/5/5d/Sussman_(2004)_Building_Adaptive_Capacity.pdf
63. Taylor-Ritzler T, Suarez-Balcazar Y, Garcia-Iriarte E, Henry DB, Balcazar FE. Understanding and measuring evaluation capacity: a model and instrument validation study. Am J Eval. 2013;34(2):190–206. https://doi.org/10.1177/1098214012471421
64. Tikka T. The process of organisational adaptation through innovations, and organisational adaptability. Espoo: Helsinki University of Technology; 2010.
65. University of South Australia. Evidence-based practice profile questionnaire. n.d.. https://docplayer.net/22724820-Evidence-based-practice-profile-questionnaire.html
66. Upton D, Upton P. Development of an evidence-based practice questionnaire for nurses. J Adv Nurs. 2006;53(4):454–8. https://doi.org/10.1111/j.1365-2648.2006.03739.x
67. van Herwerden LA, Palermo C, Reidlinger DP. Capacity assessment in public health community interventions: a systematic review. Health Promot Int. 2019;34:84–93. https://doi.org/10.1093/heapro/day071
68. Yost J, Dobbins M, Traynor R, DeCorby K, Workentine S, Greco L. Tools to support evidence-informed public health decision making. BMC Public Health. 2014;14:728. https://doi.org/10.1186/1471-2458-14-728
69. Zaff JF, Pufall Jones E, Aasland K, Donlan AE, Lin ES, Prescott JE, et al. Alignment of perceived needs across levels of a community. J Appl Dev Psychol. 2015;40:8–16. https://doi.org/10.1016/j.appdev.2015.02.002
70. Edmondson A, Moingeon B. From organizational learning to the learning organization. Manag Learn. 1998;29(1):5–20. https://doi.org/10.1177/1350507698291001
71. Bryson JM, Crosby BC, Middleton Stone M. The design and implementation of cross-sector collaborations: propositions from the literature. Public Adm Rev. 2006;66:44–55. https://doi.org/10.1111/j.1540-6210.2006.00665.x
72. Foster-Fishman PG, Watson ER. The ABLe change framework: a conceptual and methodological tool for promoting systems change. Am J Community Psychol. 2012;49:503–16. https://doi.org/10.1007/s10464-011-9454-x
73. Kania J, Kramer M, Senge P. The water of systems change. Boston, MA: FSG; 2018. p. 1–20.
74. Bratianu C. Organizational learning and the learning organization. In: Bratianu C, editor. Organizational knowledge dynamics: Managing knowledge creation, acquisition, sharing, and transformation. Hershey, PA: IGI Global; 2015. https://doi.org/10.4018/978-1-4666-8318-1.ch012
75. Mayne J. Building evaluative culture in community services: caring for evidence. Eval Program Plann. 2020;80:101450. https://doi.org/10.1016/j.evalprogplan.2017.05.011
76. Mckenzie F. Building a culture of learning at scale learning networks for systems change. 2021.
77. Milèn A. What do we know about capacity building? An overview of existing knowledge and good practice. Geneva: World Health Organization; 2001. https://www.ircwash.org/sites/default/files/Milen-2001-What.pdf
78. DeCorby-Watson K, Mensah G, Bergeron K, Abdi S, Rempel B, Manson H. Effectiveness of capacity building interventions relevant to public health practice: a systematic review. BMC Public Health. 2018;18:684. https://doi.org/10.1186/s12889-018-5591-6
79. HealthWest Partnership. Evaluation capacity health check. Melbourne: HealthWest; 2020. https://healthwest.org.au/wp-content/uploads/2020/07/Evaluation-Capacity-Health-Check.pdf
80. Mowle A, Klepac B, Craike M, Riley T. Community asset mapping in Brimbank: findings snapshot. Melbourne: Victoria University; 2021. https://doi.org/10.26196/2653-DT12
81. Beaujean AA. Latent variable modeling using R: a step-by-step guide. 1st ed. New York and East Sussex: Routledge Taylor & Francis Group; 2014.
82. DeVellis RF. Scale development: theory and applications. 4th ed. Los Angeles: Sage Publications; 2017.
83. Markon KE. Bifactor and hierarchical models: specification, inference, and interpretation. Annu Rev Clin Psychol. 2019;15:51–69. https://doi.org/10.1146/annurev-clinpsy-050718-095522
84. Revelle W, French JA. The personality project Northwestern University: personality, motivation and cognition. Personality Project. 2013. https://www.personality-project.org/
85. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care. 2000;38(9):II28–42. https://doi.org/10.1097/00005650-200009002-00007
86. Muraki E, Muraki M. Generalized partial credit model. In: van der Linden WJ, editor. Handbook of Item Response Theory. 1st ed. Boca Raton, FL: CRC Press Taylor & Francis Group; 2016. p. 12.
87. R Core Team. R: a language and environment for statistical computing [computer software]. R foundation for Statistical Computing. 2022. https://www.R-project.org/
88. Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw. 2012;48(2):1–36. https://doi.org/10.18637/jss.v048.i02
89. Chalmers RP. Mirt: a multidimensional item response theory package for the R environment. J Stat Softw. 2012;48(6):1–29. https://doi.org/10.18637/jss.v048.i06
90. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J. 1999;6(1):1–55. https://doi.org/10.1080/10705519909540118
91. Bouwen R, Taillieu T. Multi-party collaboration as social learning for interdependence: developing relational knowing for sustainable natural resource management. J Community Appl Soc Psychol. 2004;14(3):137–53. https://doi.org/10.1002/casp.777
92. Brown AD, Whelan J, Bolton KA, Nagorcka-Smith P, Hayward J, Fraser P, et al. A theory of change for community-based systems interventions to prevent obesity. Am J Prev Med. 2022;62(5):786–94. https://doi.org/10.1016/j.amepre.2021.10.006
93. Hawe P, Noort M, King L, Jordens C. Multiplying health gains: the critical role of capacity-building within health promotion programs. Health Policy. 1997;39(1):29–42. https://doi.org/10.1016/S0168-8510(96)00847-0
94. Carey G, Malbon E, Carey N, Joyce A, Crammond B, Carey A. Systems science and systems thinking for public health: a systematic review of the field. BMJ Open. 2015;5:e009002. https://doi.org/10.1136/bmjopen-2015-009002

95. Coldwell M, Pearson S, Wilson I. Evaluating local system change using a comparative maturity matrix. Evaluation. 2022;28(4):446–65. https://doi.org/10.1177/13563890221126744
96. Gates EF, Fils-Aime F. System change evaluation: insights from the Rippel Foundation and its ReThink Health initiative. New Dir Eval. 2021;2021(170):125–38. https://doi.org/10.1002/ev.20462
97. Holmes BJ, Noel K. Time to shift from systems thinking-talking to systems thinking-action. Int J Health Policy Manag. 2015;4(4):245–7. https://doi.org/10.15171/ijhpm.2015.38
98. McNeish D, Wolf MG. Dynamic fit index cutoffs for confirmatory factor analysis models. Psychol Methods. 2023;28(1):61–88. https://doi.org/10.1037/met0000425

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Craike M, Bartlett L, Mowle A, Riley T, Krahe M, Klepac B. Development and testing of the capacity of organisations for system practices scale. Health Promot J Austral. 2024. https://doi.org/10.1002/hpja.922