



Swansea University
Prifysgol Abertawe



Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in:
Evaluation and Program Planning

Cronfa URL for this paper:
<http://cronfa.swan.ac.uk/Record/cronfa49960>

Paper:

Schwarzman, J., Bauman, A., Gabbe, B., Rissel, C., Shilton, T. & Smith, B. (2019). Understanding the factors that influence health promotion evaluation: The development and validation of the evaluation practice analysis survey. *Evaluation and Program Planning*, 74, 76-83.

<http://dx.doi.org/10.1016/j.evalprogplan.2019.03.002>

Released under the terms of a Creative Commons Attribution Non-Commercial No Derivatives License (CC-BY-NC-ND).

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

Title

Understanding the factors that influence health promotion evaluation: the development and validation of the Evaluation Practice Analysis Survey

Authors

Joanna Schwarzman ^a, Adrian Bauman ^b, Belinda J. Gabbe ^{a,c}, Chris Rissel ^b, Trevor Shilton ^d and Ben J. Smith ^{a,b}

Author affiliations and postal addresses

^a School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, VIC, 3004, Australia.

^b Prevention Research Collaboration, School of Public Health, The University of Sydney, Sydney, NSW 2006, Australia.

^c Health Data Research UK, Swansea University Medical School, Swansea University, Singleton Park, Swansea, SA2 8PP, Wales UK

^d National Heart Foundation of Australia, 334 Rokeby Road, Subiaco, WA 6008, Australia.

Corresponding author

Joanna Schwarzman

School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, VIC, 3004, Australia.

Ph. +61 3 9903 0241

joanna.schwarzman@monash.edu

Funding

This work was supported by the Australian Research Council (ARC) [DP150103575]. Prof. Gabbe was supported by an ARC Future Fellowship [FT170100048] during the preparation of this manuscript.

Declarations of interest: None

Title: Understanding the factors that influence health promotion evaluation: the development and validation of the Evaluation Practice Analysis Survey

Highlights

- Measurement tools to assess the evaluation capacity or practice in health promotion organisations are lacking.
- The newly developed Evaluation Practice Analysis Survey (EPAS) is a reliable and valid instrument to assess elements of evaluation practice in health promotion organisations.
- The EPAS comprises 25 scales addressing individual, organisational and system level factors that influence health promotion evaluation.
- The instrument demonstrated good to excellent internal reliability of 23 scales, and predictive validity for a number of organisational and resource variables.
- The EPAS has potential application in the planning and evaluation of capacity building initiatives in health promotion organisations.

Abstract

The demand for improved quality of health promotion evaluation and greater capacity to undertake evaluation is growing, yet evidence of the challenges and facilitators to evaluation practice within the health promotion field is lacking. A limited number of evaluation capacity measurement instruments have been validated in government or non-government organisations (NGO), however there is no instrument designed for health promotion organisations. This study aimed to develop and validate an Evaluation Practice Analysis Survey (EPAS) to examine evaluation practices in health promotion organisations. Qualitative interviews, existing frameworks and instruments informed the survey development. Health promotion practitioners from government agencies and NGOs completed the survey (n=169). Principal components analysis was used to determine scale structure and Cronbach's α used to estimate internal reliability. Logistic regression was conducted to assess predictive validity of selected EPAS scale. The final survey instrument included 25 scales (125 items). The EPAS demonstrated good internal reliability ($\alpha > 0.7$) for 23 scales. Dedicated resources and time for evaluation, leadership, organisational culture and internal support for evaluation showed promising predictive validity. The EPAS can be used to describe elements of evaluation capacity at the individual, organisational and system levels and to guide initiatives to improve evaluation practice in health promotion organisations.

Keywords

Health promotion; evaluation practice; evaluation capacity building; measurement; reliability; validity

Introduction

There is a misalignment between the demand for rigorous evaluation of health promotion programs, and the documented quality of evaluation practice in this field (Author, et al., 2016; Chambers, Murphy, & Kolbe, 2015; Lim, Wright, Carrotte, & Pedrana, 2016; Oxman et al., 2010; Pettman et al., 2012). Health promotion programs typically aim to address socio-ecological determinants of health using multiple and varied strategies; from environmental change, to social marketing or education. In many cases health promotion agencies will work with diverse populations, on a variety of health and social issues, in a range of contexts. In Australia, these agencies are commonly responsible for planning, delivery (frequently through partner organisations) and evaluation of programs.

Consequently, program planning, implementation and evaluation are identified core competencies for health promotion practitioners (International Union for Health Promotion and Education, 2016). Practitioners and organisations often recognise the importance of evaluation, and initiatives to build evaluation capacity in health promotion organisations have been reported (Edwards, Stickney, Milat, Campbell, & Thackway, 2016; Lindeman et al., 2018; Lobo et al., 2016; Nichols, McFarlane, Gibson, Millard, Packer, & McDonald, 2018; Valenti, Campetti, Schoenborn, Quinlan, & Dash, 2017). The interest in evaluation capacity within the field is encouraging, however there is a need for robust evidence to guide these efforts.

As the evaluation and ECB field progresses from theoretical frameworks to empirically supported models of ECB, commentators are calling for the development and validation of measurement instruments that can guide the planning of comprehensive ECB strategies and enable their evaluation (Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012; Norton, Milat, Edwards, & Giffin, 2016). To date a limited number of survey instruments have been developed to assess evaluation capacity and practice within organisations. In Canada two instruments, based on a framework of capacity to undertake and use evaluation, were developed to assess evaluation capacity. The Evaluation Capacity in Organisations Questionnaire (ECOQ) was developed for use by internal evaluators within Canadian government and NGOs (Cousins et al., 2008). Nine scales measured

capacity for organisational learning, support structures, capacity to do evaluation, ECB activities, use of findings and process use, and stakeholder participation. Eight of the scales demonstrated good internal reliability ($\alpha > 0.8$) (Cousins et al., 2008) and there was found to be good fit of the nine scales to the conceptual framework (Gagnon, Aubry, Cousins, Goh, & Elliott, 2018). Another measure, the Organisational Evaluation Capacity Self-Assessment Instrument, explored six domains of evaluation, including human and organisational resources, evaluation planning and activities, evaluation literacy, organisational decision making and learning benefits (Bourgeois, Simmons, Hotte, & Osseni, 2016). The instrument demonstrated good face validity as assessed by expert review and pilot testing, and was implemented with predominantly government agencies.

The Evaluation Capacity Index (ECI) was designed to map local government evaluation capacity in Denmark across evaluation demand factors (objectives, structure and processes), and supply factors (technology, human capacity) (Nielsen, Lemire, & Skov, 2011). Construct validity of the ECI was assessed using confirmatory factor analysis and found to support the theoretical model. Taylor-Ritzler et al. (2013) developed the Evaluation Capacity Assessment Index (ECAI) which also demonstrated good internal reliability of all ten scales ($\alpha > 0.8$). The study found construct validity of the ECAI through assessment of the relationship between awareness, motivation, competence constructs and a second order latent variable of individual factors, and between leadership, learning climate and resources and a latent variable of organisational factors (Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry, & Balcazar, 2013). The measurement properties of additional evaluation practice instruments have not yet been assessed (Carman, 2007, 2008; Carman & Fredericks, 2009; Fierro & Christie, 2017).

In the health promotion field efforts to conduct and assess the impact of ECB strategies are emerging. In HIV prevention organisations a 7-item knowledge and confidence scale was adapted from existing instruments (Taylor-Ritzler et al., 2013) completed by participants of an ECB initiative ($n=33$) (Lindeman et al., 2018). The authors reported a high internal consistency of the adapted scale

in both the 'retrospective' ($\alpha=0.88$) and the 'current' ($\alpha=0.93$) application of the survey. Another case study used documentary review and qualitative interviews to report the impact of, and challenges to implementing an evaluation technical assistance and coaching initiative in substance abuse prevention organisations (Valenti et al., 2017). A further study, in an Aboriginal Community Controlled primary health care organisation, evaluated the effect of organisational systems changes, a workshop series and mentoring using pre-post workshops surveys to assess changes in skills and confidence among health promotion practitioners (Nichols et al., 2018). While these studies demonstrate motivation to conduct and evaluate initiatives to improve evaluation capacity in the health promotion field, they were conducted with small samples, and gave limited attention to the psychometric properties of survey instruments used. These studies also highlight the need for an instrument that can be used to describe and measure aspects of evaluation capacity beyond individual skills and knowledge.

Most studies of evaluation practices in the health promotion field have used qualitative approaches to explore barriers and facilitators to evaluation in homogenous organisation types (Author et al., 2015; Jolley, Lawless, Baum, Hurley, & Fry, 2007; Napp, Gibbs, Jolly, Westover, & Uhl, 2002; van Koperen et al., 2016). We recently conducted a qualitative study that explored the perceptions of 40 managers, from a range of disease and injury fields, jurisdictions and organisation types, of the key factors affecting health promotion evaluation. This revealed the powerful influence of organisational factors (e.g. leadership, culture, resource allocation, support) upon evaluation practice and use, together with staff-level (e.g. skills, attitudes) and system level (e.g. funding body requirements, political factors) determinants (Author et al., 2018). Further studies in health promotion organisations identified human resources, time and the availability of evaluation tools or templates as necessary to evaluation (Author et al., 2015; van Koperen et al., 2016).

In the wider field of research concerning evaluation in human services and other organisations, several studies and frameworks position the organisation and organisational learning as central to

evaluation capacity (Bourgeois & Cousins, 2013; Cousins et al., 2008; Nielsen et al., 2011; Preskill & Boyle, 2008). Resources, leadership, organisational culture and systems or processes that support evaluation within the organisation have been consistently described as pivotal to evaluation practice, use and/or capacity in a range of organisational settings (Author et al., 2018; Bourgeois & Cousins, 2013; Carman, 2007; Cousins et al., 2008; Fierro & Christie, 2017; Labin, 2014; Taylor-Ritzler et al., 2013). Internal and external support for evaluation has been reported to have a mediating effect on evaluation capacity building (ECB) (Labin, 2014), particularly support and expertise accessed through an internal evaluation team, external partnerships, or engagement with university or consultant evaluators (Author et al., 2018).

The organisation has been shown to play an important role in developing staff level capacities such as evaluation knowledge, skills, attitudes and beliefs, through leadership and provision of a supportive environment (Author et al., 2018; Labin, 2014; Labin et al., 2012; Taylor-Ritzler et al., 2013). These individual factors have been identified as important barriers to, and facilitators of, evaluation in health promotion organisations (Author et al., 2015; Huckel Schneider, Milat, & Moore, 2016; Lobo, McManus, Brown, Hildebrand, & Maycock, 2010). The relationship between individual and organisational level evaluation capacity has been postulated in theoretical frameworks (Preskill & Boyle, 2008), and described in a range of models underpinned by empirical research (Labin, 2014; Labin et al., 2012; Taylor-Ritzler et al., 2013).

Evaluation capacity has also been portrayed as the relationship between evaluation supply and demand in an organisation (Nielsen et al., 2011). Several qualitative studies exploring barriers and facilitators to evaluation practice have identified health policy, funding and administrative requirements, and expectations that can act as powerful drivers or hindrances to evaluation in health promotion agencies (Author et al., 2015; Lobo et al., 2010; Lobo, Petrich, & Burns, 2014; Napp et al., 2002).

The findings of research to date show that influences on health promotion evaluation practice act at the individual-, organisational- and wider prevention system- levels. Outside the health promotion field several attempts to measure evaluation practice and capacity have been made, however these instruments focus predominantly on a select number of organisational and individual factors. Given the distinctive administrative, organisational and practice characteristics of health promotion, there is a need for purposively developed instruments for measuring evaluation capacity and practice in this field. Additionally, the validity and reliability of existing instruments have not been assessed in health promotion agencies, where evaluation is often the responsibility of program staff, as opposed to specialist evaluators surveyed in previous studies (Cousins et al., 2008; Nielsen et al., 2011). In light of these limitations, the purpose of this study was to develop and evaluate the psychometric properties of a survey instrument to measure the factors that influence health promotion and disease prevention evaluation practice.

Methods

Survey development

The Evaluation Practice Analysis Survey (EPAS) was developed in multiple stages to enhance face and content validity (Figure 1). The conceptual framework that informed the content and structure of the survey was based on a synthesis of the literature presented above. Particular weight was given to evidence generated from within the health promotion field. Factors that influence evaluation practice in health promotion were grouped into three domains (Figure 2). Survey items were created for each major domain (individual, organisation and system), and grouped into sub-headings. Items from existing evaluation practice and capacity instruments were reviewed for suitability to the health promotion context.

Figure 2. Proposed framework of the factors that influence evaluation practice in health promotion.

Individual level	Organisational level	System level
Skills for evaluation Attitudes and beliefs	Organisational culture Leadership for evaluation Systems and structures Evaluation team or role Support for evaluation Partnerships External evaluators University partners Resources	Funding for evaluation Reporting requirements Funding body priorities and requirements Political and prevention context

A panel of four experts in health promotion and disease prevention evaluation and research reviewed the items for face validity, clarity of wording, and provided feedback on suggested Likert scales. The revised instrument was used in face-to-face cognitive interviews with 10 health promotion practitioners in Victoria, Australia. Further revisions were made for clarity of definitions, item wording and response options. The survey was transferred into an online platform and sent to 10 eligible practitioners outside of Victoria who agreed to provide feedback after completion of the survey. Six participants provided feedback and minor revisions were made to instructions to participants and response options.

The preliminary survey included 158 items exploring factors influencing evaluation practice in 15 groups, with all but two sections being rated on 6 point Likert scales from strongly agree to strongly disagree (Table 1). Minor adaptations were made to 17 items from the ECAI 'Awareness' and 'Competence' scales (Taylor-Ritzler et al., 2013) and included in the preliminary survey. A "don't know" response option was added to 59 items based on feedback during expert review and cognitive interviews (Table 1). Respondent and organisational characteristics and evaluation practice measures were also collected.

Table 1. Items and response scales in preliminary Evaluation Practice Analysis Survey

Initial proposed scales	N items	Response options
Self-reported evaluation practice	6	Continuous scale 0-100 ^a
<i>Individual level factors</i>		
Skills – evaluation steps	13 ^b	6 point beginner-expert scale
Skills – levels of evaluation	5	6 point beginner-expert scale
Attitudes and beliefs	13 ^c	6 point agreement scale
<i>Organisational level factors</i>		
Organisational culture	12	6 point agreement scale
Leadership	9	6 point agreement scale
Systems and structures	10	6 point agreement scale*
Evaluation team or role [#]	7	6 point agreement scale
Support	9	6 point agreement scale
Partnerships	8	6 point agreement scale
External evaluators and University partners [#]	7	6 point agreement scale
Resources	20	6 point agreement scale*
<i>System level factors</i>		
Funding for evaluation	10	6 point agreement scale*
Reporting requirements [#]	10	6 point agreement scale*
Funding body priorities and requirements	11	6 point agreement scale*
Political and prevention context	14	6 point agreement scale*

^a Presented on sliding scale with increments of 5, with allowance for free entry of number 0 – 100.

^b 6 items from the Evaluation Capacity Assessment Instrument (Taylor-Ritzler et al., 2013)

^c 11 items from the Evaluation Capacity Assessment Instrument (Taylor-Ritzler et al., 2013)

* Included "don't know" response option.

[#] Screening question to skip item set if not applicable e.g. Does your organisation have a designated evaluation team or role?

Participant recruitment

Ethics approval was received from the Monash University Human Research Ethics Committee.

Practitioner and health promotion organisation email addresses were compiled through internet searching, health promotion practitioner networks and the 2015 Australian Health Promotion

Association (AHPA) Conference delegate list. The AHPA also forwarded an invitation to members in

the Australian Capital Territory, Northern Territory, Queensland and Tasmania. In total, an invitation

was sent by email to 337 practitioners and organisational administrative contacts. To encourage

snowball recruitment, invited participants were encouraged to forward the email to eligible

colleagues. This survey development study was undertaken within the context of a larger research

project in which health promotion organisations had previously been recruited from New South Wales, South Australia, Victoria and Western Australia. The organisations recruited to the larger research project were considered ineligible to participate in this current study as they would be invited to complete the revised and validated EPAS at a later date. To avoid contamination of the participant groups, respondents to the initial invitation were screened based on their state and if necessary, their organisation. All participants were provided with an information sheet and commencement of the online survey was accepted as consent. Participants were offered an AUD\$ 20 gift voucher on completion of the survey.

Preliminary analysis

Survey data were imported into IBM SPSS 23 for analysis. All system level scales, and five items from the preliminary 'Organisational resources' scale had "don't know" options. When "don't know" was present for 25% or fewer items within a scale this was replaced by the mean of scores from the other items in the scale to retain cases for analysis (Tabachnick & Fidell, 2013). All other "don't know" and not applicable responses were managed as pairwise missing values in principal components analysis (PCA).

Spearman's correlation coefficients were generated for each preliminary scale to examine correlations between scale items and to guide removal of items with very low or very high correlations (Field, 2013). Where items demonstrated poor correlation ($R < 0.25$ for more than 50% of the scale items), these were deleted from the preliminary item pool. Preliminary scales were checked for multicollinearity, with a matrix determinant of < 0.000001 indicating multicollinearity (Field, 2013). Scales that were found not to meet the multicollinearity criteria were reviewed for highly correlating pairs and if required, an item was deleted prior to analysis.

Principal components analysis and scale internal reliability

We selected PCA as a technique that allowed us to understand the structure of each aspect of evaluation practice, extract the important components and eliminate redundant items using

principal components analysis (PCA) (Field, 2013). This approach to assessing scale structure was used as the survey items were developed on the basis theoretical frameworks and empirical evidence that identify levels and dimensions of evaluation capacity in the health promotion field. As this survey measured a wide range of influences upon evaluation practice using a large number of items relative to the sample size, each preliminary construct was analysed using separate PCA. Prior to PCA it was determined whether each scale met the required assumptions; Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) (>0.5), Bartlett's test of sphericity ($p < 0.001$), and absence of multicollinearity based on the determinant matrix. PCA was conducted with oblique (oblimin in SPSS) rotation as the components were assumed to be correlated (Field, 2013) and initial PCA set the minimum eigenvalue of 1.0 for component extraction. A minimum component loading of 0.5 was selected as this threshold is statistically significant in samples of over 100 participants (Stevens, 2002, in Field, 2013). A higher loading of 0.7 was adopted for the 'Evaluation role' scale which had a substantially lower sample size ($n=49$) due to a prior screening question (Stevens, 2002, in Field, 2013). Extracted components with fewer than three items were not considered suitable for final scale output. Component extraction criteria for PCA were used to confirm the final scales, including Scree plots, eigenvalues, total variance accounted for, as well as appropriate theoretical fit of items within a scale (Field, 2013).

For each final scale, the internal reliability was determined using Cronbach's α . A threshold of $\alpha > 0.6$ was considered acceptable for this study, particularly as the number of items in a scale influence α and many final EPAS scales consisted of fewer than seven items (Streiner & Norman, 2008).

Predictive validity

Predictive validity of the EPAS was assessed using logistic regression. The dependent variable of self-reported evaluation practice was developed using survey items that asked participants to estimate the percentage of projects evaluated at formative, process, impact and outcome levels in the past two years. Principal components analysis was undertaken using these four items. All required

assumptions for PCA were met and a single evaluation practice scale was extracted with item loadings >0.7 and $\alpha=0.79$.

The 'Self-reported evaluation practice' scale was then dichotomised at a threshold of at least 60% of programs evaluated (average of formative, process, impact and outcome evaluation). This cut point closely aligned with the highest tertile of scale distribution ($T_2=61.2$, range 0-100) (Table 3).

Sensitivity analysis was undertaken with cut points for the evaluation practice scale set at higher and lower levels (50%, 70%), and this had negligible impact upon the predictive validity findings.

Based on prior qualitative studies (Author et al., 2018), nine EPAS scales that were considered likely to influence evaluation practice were selected as independent variables for analysis. Each of the nine EPAS scales was categorised based on tertiles into high, medium or low EPAS scale scores, and was entered into a logistic regression model with self-reported evaluation practice as the outcome to establish the predictive capacity of the scales. The tertile distribution of each scale is shown in Table 3. Variables associated with both self-reported evaluation practice and each EPAS variable at the $p<0.1$ level were considered potential confounders, and included selectively in the respective logistic regression models. Odds ratios (OR) and 95% confidence intervals (CI) were reported.

Results

Participant characteristics

The final sample included the responses of 169 participants, of which 159 were complete for all items. A response rate could not be calculated due to potential overlap of participants from each recruitment channel and the use of central organisational email contacts. Participant organisation and individual characteristics are shown in Table 2.

Table 2. Organisational, program and participant characteristics

	Non-Government		Government		Total	
	n	(%)	n	(%)	n	(%)
Total	83	(49)	86	(51)	169	(100)
<i>Location</i>						
Metropolitan area	34	(41)	48	(56)	82	(48)
Non-metropolitan area	49	(59)	38	(44)	87	(51)
<i>Annual health promotion budget (Australian Dollars)</i>						
Less than \$499 999	49	(59)	42	(49)	91	(54)
\$500 000 or more	18	(22)	28	(33)	46	(27)
Don't know	16	(19)	16	(19)	32	(19)
<i>Staffing levels for health promotion (full time equivalent [FTE])</i>						
Less than 3 FTE	40	(48)	37	(43)	77	(46)
3 or more FTE	43	(52)	49	(57)	92	(54)
<i>Organisation's experience in health promotion (years)</i>						
10 years or less	27	(33)	15	(17)	42	(25)
More than 10 years	55	(66)	60	(70)	115	(68)
Don't know	1	(1)	11	(13)	12	(7)
<i>Participant characteristics</i>						
<i>Current role</i> n=159						
Manager	33	(40)	15	(17)	48	(28)
Non-manager	45	(54)	66	(77)	111	(66)
<i>Experience in health promotion</i> n=169						
Less than 10 years	45	(54)	47	(55)	92	(54)
10 or more years	38	(46)	39	(45)	77	(46)
<i>Qualifications</i> n=159						
Undergraduate or post graduate public health or health promotion	48	(62)	57	(70)	105	(66)
Other qualifications	25	(32)	22	(27)	47	(30)
None	5	(6)	2	(2)	7	(4)

Scale structure and internal reliability.

Sixteen items were deleted from ten preliminary scales prior to PCA based on the correlation analysis. A further 17 items with low loadings were excluded in the PCA process. Eight of the preliminary scales retained all items and were unchanged after PCA.

The PCA resulted in 25 EPAS scales, and retained 125 items addressing individual (four scales), organisational (13 scales) and system level factors (eight scales) (Table 3). While the majority of scales were found to be unidimensional, five preliminary scales were found to be comprised of two dimensions ('Organisational culture', 'Support', 'Attitudes' and 'External evaluation and university

partnerships') or three dimensions ('Resources'). The item list for each final scale is provided in Appendix A. Most EPAS scales demonstrated good to excellent internal reliability, with $\alpha > 0.7$ for 23 of the 25 scales. The remaining two scales ('Control and flexibility' and 'Political decision making') showing acceptable internal reliability ($\alpha > 0.6$). The properties of the dependent variable 'Self-reported evaluation practice' are also shown in Table 3.

Table 3. Descriptive and internal reliability statistics for Evaluation Practice Analysis Survey scales

Scale	Sample (n)	Items (n)	Scale range	Cronbach's α	Mean	(SD)	Ter1, Ter2
<i>Individual level</i>							
Attitudes and beliefs – program level	169	5	5 - 30	0.83	27.8	(2.5)	27, 30
Attitudes and beliefs – wider benefits	169	5	5 - 30	0.83	25.9	(3.3)	24, 28
Skills – levels of evaluation	169	5	5 - 30	0.90	18.2	(5.4)	16, 21
Skills – evaluation tasks	169	12	12 - 72	0.95	51.3	(12.4)	47, 57.3
<i>Organisational level</i>							
Leadership	168	9	9 - 54	0.96	36.5	(9.8)	34, 41
Organisational culture for the right environment	168	6	6 - 36	0.87	25.3	(5.5)	24, 27
Organisational culture for systems	168	6	6 - 36	0.91	24.5	(6.7)	21, 28
Systems and structures	167	7	7 - 42	0.85	26.2	(6.9)	23, 30
Internal support	164	4	4 - 24	0.88	15.4	(4.2)	14, 17
External support	164	3	3 - 18	0.77	11.4	(3.2)	10, 13
Tools	161	5	5 - 30	0.86	19.5	(4.8)	18, 22
Time	161	4	4 - 24	0.87	15.3	(3.8)	14, 18
Resources dedicated to evaluation	144	4	4 - 24	0.85	12.9	(4.8)	10, 15
Evaluation role	49	4	4 - 24	0.86	18.8	(3.7)	18, 21.3
Partnerships	162	6	6 - 36	0.86	27.3	(4.8)	26, 30
University evaluators	94	3	3 - 18	0.75	12.6	(2.7)	11, 14
External evaluators	95	3	3 - 18	0.73	14.4	(2.5)	14, 15
<i>System level</i>							
Adequate funding for evaluation	141	5	5 - 30	0.94	16.7	(6.5)	15, 20
Clear and realistic expectations	96	4	4 - 24	0.78	15.8	(4.0)	15, 17.7
Control and flexibility	128	3	3 - 18	0.66	11.5	(2.9)	10, 13
Competing demands	109	5	5 - 30	0.77	15.2	(4.7)	12, 17
Organisational influence	96	4	4 - 24	0.74	13.9	(4.0)	12.3, 15
Learning and sharing	112	4	4 - 24	0.76	14.5	(3.7)	13, 16
Political decision making	116	5	5 - 30	0.63	24.0	(3.2)	23, 25
Reporting facilitates evaluation	94	4	4 - 24	0.78	27.8	(5.2)	26, 30
<i>Evaluation practice</i>							
Self-reported evaluation practice	169	4	0-100	0.79	46.6	(27.7)	25, 61.2

SD = standard deviation, Ter1 = 1st tertile, Ter2 = 2nd tertile.

Predictive validity of key EPAS variables on evaluation practice

Chi-square analysis identified that annual health promotion budget and/or the respondent's role within their organisation were potential confounders of the relationship between four of the EPAS scales and the dependent evaluation practice variable. These potential confounders were included as covariates in logistic regression analysis for the four identified EPAS scales (Table 4). The remaining five EPAS scales were analysed using univariate logistic regression. The logistic regression results found seven of the nine of selected EPAS scales to be significantly associated with evaluation practice (Table 4).

Table 4. Odds ratios for selected EPAS scales to predict program evaluation conducted for more than 60% of projects.

EPAS scale		OR (95% CI)	p value
Adequate funding for evaluation	Low	Ref	
	Medium	1.43 (0.52-3.90)	0.49
	High	5.19 (2.10-12.80)	0.00
Clear and realistic expectations	Low	Ref	
	Medium	1.80 (0.64-5.09)	0.27
	High	2.77 (0.98-7.85)	0.06
Resources dedicated to evaluation ^a	Low	Ref	
	Medium	4.73 (1.58-14.19)	0.01
	High	4.76 (1.62-14.00)	0.01
Time	Low	Ref	
	Medium	2.26 (0.98-5.22)	0.06
	High	3.14 (1.39-7.12)	0.01
Leadership	Low	Ref	
	Medium	2.12 (0.94-4.80)	0.07
	High	2.69 (1.19-6.10)	0.02
Organisational culture for the right environment	Low	Ref	
	Medium	0.81 (0.36-1.82)	0.61
	High	2.26 (1.03-4.93)	0.04
Organisational culture for systems ^b	Low	Ref	
	Medium	1.92 (0.80-4.60)	0.14
	High	3.75 (1.61-8.71)	0.00
Internal support ^b	Low	Ref	
	Medium	1.66 (0.71-3.89)	0.24
	High	2.89 (1.22-6.83)	0.02

Systems and structures ^{a, b}	Low	Ref	
	Medium	1.51 (0.55-4.13)	0.42
	High	1.57 (0.62-4.03)	0.34

OR = odds ratio, CI = confidence interval, Ref = Reference category

^a Adjusted for *Annual health promotion budget*

^b Adjusted for *Current role*

High levels of resources for evaluation, including adequate funding, dedicated resources and time for evaluation were associated with between three and fivefold greater odds of evaluating of more than 60% of projects (Table 4). Reporting high levels of leadership, organisational culture to create a supportive environment, organisational culture that facilitates systems and internal support were also associated with at least two times greater odds of evaluating at least 60% of projects. Variables 'Clear and realistic expectations', and 'Systems and structures' were not found to be significantly associated with evaluation practice at the $p < 0.05$ level.

Discussion

The availability of reliable and valid measures to describe evaluation practice is important for the initiation, design and evaluation of ECB strategies (Labin et al., 2012; Norton et al, 2016). This current study describes the development of the first instrument designed to measure evaluation practice in health promotion agencies. In particular, the EPAS uniquely contributes a wider range of reliable and valid scales that measure the individual, organisational and system-level dimensions of health promotion evaluation capacity than previous instruments. Principal components analysis was used to identify the dimensionality of the EPAS scales and reduce the number of survey items, resulting in an instrument that demonstrated good to excellent internal reliability for all but two of the 25 scales. Further, seven of the nine key EPAS scales showed promising predictive validity of evaluation practice in health promotion. This study is the first to our knowledge to assess the predictive validity of evaluation capacity measures upon reported evaluation practice.

To progress understanding of the influences on evaluation practice, this study builds on qualitative research exploring evaluation capacity in health promotion, theoretical frameworks of evaluation capacity and previous attempts at evaluation capacity measurement. We developed several scales that are consistent with instruments and frameworks published to date, especially within the individual and organisational domains. To our knowledge the ECAI is the only other instrument to assess the measurement properties of individual level factors influencing evaluation practice, including awareness, motivation and competence, and we adapted items from the 'Awareness' and 'Competence' scales for the EPAS (Taylor-Ritzler et al., 2013). In contrast with Taylor-Ritzler's study, we found that the individual level factors of attitudes and beliefs about evaluation had distinct dimensions, each demonstrating high levels of internal reliability ($\alpha > 0.8$). The first dimension described perceived benefits at the program level, the second addressed perceived organisational and system wide benefits of evaluation.

Organisational level factors contributed the highest number of scales to the EPAS, and the majority of these had high levels of internal reliability (10 of 13 scales $\alpha > 0.8$). This was unsurprising given the range of existing instruments and theories that position the organisation as central to evaluation capacity. There are some similarities between the EPAS and the three organisational constructs measured in the ECAI, namely 'Resources', 'Leadership' and 'Learning climate' (Taylor-Ritzler et al., 2013). While we also developed scales for resources, PCA identified three resource sub-scales that delineated the elements of time, evaluation tools, and dedicated resources for evaluation within an organisation. Another distinction between the ECAI and the EPAS is the nature of the leadership constructs. The ECAI 'Leadership' scale (5 items, $\alpha = 0.82$) addresses the evaluation leadership characteristics of program managers only (Taylor-Ritzler et al., 2013), whereas the EPAS 'Leadership' scale encompasses more diverse sources of evaluation leadership. This interpretation of leadership reflects a breadth of examples of leadership for evaluation (e.g., champion evaluation, research and evaluation support role, facilitate partnerships) as identified in qualitative research in health promotion organisations (Author et al., 2018).

The identification of suitable groupings of system level items for PCA was more challenging compared to organisational and individual factors. While most system level scales still demonstrated good internal reliability, the scales for 'Control and flexibility' and 'Political decision making' reached only acceptable levels of internal reliability ($\alpha > 0.6$). Beyond identification of funding body requirements as influential on health promotion evaluation (Author et al., 2015; Lobo 2010), evidence of how the health promotion system influences evaluation has not been extensively explored. Therefore, the constructs relating to system level influences on evaluation practice may evolve with further qualitative and quantitative exploration of the health promotion system.

In our analysis of the predictive validity of selected EPAS scales we found that 'Dedicated resources' and 'Time' for evaluation were positively associated with conducting program evaluation. Resources are frequently described as essential to evaluation practice, particularly for health promotion practitioners who have heavy workloads and are expected to be specialists across many areas of program management, including evaluation (Author et al., 2015; Author et al., 2018; Jolley et al., 2007; Lobo et al., 2010).

Leadership and organisational culture were also associated with evaluation practice in this study. Qualitative research has identified the pivotal role of leadership in creating a culture of learning, establishing and maintaining systems for communication and advocating for resource allocation to evaluation in health promotion (Author et al., 2018). Two dimensions of organisational culture for which we developed reliable and valid scales, namely 'Organisational culture to create the right environment' and 'Organisational culture for systems and structures', have been found to be important in previous studies concerning evaluation capacity. An organisational culture that is supportive of learning, innovation and evaluation has been described as a key facilitator of ECB (Preskill & Boyle, 2008; Taylor-Ritzler et al., 2013), as has an organisational culture that develops systems and processes for shared learning and reporting to actively encourage evaluation (Bourgeois & Cousins, 2013). Embedding evaluation firmly within organisational culture and systems should also

be an important goal of ECB, especially to promote use of evaluation (Cousins, Goh, Elliott, Aubry, & Gilbert, 2014).

To date ECB initiatives have tended to focus on training or technical support for the development of knowledge and skills, and the provision of resources in the form of evaluation guides and tools (Cousins et al., 2014; Norton et al., 2016). The EPAS shows promising utility for planning more comprehensive ECB initiatives in health promotion organisations that address skills, organisational resources, time and tools, together with organisational culture and support, and leadership for evaluation. Additionally, the EPAS may be useful in furthering research and evaluation of the impact of ECB strategies, which has been identified as an important next step in ECB research (Norton et al., 2016).

We also acknowledge limitations to this study. The sample size for this study may be considered low, given PCA is generally a larger sample technique (Tabachnick & Fidell, 2013). However in samples of less than 300 participants a combination of loadings, communalities, items, and factors should be considered when interpreting the components extracted (Field, 2013; Guadagnoli & Velicer, 1988; MacCallum, Widaman, Preacher & Hong, 2001; Tabachnick & Fidell, 2013). In our study, the scale 'Political decision making' should be interpreted cautiously as it did not meet the additional criteria to justify a sample of less than 300 based on item communalities of 0.5 (Guadagnoli & Velicer, 1988) and at least four items with loadings >0.6 (MacCallum, Widaman, Zhang & Hong, 1999), despite meeting the KMO measure of sampling adequacy. Due to the sample size of the study several of the logistic regression models returned wide confidence intervals in the predictive validity analysis. We considered reducing categories within each independent variable to increase cell size, however decided that the tertile categories used more of the information available with only a modest loss of precision. Despite sample size limitations, we believe we were able to capture a representative sample of eligible health promotion practitioners through a comprehensive sampling strategy and follow-up emails and phone calls.

The use of self-report to measure evaluation practice for the predictive validity analysis may also be considered a limitation, however, given the lack of alternatives, we considered it an adequate proxy for this study and the scale to measure this dependent variable showed good internal reliability.

Establishing valid and robust measures of evaluation practice would be beneficial for further validation of the EPAS. Additionally, while we explored the relationship of selected constructs with evaluation practice, there is also a hypothesised relationship between the EPAS scales that would benefit from further testing to help understand how each factor interacts to influence evaluation practice.

Conclusions and lessons learned

The EPAS, based on theoretical and empirical foundations, is a comprehensive measurement tool examining individual, organisational and system level aspects of health promotion evaluation capacity. The EPAS demonstrated good internal reliability and validity to predict evaluation practice. It also shows promising utility as a tool for planning more ECB initiatives in health promotion that include skill building, development of organisational resources, tools, organisational culture and leadership development in evaluation. There is potential for practitioners and organisations to use the EPAS to deepen understanding of the influences on evaluation practice and ultimately develop and evaluate effective strategies to guide ECB in health promotion organisations.

Acknowledgements

[Removed for blind review]

Funding

[Removed for blind review]

References

Author et al. (2015). [Removed for blind review]

Author, et al. (2016). [Removed for blind review]

Author, et al. (2018). [Removed for blind review]

Bourgeois, I., & Cousins, J. B. (2013). Understanding Dimensions of Organizational Evaluation

Capacity. *American Journal of Evaluation*, 34(3), 299-319. doi:10.1177/1098214013477235

Bourgeois, I., Simmons, L., Hotte, N., & Osseni, R. (2016). Measuring Evaluation Capacity in Ontario

Public Health Units. *The Canadian Journal of Program Evaluation*, 31(2), 1.

Carman, J. G. (2007). Evaluation practice among community-based organizations - Research into the

reality. *American Journal of Evaluation*, 28(1), 60-75. doi:10.1177/1098214006296245

Carman, J. G. (2008). Nonprofits, Funders, and Evaluation: Accountability in Action. *The American*

Review of Public Administration, 39(4), 374-390. doi:10.1177/0275074008320190

Carman, J. G., & Fredericks, K. A. (2009). Evaluation Capacity and Nonprofit Organizations: Is the

Glass Half-Empty or Half-Full? *American Journal of Evaluation*, 31(1), 84-104.

doi:10.1177/1098214009352361

Chambers, A. H., Murphy, K., & Kolbe, A. (2015). Designs and methods used in published Australian

health promotion evaluations 1992-2011. *Australian and New Zealand Journal of Public*

Health, 39(3), 222-226. doi:10.1111/1753-6405.12359

Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Chouinard, J., Goh, S. C., et al. (2008). Organizational

Capacity to Do and Use Evaluation: Results of a Pan-Canadian Survey of Evaluators. *The*

Canadian Journal of Program Evaluation, 23(3), 1-35.

Cousins, J. B., Goh, S. C., Elliott, C., Aubry, T., & Gilbert, N. (2014). Government and voluntary sector

differences in organizational capacity to do and use evaluation. *Evaluation and Program*

Planning, 44, 1-13.

- Edwards, B., Stickney, B., Milat, A., Campbell, D., & Thackway, S. (2016). Building research and evaluation capacity in population health: the NSW Health approach. *Health Promotion Journal of Australia*, 27(3), 264-267. doi:10.1071/HE16045
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics: (and sex and drugs and rock 'n' roll)* (4th ed. ed.). Los Angeles:Sage.
- Fierro, L. A., & Christie, C. A. (2017). Evaluator and program manager perceptions of evaluation capacity and evaluation practice. *American Journal of Evaluation*, 38(3), 376-392.
- Gagnon, F., Aubry, T., Cousins, J. B., Goh, S. C., & Elliott, C. (2018). Validation of the evaluation capacity in organizations questionnaire. *Evaluation and Program Planning*, 68, 166-175. doi:<https://doi.org/10.1016/j.evalprogplan.2018.01.002>
- Guadagnoli, E., & Velicer, W. (1988). Relation of Sample Size to the Stability of Component Patterns. *Psychological Bulletin*, 103(2), 265. doi:10.1037/0033-2909.103.2.265
- Huckel Schneider, C., Milat, A. J., & Moore, G. (2016). Barriers and facilitators to evaluation of health policies and programs: Policymaker and researcher perspectives. *Evaluation and Program Planning*, 58, 208-215. doi:10.1016/j.evalprogplan.2016.06.011
- International Union for Health Promotion and Education (2016). *Core Competencies and Professional Standards for Health Promotion*. Accessed online 17th January, 2018. https://www.iuhpe.org/images/JC-Accreditation/Core_Competencies_Standards_linkE.pdf
- Jolley, G. M., Lawless, A. P., Baum, F. E., Hurley, C. J., & Fry, D. (2007). Building an evidence base for community health: a review of the quality of program evaluations. *Australian Health Review*, 31(4), 603-610.
- Labin, S. N. (2014). Developing Common Measures in Evaluation Capacity Building: An Iterative Science and Practice Process. *American Journal of Evaluation*, 35(1), 107-115. doi:10.1177/1098214013499965

- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A Research Synthesis of the Evaluation Capacity Building Literature. *American Journal of Evaluation*, 33(3), 307-338. doi:10.1177/1098214011434608
- Lim, M. S. C., Wright, C. J. C., Carrotte, E. R., & Pedrana, A. E. (2016). Reach, engagement, and effectiveness: a systematic review of evaluation methodologies used in health promotion via social networking sites. *Health Promotion Journal of Australia*, 27(3), 187-197. doi:10.1071/HE16057
- Lindeman, P. T., Bettin, E., Beach, L. B., Adames, C. N., Johnson, A. K., et al. (2018). Evaluation capacity building—Results and reflections across two years of a multisite empowerment evaluation in an HIV prevention context. *Evaluation and Program Planning*, 71, 83-88. doi:https://doi.org/10.1016/j.evalprogplan.2018.09.001
- Lobo, R., Crawford, G., Hallett, J., Laing, S., Mak, D. B., Jancey, J., et al. (2016). A research and evaluation capacity building model in Western Australia. *Health Promotion International*. Online publication, 1-11. doi:10.1093/heapro/daw088
- Lobo, R., McManus, A., Brown, G., Hildebrand, J., & Maycock, B. (2010). Evaluating peer-based youth programs: barriers and enablers. *Evaluation Journal of Australasia*, 10(1), 36-43.
- Lobo, R., Petrich, M., & Burns, S. K. (2014). Supporting health promotion practitioners to undertake evaluation for program development. *BMC Public Health*, 15(1315), 1-8. doi:10.1186/1471-2458-14-1315
- MacCallum, R., Widaman, K., Zhang, S., & Hong, S. (1999). Sample Size in Factor Analysis. *Psychological Methods [PsycARTICLES]*, 4(1), 84. doi:10.1037/1082-989X.4.1.84
- MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample Size in Factor Analysis: The Role of Model Error. *Multivariate Behavioral Research*, 36(4), 611-637. doi:10.1207/S15327906MBR3604_06

- Napp, D., Gibbs, D., Jolly, D., Westover, B., & Uhl, G. (2002). Evaluation barriers and facilitators among community-based HIV prevention programs. *AIDS Education and Prevention*, 14(3 Suppl A), 38-48.
- Nichols, N., McFarlane, K., Gibson, P., Millard, F., Packer, A., & McDonald, M. (2018). Skills, systems and supports: An Aboriginal Community Controlled Health Service (Apunipima) approach to building health promotion evaluation capacity of staff. *Health Promotion Journal of Australia*, 29(1), 18-22. doi:doi:10.1002/hpja.18
- Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring Evaluation Capacity—Results and Implications of a Danish Study. *American Journal of Evaluation*, 32(3), 324-344. doi:10.1177/1098214010396075
- Norton, S., Milat, A., Edwards, B., & Giffin, M. (2016). Narrative review of strategies by organizations for building evaluation capacity. *Evaluation and Program Planning*, 58, 1-19. doi:10.1016/j.evalprogplan.2016.04.004
- Oxman, A. D., Bjørndal, A., Becerra-Posada, F., Gibson, M., Block, M. A. G., Haines, A., et al. (2010). A framework for mandatory impact evaluation to ensure well informed public policy decisions. *The Lancet*, 375(9712), 427-431. doi:10.1016/S0140-6736(09)61251-4
- Pettman, T. L., Armstrong, R., Doyle, J., Burford, B., Anderson, L. M., Hillgrove, T., et al. (2012). Strengthening evaluation to capture the breadth of public health practice: ideal vs. real. *J Public Health (Oxf)*, 34(1), 151-155. doi:10.1093/pubmed/fds014
- Preskill, H., & Boyle, S. (2008). A Multidisciplinary Model of Evaluation Capacity Building. *American Journal of Evaluation*, 29(4), 443-459. doi:10.1177/1098214008324182
- Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales a practical guide to their development and use* (4th ed.). New York: Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.

Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D. B., & Balcazar, F. E. (2013).

Understanding and Measuring Evaluation Capacity: A Model and Instrument Validation Study. *American Journal of Evaluation*, 34(2), 190-206. doi:10.1177/1098214012471421

Valenti, M., Campetti, R., Schoenborn, N., Quinlan, K., & Dash, K. (2017). Building evaluation capacity of local substance abuse prevention programs serving LGBTQ populations. *Evaluation and Program Planning*, 63, 101.

van Koperen, T. M., Renders, C. M., Spierings, E. J. M., Hendriks, A.-M., Westerman, M. J., Seidell, J. C., et al. (2016). Recommendations and Improvements for the Evaluation of Integrated Community-Wide Interventions Approaches. *Journal of Obesity*, 2016, 1-13.
doi:10.1155/2016/2385698

Author biographies

Joanna Schwarzman is a final year PhD candidate at the School of Public Health and Preventive Medicine at Monash University, and holds a MPH from Flinders University and Bachelor Physiotherapy from University of Melbourne. Joanna's career interests focus on the use of evidence in health promotion and disease prevention and building workforce and organisational capacity for evaluation in order to more effectively address population health challenges.

Adrian Bauman is Professor of Public Health and co-directs the WHO Collaborating Centre on Physical Activity, Nutrition and Obesity. He has a medical degree, MPH and PhD from Sydney University. His research is in chronic disease prevention and public health research translation and in the evaluation methods for complex public health programs. His expertise includes physical activity research, epidemiological studies, and the evaluation of population interventions and behaviour change programs. He is well published in research, and was on the 2015, 2016 and 2017 Thompson-Reuter (Clarivate) lists of the most cited 1% of researchers in any discipline.

Professor Belinda Gabbe is the Head of the Prehospital, Emergency and Trauma Research Unit in the Department of Epidemiology and Preventive Medicine at Monash University. She is an Australian Research Council Future Fellow, and holds a Professorial appointment in the Health Data Research UK Institute at Swansea University (UK). Prof Gabbe holds an Honours degree in Physiotherapy, a PhD in Epidemiology and a Master of Biostatistics.

Chris Rissel is Director of the NSW Office of Preventive Health and Professor of Public Health with the School of Public Health, The University of Sydney. He has a BSc (with Honours) in Psychology from UNSW, a Master of Public Health degree from Sydney University, and a PhD in Epidemiology (with a focus on community health promotion) from the University of Minnesota (USA). He works at the interface between research, policy and program implementation and is interested in the role of evidence in advocacy and how to best scale-up interventions for maximum health outcomes.

Trevor Shilton is National Director, Active Living with the National Heart Foundation of Australia. He is Vice President with the International Union for Health Promotion and Education (IUHPE) and has played a prominent role in developing national and global health promotion competencies and in advancing the profession of health promotion. He was the Founding National President of the Australian Health Promotion Association (AHPA). He has a Masters Degree in Health Promotion and a Bachelor of Education. His research, policy and practice interests are in chronic disease prevention, health promotion, advocacy and physical activity.

Ben Smith is an Associate Professor in the School of Public Health at the University of Sydney, and holds a BSW (Hons) degree (University of New South Wales), along with an MPH and PhD from the University of Sydney. Prior to his academic career he worked in community development and program management roles addressing health issues in high needs urban populations. His research interests concern chronic disease prevention, physical activity promotion, and capacity building for health promotion. During his career he has worked with a wide range of government and non-government agencies in the design and conduct of project evaluations.

Appendix A. Final Evaluation Practice Analysis Survey (EPAS): scales and survey items

Survey domains Scale names	Cronbach's α	Component loading
Evaluation practice		
Self-reported evaluation practice	0.79	
a) Percentage of projects evaluated in the past two years at the formative level		0.74
b) Percentage of projects evaluated in the past two years at the process level		0.79
c) Percentage of projects evaluated in the past two years at the impact level		0.84
d) Percentage of projects evaluated in the past two years at the outcome level		0.75
Individual level		
Attitude and beliefs - Program and staff level benefits <i>I think that evaluation...</i>	0.83	
a) Will help me understand my program		0.83
b) Will inform the decisions I make about my program		0.91
c) Is absolutely necessary to improve my program		0.56
d) Should involve program participants in the evaluation process		0.73
e) Is worth spending time on		0.67
Attitude and beliefs - Organisation and wider benefits <i>I think that evaluation...</i>	0.83	
a) Will influence policy relevant to my program (e.g. stakeholder/partnership, local, state or national policy)		0.87
b) Will help improve programs for people from minority and hard to reach populations		0.78
c) Will justify funding for my program		0.50
d) Will help to convince managers that changes are needed in my program		0.75
e) Will inform changes in our internal processes and documentation		0.66
Skills for different levels of evaluation	0.90	
a) Formative evaluation (i.e. pretesting or piloting)		0.87
b) Process evaluation (i.e. monitoring of implementation)		0.90
c) Impact (i.e. achievement of objectives)		0.94
d) Outcome evaluation (i.e. achievement of project goals)		0.93
e) Economic evaluation (eg. cost benefit analysis)		0.60
Skills for different evaluation tasks	0.95	
a) Clearly stating goals and objectives for my project		0.74
b) Deciding what questions to answer in an evaluation		0.86
c) Choosing designs to evaluate project impacts or outcomes		0.86
d) Developing data collection tools		0.85
e) Identifying from whom to collect the information (i.e. a sampling strategy)		0.83
f) Implementing or appraising ethical standards in evaluation		0.68
g) Analysing quantitative evaluation data		0.79
h) Analysing qualitative evaluation data		0.79
i) Developing recommendations based on evaluation results		0.82
j) Writing an evaluation report		0.85
k) Disseminating findings to relevant stakeholders		0.83
l) Managing an evaluation		0.89
Organisational level		
Leadership for evaluation <i>The leaders in my organisation...</i>	0.96	
a) Have a clear vision for evaluation		0.91
b) Demonstrate a good understanding of program evaluation		0.92
c) Are able to build good partnerships with those who can support evaluation when required		0.87
d) Can successfully negotiate evaluation and reporting requirements with our funding source		0.83
e) Ensure our evaluation findings are useful to decision makers		0.90
f) Champion evaluation practice and use internally		0.93
g) Have a strong desire to share our findings		0.86
h) Lack interest in evaluation		0.78
i) Advocate in support of our evaluation efforts when required		0.85

Understanding the factors that influence health promotion evaluation

Organisational culture to create the right environment	0.87	
<i>My organisation...</i>		
a) Encourages a learning environment		0.77
b) Is an innovative organisation		0.76
c) Uses evidence from evaluations to inform decision making		0.69
d) Shares evaluation findings with other organisations		0.74
e) Has a system that allows us to learn successful practices from other organisations		0.65
f) Is worried about evaluation findings making the organisation look bad		0.69
Organisational culture for systems and structures	0.91	
a) Evaluation is an important part of our organisational culture		0.81
b) Evaluation of our programs is part of our core business		0.92
c) Everyone in the organisation has some responsibility to contribute to evaluation		0.83
<i>My organisation...</i>		
d) Embeds evaluation in everything we do		0.67
e) Uses evaluation findings to inform the next phase of work		0.58
f) Actively provides staff with opportunities to upskill in the area of evaluation		0.54
Systems and structures	0.85	
a) It is clear from my job description what is expected of me in terms of evaluation		0.74
b) We have an evaluation framework in place for the health promotion work that we do		0.82
c) We have a user friendly way to store and manage our program data (electronic, or paper based)		0.63
d) It is easy to access evaluation expertise from within the organisation		0.75
e) There is a process for seeking advice from external evaluation experts if needed		0.71
f) We have regular meetings to facilitate shared understanding regarding the evaluation plan		0.77
g) Our organisation has clear expectations and timeframes for reporting		0.68
Internal support	0.88	
a) There is adequate support within my organisation for the evaluation I'm involved in		0.90
b) My organisation arranges structured support and training for evaluation as required		0.82
c) Management is happy to support evaluation and allows time and resources for it		0.94
d) We have a strong working relationship with people who can support our team in evaluation		0.71
External support	0.77	
a) I am comfortable contacting people outside of my organisation for evaluation support and advice		0.94
b) Formal arrangements with other organisations support our evaluation practice		0.66
c) I get evaluation support from a network of practitioners I am involved with		0.74
Tools for evaluation	0.86	
a) We have access to data collection instruments appropriate to our work		0.75
b) We have a functioning database to collate information		0.87
c) We have appropriate tools for evaluation involving hard to reach and vulnerable groups		0.82
d) Our organisation provides us with the basic resources to support evaluation (e.g., computers, software, copying, administrative support)		0.78
e) We can access a range of good templates for reporting on our evaluation		0.70
Time for evaluation	0.86	
a) We have time to engage with our community and stakeholders for evaluation		0.60
b) We have adequate time for evaluation planning		0.73
c) We have enough time to think and reflect on our work		0.97
d) We have time to write up our findings to share them		0.92
Resources dedicated to evaluation	0.85	
a) Our organisation has long-term, dedicated financial support for our evaluation activities		0.81
b) Evaluation funding is set aside as part of every program budget		0.93
c) Even when there are budget constraints we still allocate funds to evaluation		0.91
d) Staff are encouraged to spend time on program evaluation		0.51

Understanding the factors that influence health promotion evaluation

Evaluation role in the organisation	0.86	
<i>The evaluation person or team...</i>		
a) Champions evaluation of health promotion projects in our organisation.		0.82
b) Supports the health promotion program staff to undertake elements of the evaluation		0.89
c) Mentors staff members in evaluation practice		0.75
d) Is involved in helping to plan and design our evaluations right from the beginning		0.76
Partnerships	0.86	
a) Partnerships allow us to attract additional resources for evaluation		0.77
b) Involving our partners in evaluation provides additional skills for evaluation		0.81
c) Our partners help to facilitate evaluation involving hard to reach population groups		0.83
d) Our partnerships are valuable for obtaining data needed in our evaluation		0.81
e) There are clear expectations about evaluation between our partners		0.62
f) Involving our partners in evaluation of the programs is worth the effort		0.74
External Evaluators	0.73	
<i>We engage external evaluators...</i>		
a) To ensure the findings are independent		0.69
b) To ensure a high quality of evaluation		0.80
c) When we don't have the required skills internally		0.87
University partnerships	0.75	
a) We work with university partners to help obtain funds for evaluation		0.81
b) Our university partners understand how we work		0.84
c) We have the opportunity to build evaluation capacity through our contacts with university partners		0.72
System level		
Adequate funding for evaluation	0.94	
<i>We have sufficient funding for...</i>		
a) Health promotion evaluation (overall)		0.91
b) Formative evaluation (i.e. pretesting and piloting)		0.88
c) Process evaluation (i.e. monitoring of implementation)		0.90
d) Impact and/or outcome evaluation (i.e. achievement of objectives)		0.92
e) Outcome evaluation (i.e. achievement of project goals)		0.89
Clear and realistic expectations	0.78	
a) There are clear evaluation requirements from our funding body		0.78
b) There are clear expectations from our funding body about the level of investment required in evaluation		0.61
c) The evaluation activities I engage in are consistent with funders' expectations		0.76
d) The funding body has realistic expectations for evaluation given available resources		0.90
Control and flexibility	0.66	
a) We have good flexibility from our funding body to evaluate the way we think is most appropriate		0.75
b) We have a high level of control over the indicators chosen to evaluate and report on		0.80
c) There is sufficient time available to develop an evaluation plan that meets stakeholder requirements		0.77
Competing demands	0.77	
a) It is challenging to meet funding body requirements for evaluation due to the complex nature of our programs		0.72
b) There is pressure to deliver the programs at the expense of quality evaluation		0.80
c) Evaluation is cut or reduced when the budget is tight overall		0.71
d) Our funding body is generally more concerned with health service delivery than health promotion or prevention		0.62
e) There is tension between what is important for us to evaluate locally and the funding body's priorities for evaluation		0.75
Organisational influence in the system	0.74	
a) We've built political support for our health promotion or disease prevention work to continue and grow		0.68

Understanding the factors that influence health promotion evaluation

b) We go to extra lengths to secure funding for evaluation beyond the usual program funding	0.80
c) We distil our evaluation findings into key messages that help us engage politicians	0.66
d) We have had success in securing additional funds for program evaluation	0.83
Learning and sharing	0.76
a) Our funding body communicates clearly about how they will use our evaluation	0.77
b) Our funding body encourages us to share our evaluation findings	0.77
c) Our funding body provides leadership for conducting evaluation	0.78
d) The funding body supports learning from evaluation findings	0.73
Political decision making	0.63
a) Changes in funding priorities have impacted on the evaluation of our programs	0.53
b) Despite having evaluation findings available, decisions appear to be made based on political factors	0.58
c) Despite demonstrating effectiveness, projects have not been refunded	0.51
d) It's more important for us to evaluate using rigorous methods in the current political climate	0.64
e) Evaluation findings that include an economic component are used more often by decision makers	0.79
Reporting facilitates evaluation	0.78
a) I am confident to use the reporting templates that we have been provided with	0.74
b) Evaluation is expected as part of the funding arrangements	0.83
c) Our reporting requirements and key performance indicators drive our evaluation work	0.78
d) Our funding body provides a reporting template	0.76

Figure 1. Evaluation Practice Analysis Survey: Stages of development

