



The evidence needed to demonstrate impact: A synthesis of the evidence from a phased social and emotional wellbeing intervention



Leigh-ann Onnis^{a,*}, Helen Klieve^b, Komla Tsey^c

^a Indigenous Education and Research Centre, James Cook University, PO Box 6811, Cairns, QLD 4870, Australia

^b School of Education and Professional Studies, Griffith University, Australia

^c College of Arts, Society & Education, James Cook University, Australia

ARTICLE INFO

Keywords:

Family wellbeing
Aboriginal health
Social and emotional wellbeing
Phased interventions

ABSTRACT

Policy decisions are based on evidence that demonstrates the effectiveness of interventions; however, the quantity and type of evidence that is needed to demonstrate the effectiveness of an intervention is not universally agreed upon. The aim of this study was to collaborate with researchers who have not been involved directly in Family Wellbeing interventions to lead a review of characteristics of the Family Wellbeing intervention evaluation output to date, and to assess for evidence of the FWB intervention's impact on participants and their communities. The study found that where it is not appropriate or viable to conduct research, such as randomised control trials, alternative ways of providing evidence to demonstrate the effectiveness of an intervention is vital. This review suggests that Family Wellbeing interventions are having a positive impact and promoting change in the lives of participants, their families and their communities. Hence, for complex interventions, such as those with Indigenous populations, evidence should demonstrate effectiveness against prescribed outcomes, as well as critical aspects behind how and why a complex intervention was successful.

1. Introduction

Best practice is to develop interventions systematically, using the best available evidence and appropriate theory, then to test them using a carefully phased approach, starting with a series of pilot studies targeted at each of the key uncertainties in the design, and moving on to an exploratory and then a definitive evaluation (Craig et al., 2006, p.8).

Policy decisions should be based on evidence that demonstrates the effectiveness of interventions; however, the quantity and type of evidence that best supports the effectiveness of an intervention is not universally agreed upon and at times the evidence is not available. Generally, the strongest evidence is demonstrated through gold standard research interventions, most notably the Randomised Control Trial (RCT) (Campbell et al., 2000). Policy decisions based on strong evidence are increasingly more desirable as decision-makers operate in risk adverse public policy areas and have public accountability. Yet in areas of complex public policy, such as Indigenous health, waiting for the right type of scientific evidence, while other specialist knowledge is available, could be harmful. Perverse incentives, contrary to the decision-makers desired outcomes may occur where inaction results in continued or escalating poor health outcomes. Public health policy-

makers, attuned to the precautionary principle, need to consider how to account for growing complexity and uncertainty where complex interventions cannot provide the preferred gold standard evidence (WHO, 2004).

The concepts of precaution and prevention are at the heart of public health policy where identifying and avoiding risks is as critical as identifying and implementing protective measures. WHO (2004, p.3) explain that in the past, 'public health interventions focused on removing hazards that had already been identified and "proven" (even if the etiological mechanisms were not well understood).' In contrast, as potential risk factors increase in complexity, the precautionary principle seeks to tackle uncertain risk; hence, changing how science informs policy. Thus, this paper, in focusing on a complex situation, explores alternative approaches for the provision of quality evidence to support decision making.

Vujcich, Rayner, Allender, and Fitzpatrick (2016) demonstrate this principle emphasising that at times, policy makers must make policy decisions in the absence of strong evidence, particularly in circumstances where a lack of action is socio-politically unacceptable. This is particularly the case where practical constraints often make gold standard research methodologies impractical, such as a RCT in Australia's Northern Territory where a population level intervention, would

* Corresponding author.

E-mail addresses: leighann.onnis@jcu.edu.au (L.-a. Onnis), h.klieve@griffith.edu.au (H. Klieve), komla.tsey@jcu.edu.au (K. Tsey).

have required a sample of 30,000 Indigenous people, which in that context was impractical (Vujcich et al., 2016).

In their examination of the Indigenous Tobacco Control initiative Vujcich et al. (2016) found that policies were not made recklessly, instead, evidence from other settings and expert opinions were considered to generate conceptually plausible policy responses aligned to the preferences of members representing Indigenous communities. Furthermore, they found that '[e]vidence cannot trump all other factors; rather it must coexist with other inputs in the policy process' which includes the consideration of the needs of the target population (Vujcich et al., 2016, p.12). Similarly, Hudson (2016, p.1) reported that '[o]ne of the reasons given for the low return on Indigenous investment is that the money is not going to where it is needed most, or used in ways that respect Indigenous input into program design and delivery'. Therefore, Hudson (2016) reinforces Vujcich et al.'s findings proposing that interventions are unlikely to meet the needs of Indigenous communities unless Indigenous people are involved in the program's design and delivery. Generally, health intervention responses acceptable to Indigenous populations will also be supported by strong evidence; however, as demonstrated by Vujcich et al. (2016) there are circumstances in which it may be justifiable to enact interventions in the absence of gold standard evidence. In such circumstances, a better approach is to start with the available evidence, scientific and valued by Indigenous people, and incrementally strengthen the evidence base through approaches that provide complementary evidence, such as continuous quality improvement (CQI).

Systematic reviews, and reviews of reviews, reveal the variety of ways in which programs and services can be rated for effectiveness and impact (Day and Francisco, 2013; Dudgeon et al., 2014; Sanson-Fisher et al., 2006). While gold standard evidence is being sought for policy decisions, the path to such evidence should be suitable to the context and target populations. Often, systematic reviews use study quality assessment measures developed in the clinical context and may be, less suitable for evaluating the impact of interventions which have the complexity of psycho-social interactions. Moreover, when working with Indigenous populations, appropriate interventions must include Indigenous perspectives in the planning, developing, implementation and evaluation stages if the desired outcomes are to be achieved (Dudgeon et al., 2014; Hudson, 2016). Qualitative evaluations align with Indigenous traditions of storytelling; yet such approaches do not necessarily generate the evidence for policy decisions where more objective measures are preferred. As such, research techniques that endeavour to interrogate qualitative data to provide evidence of impact beyond narrative and thematic analyses are beneficial. Therefore, a detailed review of valued Indigenous interventions is needed, so that the uptake of programs that Indigenous people believe have impact is not constrained by the lack of the type of evidence required by decision-makers.

While the requirement for quality evidence and associated research impact is inevitably increasing, there are challenges in determining the appropriate methods and indicators for assessing research impact, particularly the attribution of a given outcome for a given intervention (ARC, 2017). Hence, developing such evidence requires time, so it is also imperative that interventions are sustainably supported while incrementally improving the evidence-base. One intervention program that has sought to address the challenges of social health research in providing evidence of impact is the Family Wellbeing Empowerment (FWB) Program which has adopted a phased approach to evaluating complex interventions (Campbell et al., 2000; Tsey et al., 2007; Tsey, 2015).

1.1. Family wellbeing empowerment program (FWB)

The Family Wellbeing Empowerment Program was created by a group of 'stolen generation' Indigenous people who felt that Indigenous people and families needed support in developing the skills and

capacity needed to face the daily challenges of being 'relatively marginalized minority peoples in a highly affluent Australian society' (Tsey et al., 2007, p.S35). Hence, the program addresses socio-economic disadvantage, and health inequality through fostering personal empowerment, social cohesion and community connectedness (Haswell et al., 2010; Kinchin, Jacups, Tsey, & Lines, 2015).

From modest beginnings in 1993, the program developed into a flexibly structured group learning experience combining Indigenous survival experiences and psycho-social holistic approaches for health and wellbeing, into the four stages of FWB: Foundations in Counselling; Coping with Grief and Loss; Changing and Working Together; and Moving Forward (Tsey et al., 2007; Tsey, 2000). While FWB has been the subject of ongoing evaluations, these have generally been qualitative focusing on individual perceptions of change and were necessarily based on small sample sizes. More recently, mixed methods evaluations have been conducted; however, these are also based on self-reported data from small samples. Consequently, the strength of the evidence of impact is often dismissed as insufficient to link the intervention directly to positive outcomes. The adoption of approaches to synthesise similar interventions to provide stronger evidence for impact highlights the complexity and additional challenges in establishing strong evidence for the effectiveness in social health research. Efforts to review and synthesise FWB 'research to date' by Tsey, Harvey, Gibson, and Pearson (2009a, 2009b) were in essence a review of the evidence being carried out by researchers already familiar with the intervention. Hence, the aim of this review was to collaborate with independent researchers (authors one and two) who have not been involved directly in FWB interventions to: 1) lead a review of characteristics of the FWB evaluation output to date, and 2) to assess for evidence of the FWB intervention's impact on participants and their communities. In leading the review, these independent researchers developed the inclusion criteria, conducted the article selection, analysed the characteristics of the selected articles, developed the matrix protocol and analysed the data to populate the matrix. All three authors contributed to the development of the program logic, the discussion section and the explanation of the implications for future research in this area.

In this review, we propose that while policy makers need rigorous evidence to support decision making; such evidence is not always available due to ethical, cultural and practical situations. Furthermore, Indigenous issues require Indigenous involvement and Indigenous perspectives to ensure that policy decisions and strategies address the real issues as identified by Indigenous people. Therefore, a way forward is to conduct research where multiple sources of evidence are collected and analysed to develop a rich picture of the situation, and opportunities for data triangulation from multiple sources. This approach offers more than the sum of the quantitative and qualitative evidence; the synthesis can provide more depth which is likely to improve understanding and support implementation. Thus, this review uses practical solutions to achieve outcomes that increase the evidence-base, which is more pragmatic than demanding the perfect evaluation which cannot be delivered.

2. Method

A literature review was conducted in a manner consistent with the PRISM statement (Fig. 1). This used the search term 'family wellbeing program AND Aboriginal' using *One Search* which searched multiple databases including: informit, CINAHL, EBSOHost, OvidSP, OvidMP, PubMed, ProQuest and the Wiley online library. The 2982 matches were screened by title and abstract against the inclusion criteria below.

Inclusion criteria for the review of characteristics of the FWB evaluation output to date:

1. The primary focus of the paper was improving health and wellbeing outcomes through a FWB intervention; or
2. The paper contains a multi-component study that includes health

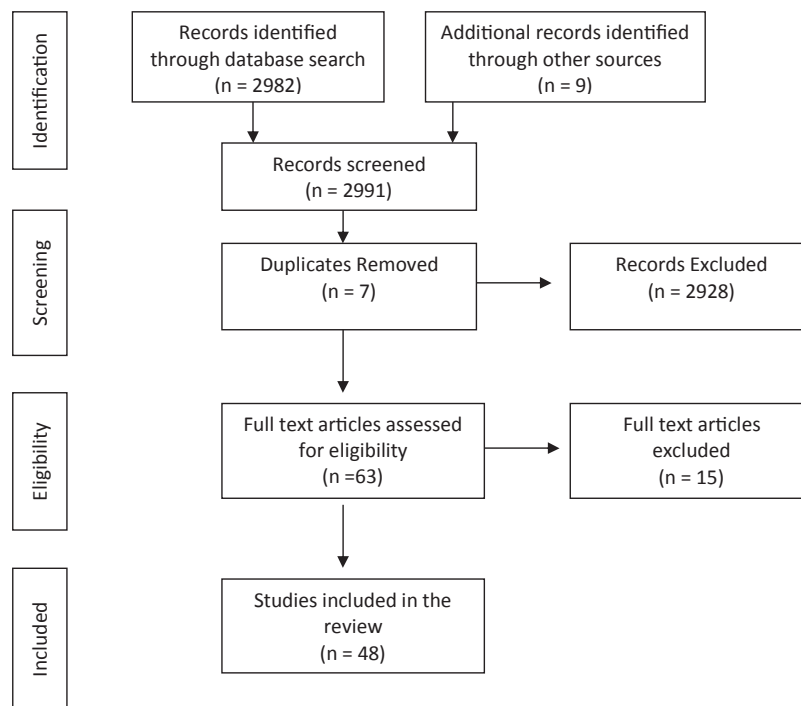


Fig. 1. Literature review using PRISMA statement.

and wellbeing outcomes for a FWB intervention; and
3. The paper was published, or prepared for peer-review publication

Additionally, nine papers were identified from key authors in this field that were under review or prepared for publication at the time of writing this article. Following screening, 63 papers were included in the full paper review, with 48 papers selected for inclusion in this review.

The review commenced with an analysis of the broader characteristics of the publications. Data analysis was conducted using Microsoft Excel and IBM SPSS 23. Next, the program logic framework was used which draws on the underlying assumptions for the FWB Program to define the needs, resources, activities, outputs and outcomes of the FWB intervention providing the logic to the anticipated outcomes and outcome usage beyond the program (Inline Supplementary Fig. 1). Program logic enables researchers to consider the logic of the progression of a program from needs to outcomes, and can assist the development of strategy and policy (Day & Francisco, 2013).

The FWB program responded to a need identified by Indigenous Australians, and particular inputs and resources are required for the intervention. Through the program logic, it is apparent that the inputs and activities result in a range of outputs (e.g. trained FWB facilitators and research publications); and program level outcomes, (e.g. improvements in individual health), as well as outcomes beyond the project level (e.g. documented uptake of FWB, an Aboriginal developed and facilitated program, uptake across Australia and beyond). This program logic then guided the exploration of the evidence with, the results section examining evidence of engagement and impact in the selected publications (Fig. 2). This evidence is reported using the source of identified needs, funding sources, industry engagement, research design and methodology, outputs and outcomes to structure the findings.

It was evident from the identified studies that a larger proportion were based on qualitative than quantitative techniques, constraining the development of a synthesis of impact. To best capture the information from the analysis, the subsequent analysis for evidence of the FWB intervention's impact on participants and their communities used a subset of papers for more in-depth consideration regarding impact using the following inclusion criteria:

1. The paper included empirical research that collected data from; and
2. The paper was an evaluation, including demonstrated experience (DE) where the publication reported on the impact of FWB on the authors own life, which is in itself was evidence of evaluation, and economic evaluation (EE).

This process left 29 papers identified for further assessment of intervention impact (Inline Supplementary Table 1 contains information about the papers identified as containing empirical evaluations). Both quantitative and qualitative studies reported on smaller pilot studies but the latter provided rich, in-depth assessments of programs and thus gave insights into the how and why impacts from the interventions occurred. It was apparent that a strategy was needed to synthesise the largely qualitative evidence into more objective measures without compromising the richness of the qualitative studies to enable comparison with the quantitative studies. Two tools that assess impact and change – Bloom's Taxonomy, and the Prochaska and DiClemente's Stages of Change model – were used to assess the reported findings in each publication for evidence of positive change from the intervention attributed to the FWB program.

Bloom's Taxonomy is a widely used hierarchy, from lower-order cognitive processes classified as *remember*, *understand*, through to higher-order cognitive processes such as *apply* and *analyse*, *evaluate* and *create* (Bush, Daddysman, & Charnigo, 2014). Prochaska and DiClemente's Stages of Change model moves through six steps from: *pre-contemplation*, *contemplation*, *preparation*, *action*, *maintenance*, and then *relapse* (Prochaska & Norcross, 2001). Prochaska and Norcross (2001) report that this model has been widely used to explore behavioural change over a 20 year period, and Campbell et al. (2013) used Prochaska and DiClemente's model in their study with Indigenous Australians to examine the impact on health promotion programs on tobacco smoking cessation.

Firstly each paper was assessed against the four Bloom's Taxonomy levels for the key dimensions of the FWB stage one program: leadership, basic human needs, relationships, life journey, conflict resolution, understanding emotions, dealing with crises, beliefs and attitudes (Inline Supplementary Table 2). Following this step, using the same coding methods, the publications were then assessed against the six stages in

Program Logic Steps	Identified needs	Funding	Industry engagement	Design and Methodology	Research Outputs	Blooms Taxonomy	Cycle of Change
Identified Need	■						
Inputs/Resources			■				
Activities				■	■		
Outputs					■		
Program level outcomes						■	■
Beyond Program level outcomes						■	■

Fig. 2. Method for exploring the evidence of FWB impact.

the Prochaska and DiClemente’s Stages of Change model (Inline Supplementary Table 3). From this a matrix presenting, the evidence of impact from the learning, and change indicators was produced to identify publications that reported evidence of both learning and change (presented in the results section).

3. Results

3.1. Characteristics of the studies

An examination of the selected publications illustrates the breadth and depth of the research on FWB, and demonstrates a continued research effort over a 17 year period. Using Sanson-Fisher et al.’s (2006) classification of publications, the majority were original research (79%), with the remainder classified as case studies (19%) and program descriptions (2%)(Inline Supplementary Table 4). The publications provide an account of the development, adaptation and impact of an intervention program developed by Indigenous people that inclusively enables Indigenous participants, organisations and researchers to collaborate with university-researchers. The university-based researchers were affiliated with 12 different universities in a variety of contexts and locations in four countries: Australia (88%), PNG (8%), China (2%) and Timor Leste (2%). The majority of the Australian research sites were in Queensland (58%), Northern Territory (19%), Victoria (4%) and New South Wales (2%), with remainder not citing a specific location (17%).

The interventions varied, with the scope of the FWB intervention ranging from a 1-day workshop to an accredited Certificate III in FWB. Sample sizes ranged from 1 to 158 (excluding the synthesis and tool validation studies), with the average sample size being 33 participants. Participants were aged from 9 to 60 years. There were mixed groups of both male and female participants; and gender-specific groups (e.g. men’s groups). Overall, where the gender of participants is reported, there were more female (58%) than male (42%) participants. Most frequently, the publications reported a FWB stage one intervention (21%); and versions of the stage one intervention tailored to meet group needs (25%). There were 34 papers that contained details about participant groups to whom FWB has been delivered: community service employees (35%), residents of Indigenous communities (32%), FWB facilitators and researchers (including community-based researchers) (26%); university students (15%); young people (6%); users of community services (3%); men’s groups (15%); school children (12%) and department of education employees (3%). Approximately, one-quarter of the papers (27%) reported findings from pilot studies suggesting CQI approaches were used to incrementally build evidence. Less than one third (31%) of the papers contained data that was collected post-intervention, and of those that contained post-intervention data, almost half were published since 2015. The duration post intervention for data collection ranged from immediately after the completion of the intervention to more than one year later (see Inline Supplementary Table 4). One study (Tsey et al., under review) intended to collect post-intervention data; however, were unable to collect the data ‘due to a lack of

dedicated evaluation resources to engage and negotiate follow-up data.’

3.1.1. Identified need

The evidence of engagement, usage and impact for FWB interventions is outlined in this section, as well the inputs/resources, and activities sections (see Inline Supplementary Table 1 and Inline Supplementary Table 4). Approximately half (54%) of the papers reported an invitation to conduct the research (e.g. Aboriginal community, men’s group), with many papers not explicitly stating how the need was identified. Also, many described ongoing partnerships and that the research continued due to previous or current university-industry partnerships suggesting that the ongoing partnerships were the conduit to continued research with these industry and community partners.

3.1.2. Inputs/resources

Four key authors were identified: Tsey (44 papers since 2000), Whiteside (24 papers since 2003), McCalman (17 papers since 2005) and Cadet-James (12 papers since 2004). These authors have consistently co-authored papers for more than ten years, and have collaborated with many industry and university authors, with 62 individual authors named across the publications. While most of the authors were university-based, one third (34%) were from industry and more than half (57%) of these industry authors were Indigenous. Overall, one third (35%) of all authors were Indigenous, with five publications having an Indigenous first author.

The analysis revealed that 65% of the papers reported financial support: Australian Competitive Grants (40%); other public sector research income (21%) and industry and other research income (4%). While only one third of the papers reported in-kind support: Industry (25%), CRC (4%), university (4%); this support came from 23 sources (university (65%), industry (26%), and CRC (9%)).

3.1.3. Activities

3.1.3.1. Purpose. The FWB-related activities were carried out for a range of reasons, with most papers reporting the purpose was to address social, health and/or political challenges. There were 32 papers that explained why FWB was delivered. The most frequent purpose was empowerment (28%) followed by improving health and wellbeing (22%); improving employment outcomes (16%); addressing interpersonal violence (6%); addressing substance abuse (6%); developing resilience (6%); promoting suicide prevention (16%); and improving education outcomes (6%).

The purpose of the research was also explored, (see Inline Supplementary Table 1) and while the reasons varied, most frequently, research was conducted to determine effectiveness (37%), build theories/models (19%), evaluate impact (13%), describe the intervention (8%), assess feasibility in another context (17%), or to engage with industry (6%).

3.1.3.2. Scope of intervention. The responses regarding the scope of the

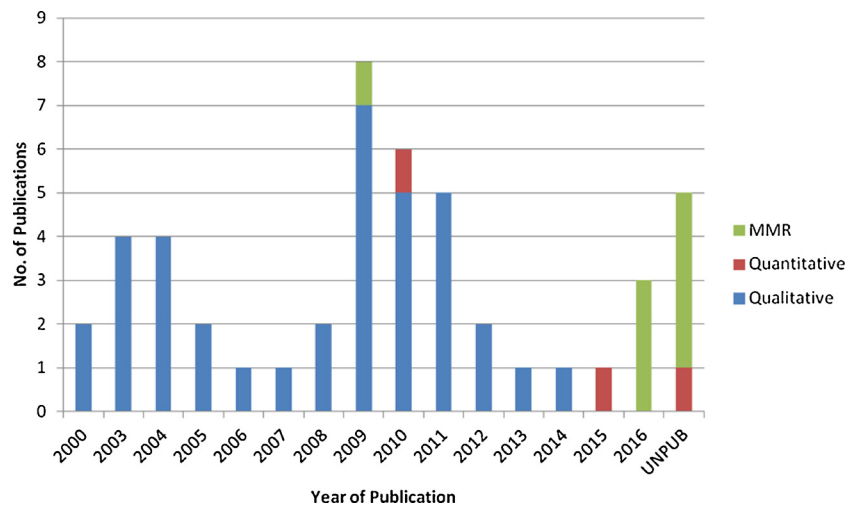


Fig. 3. Research Methods (n = 48).

FWB intervention were quite varied. There were 28 papers that described the scope, these varying from one-day training to studies conducted over many weeks (see Inline Supplementary Table 4). Most interventions involved 21–40 h (36%), some were more than 40 h (21%), and few were less than 20 h (4%). One quarter of the papers did not state the length of the intervention; instead they reported that they delivered FWB stage one (14%) or an adapted version of FWB (11%). Finally, two papers reported on participants who undertook accredited training and two papers reported on FWB as it was integrated into university curriculum.

3.1.4. Outputs

The first publication was in 2000, with consistent publications since 2003, and a sharp increase from 2009 to 2011 (Fig. 3). Analysis revealed that the increase was consistent with an increase in research activity associated with increased funding and an Industry partnership during this period (Fig. 3). There also has been a shift in study type from the initial qualitative research (2000–2008). While the greater majority of papers reported on qualitative studies (77%), the first of the three quantitative studies commenced in 2010, then 2015 and one in preparation for publication. One mixed methods study was published in 2009; however, the majority of such studies were more recent (2016) or are yet to be published.

3.1.5. Outcomes

Publications reported positive outcomes through statistically significant changes in self-reported quantitative data, or by referencing stated observations of changes in self-reported qualitative reports. A selection of the evidence suggesting that FWB had a positive impact is presented below.

Participants and FWB facilitators directly attributed change to the FWB program that they attended or facilitated, with comments recognising this change:

‘I am a changed person, because this course helped me to evaluate myself and at the same time, has empowered me to do more for other people’ (Kitau, Kinchin, Whiteside, Hane-Nou, and Tsey, 2016, p.24).

‘FWB has given me a strength to identify real needs in the community’ (McCalman, McEwan, Tsey, Blackmore, & Bainbridge, 2010a, 2010b, p.15).

Others recognised their ownership of this program and how it has moved their perspectives:

‘Our mob when they hear that it’s been developed by our own

people, that’s the only reason why sometimes I think they come along to it. So I think that’s the most critical thing. And that it works of course, but you know, people don’t know that it’s going to work until they’ve done it’ (McCalman, 2013, p. 5).

‘Why didn’t I do this sooner? We actually find ourselves beginning to live FWB. It is something that you are doing, you are teaching that to people but you are actually living it’ (Mayo & Tsey, 2009, p. S69).

‘Doing FWB has taught me skills that enabled me to realise that I need to meet my own needs. I need to look after myself rather than looking after everyone else....I wanted to stop feeling like I was a victim. I wanted to stop taking on board other people’s problems when it wasn’t my problem. I wanted to be more assertive, not to be demanding. I wanted a good home life. I was depressed sometimes. Now when this happens I can examine what’s making me feel like that. I needed to learn to look after myself’ (Tsey, 2008, p.12).

The quantitative and mixed methods research studies piloted research tools, including the development and validation of a research tool, the Growth and Empowerment Measure (GEM), specifically designed to measure impact for Indigenous populations. Other studies used validated measures such as versions of the Kessler Psychological Distress Scale (K5, K10 and the K9), the Australian Unity Wellbeing Index (AUWI) and questionnaires designed to measure change in social, emotional and physical health (e.g. physical activity, health eating and alcohol intake). Several, quantitative and mixed methods studies identified evidence of significant positive impact, for example:

Whiteside et al., 2016 Whiteside et al. (2016, p.248) reported that the K5 ‘showed a highly significant reduction in psychological distress across the time of the study ($t(12) = 3.67$ $p = .003$) with a very strong effect size ($d = 1.02$).’

Kinchin et al., 2015 Kinchin et al. (2015, p.4) found positive change for participants post-intervention, reporting that ‘the GEM responses on self-capacity, inner peace, strength, happiness and connectedness (questions 1–14) indicated a 17% positive change in the mean scores’ and that for the AUWI ‘the most satisfactory post-intervention response was provided on Future security which was estimated even higher than the national benchmark’.

Lui (unpublished) found evidence of FWB program impact in the FWB study conducted in China, reporting that ‘All aspects of the wellbeing scores before and after the training significantly improved (all p -values < 0.001) in both rounds of data collections.’ In addition, they found that the means for the leadership scores were ‘highly significantly different (p -value < 0.001) before and after the

training in both round of data collections.’

3.1.6. Assessing impact and change

The assessment of impact and change was undertaken on the 29 identified papers using Bloom’s Taxonomy, and Prochaska and DiClemente’s Stages of Change model, with scoring used to capture evidence of positive change identified through the research. For example, in a quantitative study, the identification of a statistically significant change in participants’ satisfaction with life (Kitau et al., 2016) provided evidence of impact for a positive change in ‘beliefs and attitudes’. Similarly, in a qualitative study, a story describing how a participant came to see ‘that to every problem you find a solution ...you can achieve whatever you want to in life’ (Whiteside, Tsey, Cadet-James, & McCalman, 2014) provided evidence of impact for a positive change in ‘beliefs and attitudes’.

Following the assessment of all papers, the highest score for each measure (ie. Bloom’s Taxonomy and Prochaska and DiClemente’s Stages of Change) was identified from the eight different dimensions of FWB: leadership; basic human needs; relationships; life journey; conflict resolution; understanding emotions; dealing with crises; and beliefs and attitudes assessments. This recognised the level of change observed on one or more of the eight dimensions of FWB. From this the publications fall into two distinct categories summarised in Fig. 4. Group 1, with low scores (score ≤ 3 4) and Group 2 with high scores (score between 4 and 6). Group 1 contained 13 publications, the majority of which were qualitative (77%), and the remainder were mixed methods studies (23%). Group 2 contained 16 papers, once again, the majority were qualitative (75%); however, there were also several mixed methods (19%) and one quantitative study (6%).

The publications that scored the highest across both measures provide interesting results for how we consider FWB impact. Four publications reported findings from qualitative studies that analysed data from in-depth interviews, reflective diaries, and existing documentation, and three (e.g. P10, P4 and P28) were written by Indigenous authors (published between 2004 and 2013). Interestingly, this synthesis revealed that evidence of impact, in terms of comprehension and behavioural change, was more likely found in the qualitative studies.

Fig. 4 reveals that 28 publications contained evidence of comprehension with the highest mean score for the publications translating into evidence of impact, in order of hierarchy, with evidence that participants remembered (25%), understood (32%), analysed (29%) and created (14%) which they attributed to the FWB program. Similarly, there were 22 publications that revealed evidence of change, with 77% containing evidence of action (including maintenance), 9% containing evidence of preparation and 14% containing evidence of contemplating change.

Overall, the review of the publications suggests that a staged approach to evaluating complex social interventions can provide evidence of program impact for participants, their families and communities. The results presented in this section provide a picture of the development and growth of an intervention that has systematically built an evidence-base over many years, often using CQI approaches, such as participatory action research (PAR) and pilot studies.

4. Discussion

The innovative approach used in analysing the data to assess the impact and change evidence from FWB interventions provided insights that extended from the original papers and synthesis of the study’s characteristics. This approach revealed evidence of action, that is, evidence of participants actually using the knowledge and skills developed from the intervention in their lives at the time of data collection. Further, the four papers that scored highest on Bloom’s Taxonomy (scored a six) and highest on Prochaska and DiClemente’s Stages of Change (scored a four or five), included post-intervention data. Thus, contributing to the argument that to demonstrate impact sufficient time must pass between the intervention and the post-intervention data collection to allow participants an opportunity to change. While this review revealed that continuous staged evaluation of FWB over seventeen years has provided evidence of its impact, it also highlighted further considerations for use and interpretation of findings and also future research design to contribute to enhanced evidence from such studies.

We posed the question; how much evidence is needed to secure the

		Bloom’s Taxonomy							
			Remember	Understand	Analyse	Apply	Evaluate	Create	
		Average score	0	1	2	3	4	5	6
		0	P27 P47	P17 P22 P26 P29	P18 P36				
Prochaska & DiClemente	Pre-contemplation	1							
	Contemplation	2	P44	P15	P6				
	Preparation	3		P46	P25				
	Action	4	P41	P31 P32 P48	P3 P14 P20	P1 P9 P19 P23 P45		P10	
	Maintenance	5						P4 P7 P28	
	Relapse	6							

Fig. 4. Synthesis of Bloom’s Taxonomy and Prochaska and DiClemente’s Stages of Change (n = 29).

confidence of policy-makers to support complex interventions that report positive results over a long period, across varied populations where access to culturally appropriate interventions is limited? As this review revealed, in addition to the usual practical and methodological difficulties, many of the added challenges for interventions, as seen with FWB, are associated with the difficulties in obtaining sufficiently sized samples of comparable data to demonstrate impact. The review findings highlight the range of studies conducted over the 17 year timeframe, and while the CQI and phased approaches have supported the growth and increased sophisticated in evaluation methodologies during the period, an absence of consistent reporting of aspects of the research (e.g. in-kind support, how the need was identified) was seen. Further, the variety of data collection tools used (e.g. K5, K10, modified questionnaires) reduces the benefits that a synthesis can bring to providing evidence of intervention impact. While the key consideration when evaluating complex interventions is, pragmatically, more about practical effectiveness, thus whether the intervention works in regular practice, benefit is seen in researchers using a standardised set of core measures to enable comparable data to be analysed from small interventions, to support the development of a robust evidence-base for complex phased interventions across multiple locations. This would contribute to the evidence of impact beyond the program-level, where evaluating the effectiveness of an intervention is important to our understanding about how the range of effects of the intervention, vary over time and between locations, as well as the causes of that variation (Craig et al., 2006).

The review revealed that for phased interventions, such as FWB, mixed methods research may contribute to strengthening the evidence for impact that satisfies funders, researchers and target populations, while continuing to collect the rich qualitative data. For example, a standard set of questions which would contribute to a set of core measures used across all interventions. In terms of the FWB program, the findings from this review suggest that a standard set of core measures including the Kessler Psychological Distress Scale, Australian Unity Wellbeing Index, and a small set of common questions to address key indicators that reflect changes in behaviour would provide more consistent findings from each small study to be synthesised to provide more understanding of the impact beyond each individual program. Additionally, it would be valuable for researchers to access publicly available community level data in factors such as school attendance, or child welfare notifications, before and after interventions as an additional measure of change. With policy makers requiring more evidence from research using methodologies more suited to medical research, it can be difficult for social health research which uses more qualitative methodologies to develop an evidence-base that is judged sufficiently rigorous to attract similar support and funding. Hence, this approach would enable the flexibility for each pilot study to tailor and adapt the evaluation using CQI approaches based on context and purpose of the intervention, while at the same time developing a larger quantitative dataset to analyse, all of which builds evidence for the impact of FWB, while maintaining the integrity of the intervention.

In their investigation about how research evidence influences health policy decisions, Vujcich et al. (2016) report that their literature review found that ‘most studies had methodological limitations’ and that few studies used empirical evidence ‘to demonstrate why evidence-based policy making is superior to other forms of decision-making’ (Vujcich et al., 2016, p.5). Furthermore, there can be negative implications from decision making relying on, for example, gold standard evidence. Vujcich et al. (2016, p.13) explained that when applied to a context such as Indigenous tobacco control, the position that ‘there should be “no policy without evidence” could feasibly have resulted in a widening of the health gap between Indigenous and non-Indigenous Australians simply because there was more research to demonstrate the effectiveness of interventions in the latter population’. Moreover, Birch (1997) describes the paradox of this approach where a lack of evidence of effectiveness restricts resources to a disadvantaged group, thus reducing

the resources available to provide evidence of effectiveness and further disadvantages the group who were the target of the initial intervention. Where effectiveness is demonstrated in a mainstream, less disadvantaged population, resources are redirected to the population where the evidence of effectiveness can be demonstrated (Birch, 1997). Ultimately, this contributes to further social inequalities in the health of the disadvantaged population.

Irrespective of what is deemed to be sufficient evidence for decision-makers and academics alike, the demand for FWB remains strong. In fact, the demand for FWB which arises from Indigenous people completing the FWB program and subsequently suggesting that friends and family do the program as well, results in strong support for the FWB intervention at an individual, organisational and community level. Furthermore, the spread of FWB internationally, together with the demand for FWB across Australia ‘clearly shows the extent to which Indigenous people are willing to vote with their feet in order make a program that they value available to themselves’ (Whiteside et al., 2014, p.5).

With increased demand for available resources, it is essential that resources are focused in areas that are not only perceived as important for Indigenous populations, but also are demanded and deemed to be effective by the same populations (Vujcich et al., 2016). To date, the success of FWB is in part driven by the fact that participants encourage others to attend and find ways for FWB to become part of their community and/or workplace. There is undeniably ‘consumer demand’ for FWB, surely an essential first step in any intervention program. Hence, strategies are necessary to enable and support implementation and up-scaling of programs that work, such as FWB, which will require a re-think of the way in which decision-makers view the evidence that shows that a complex intervention is effective.

4.1. Lessons learned – mapping the way forward

Complex intervention programs, such as FWB, need strategies to build evidence for their effectiveness appropriate to the context of the intervention. As such, it is unlikely the future evidence will be in the form of large scale quantitative studies or RCTs for the reasons explained earlier in this review (Campbell et al., 2000; Craig et al., 2006). Therefore, it is imperative that researchers undertaking complex interventions conduct and publish evaluations so that a stronger evidence-base can inform a more strategic way forward for implementing sustainable complex interventions. Future strategies that could improve the available evidence about the effectiveness of complex intervention should include:

- Longitudinal studies that can provide evidence about long term change resulting from the intervention;
- More consistent reporting of funding and in-kind support for the intervention;
- A more structured approach in the development of evaluation tools for collecting quantitative data so that the data from multiple studies can be synthesised and analysed. This would enable data analysis with larger sample sizes; and may enable comparisons with other datasets, such as, routinely available data, (e.g. national datasets, health and education datasets specific to Indigenous populations);
- More focused effort in reporting attribution of the intervention for impact at both a program level and beyond the program level; and
- Recognition that for social interventions the benefits may not be seen for months or years after the intervention, particularly where there is a focus on improving impact through engaging and partnering with Indigenous populations and industry. Consequently, funding should be allocated for the development and engagement costs for interventions that are required prior to the research being conducted, and for post intervention research to capture impact months and years after the intervention, if the population level

impact of the intervention is to be reported.

In addition, the CQI approach of staged evaluation and building feedback into future iterations of the programs is a sensible path to building robust evidence (Laycock, Bailie, Matthews, and Bailie 2016). However, this approach takes time and needs commitment from governments, as well as university and industry partners as it often does not generate the positive results against quantitative indicators in the short-term that constitute the preferred evidence for reporting on intervention impact.

4.2. Limitations

There was a limitation in that the majority of the research conducted has been qualitative with self-reported changes attributed to FWB quoted as evidence of FWB impact. The validation of research tools (e.g. GEM), and use of independently developed scales (e.g. Kessler, Australian Unity Wellbeing Index) were increasingly used in the most recent publications suggesting that while self-reported data remains a limitation, the studies are now incorporating widely accepted, robust scales into data collection tools.

The wide variation in the scope of the intervention (e.g. one day workshop to accredited certificates) is also acknowledged as a limitation; however, the customisation of the FWB intervention which is tailored according to need, means that the scope of the intervention may not be as important as the quality, thus a focused brief intervention may meet participants needs as well as a generic two-week intensive certificate course. Furthermore, there was evidence of impact in publications that reported brief interventions (< 20 h) as well as accredited certificate training. Hence, while scope of intervention is acknowledged as a limitation, it was not believed to be sufficient to influence the findings from this review.

While the synthesis aimed to treat the qualitative and quantitative data in a similar manner to aid comparison, differences in the publication type and style are acknowledged as a limitation for this review. In particular, some publication types, for example, reports, books and longer journal articles provided more opportunity to present evidence. This review did not count the frequency of evidence, it was either present or not, which minimised the influence of publication type and style; however, it remained a limitation for the review.

5. Conclusion

Where it is not appropriate or viable to conduct research, such as RCT, that provides evidence at the level preferred by decision-makers, alternative ways of providing evidence for effectiveness of an intervention is vital. This review suggests that FWB is having a positive impact and promoting change in the lives of participants, their families and their communities. Therefore, it is reasonable to continue by using the available evidence, while continuing to evaluate and implement recommendations to provide a deeper understanding of the long term benefits of FWB. However, to do this it is vital that decision-makers understand the limitations in either the practicality or application of approaches such as RCT assessments in some situations and thus recognise that where alternative research approaches are appropriate and methodologically sound, it is sensible to continue to evaluate, examine and interpret the available evidence through such approaches. For complex interventions, such as with Indigenous populations, evidence should demonstrate effectiveness against prescribed outcomes, as well as critical aspects behind how and why a complex intervention was successful.

6. Funding

This work was supported by funding from the Lowitja Institute, and the Centre for Research Excellence in Integrated Quality Improvement

(CRE-IQI) (NH&MRC ID 1078927), and in-kind support from James Cook University and Griffith University.

7. Declaration of interest

The third author declares that they have received grants to conduct some of the family wellbeing intervention research evaluated in this study; however, there are no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.evalproplan.2018.05.003>.

References

- Australian Research Council (2017). *ARC engagement and impact consultation paper*. [Available: http://www.arc.gov.au/sites/default/files/filedepot/Public/ARC/consultation_papers/ARC_Engagement_and_Impact_Consultation_Paper.pdf (Accessed 29 August 2017)].
- Birch, S. (1997). As a matter of fact: Evidence-based decision-making unplugged. *Health Economics*, 6(6), 547–559.
- Bush, H. M., Daddysman, J., & Charnigo, R. (2014). Improving outcomes with bloom's taxonomy: From statistics education to research partnerships. *Journal of Biometrics & Biostatistics*, 5.
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., et al. (2000). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal*, 321(7262), 694–696.
- Campbell, S., Bohanna, I., Swinbourne, A., Cadet-James, Y., McKeown, D., & McDermott, R. (2013). Stages of change: Smoking behaviour and readiness to quit in a large sample of indigenous Australians living in eight remote north Queensland communities. *International Journal of Environmental Research and Public Health*, 10, 1562–1571.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2006). *Developing and evaluating complex interventions: New guidance*. UK: Medical Research Council [Available at: www.mrc.ac.uk/complexinterventionsguidance (Accessed 29 August 2017)].
- Day, A., & Francisco, A. (2013). Social and emotional wellbeing in Indigenous Australians: Identifying promising interventions. *Australian and New Zealand Journal of Public Health*, 37(4), 350–355.
- Dudgeon, P., Walker, R., Scrine, C., Shepherd, C., Calma, T., & Ring, I. (2014). Effective strategies to strengthen the mental health and wellbeing of Aboriginal and Torres Strait Island People. *Close the Gap clearinghouse*.
- Haswell, M. R., Kavanagh, D., Tsey, K., Reilly, L., Reilly, L., Cadet-James, Y., et al. (2010). Psychometric validation of the growth and empowerment measure (GEM) applied with indigenous australians. *Australian and New Zealand Journal of Psychiatry*, 44, 791–799.
- Hudson, S. (2016). *Mapping the indigenous program and funding maze*. Research Report 18 [ISBN: 9781922184726. Available: <https://www.cis.org.au/app/uploads/2016/08/rr18-Full-Report.pdf> (Accessed 29 August 2017)].
- Kinchin, I., Jacups, S., Tsey, K., & Lines, K. (2015). An empowerment intervention for Indigenous communities: An outcome assessment. *BMC Psychology*, 3, 29.
- Kitau, R., Kinchin, I., Whiteside, M., Hane-Nou, G., & Tsey, K. (2016). Effectiveness of the uptake and implementation of an Aboriginal Australian empowerment program in the context of public health training in Papua New Guinea. *Pacific Journal of Medical Sciences*, 16(2).
- Laycock, A., Bailie, J., Matthews, V., & Bailie, R. (2016). Interactive dissemination: Engaging stakeholders in the use of aggregated quality improvement data for system-Wide change in Australian indigenous primary health care. *Frontiers*, 4, 84.
- Lui, S., [unpublished]. Promoting wellbeing and leadership competencies among university students in China: The feasibility of an Australian approach.
- Mayo, K., & Tsey, K. (2009). Reflections on a community and university research collaboration. *Australasian Psychiatry*, 17, S68–S71.
- McCalman, J., McEwan, A., Tsey, K., Blackmore, E., & Bainbridge, R. (2010a). Towards social sustainability: The case of the Family Wellbeing community empowerment education program. *Journal of Social and Economic Policy*, 13.
- McCalman, J., Tsey, K., Wentong, M., Wilson, A., McEwan, A., Cadet-James, Y., et al. (2010b). Indigenous men's support groups and social and emotional wellbeing: A meta-synthesis of the evidence. *Australian Journal of Primary Health*, 16, 159–166.
- McCalman, J. R. (2013). The transfer and implementation of an Aboriginal Australian wellbeing program: A grounded theory study. *Implementation Science*, 8, 1–9.
- Prochaska, J. O., & Norcross, J. C. (2001). Stages of change. *Psychotherapy: Theory, Research, Practice, Training*, 38, 443–448.
- Sanson-Fisher, R. W., Campbell, E. M., Perkins, J. J., Blunden, S. V., & Davis, B. B. (2006). Indigenous health research: A critical review of outputs over time. *Medical Journal of Australia*, 184(10), 502.

- Tsey, K., Wilson, A., Haswell-Elkins, M., Whiteside, M., McCalman, J., Cadet-James, Y., et al. (2007). Empowerment-based research methods: A 10-year approach to enhancing Indigenous social and emotional wellbeing. *Australasian Psychiatry*, 15, S34–S38.
- Tsey, K., Harvey, D., Gibson, T., & Pearson, L. (2009a). The role of empowerment in setting a foundation for social and emotional wellbeing. *Australian e-Journal for the Advancement of Mental Health*, 8, 1–10.
- Tsey, K., Whiteside, M., Haswell-Elkins, M., Bainbridge, R., Cadet-James, Y., & Wilson, A. (2009b). Empowerment and Indigenous Australian health: A synthesis of findings from Family Wellbeing formative research. *Health and Social Care in the Community*, 18, 169–179.
- Tsey, K. (2000). An innovative family support program by and for indigenous australians: Reflections in evaluation practice. *Journal of Family Studies*, 6(2), 302–308.
- Tsey, K. (2008). Tackling endemic substance abuse among Indigenous Australians: The contribution of values-based family empowerment education (Part 2 of 2). *Forum on Public Policy*, 2008, 1–24.
- Vujcich, D., Rayner, M., Allender, S., & Fitzpatrick, R. (2016). When there is not enough evidence and when evidence is not enough: An Australian indigenous smoking policy study. *Frontiers in Public Health*, 4.
- WHO (2004). In M. Martuzzi, & J. A. Tickner (Eds.). *The precautionary principle: Protecting public health, the environment and the future of our children* [Available: http://www.asser.nl/media/2227/cms_eel_96_1_book-precautionary-principle-protecting-public-health-the-environment.pdf (Accessed 29 August 2017)].
- Whiteside, M., Tsey, K., Cadet-James, Y., & McCalman, J. (2014). *Promoting aboriginal health: The family wellbeing empowerment approach*. New York, NY, USA: Springer.
- Whiteside, M., Klieve, H., Millgate, N., Webb, B., Gabriel, Z., McPherson, L., et al. (2016). Connecting and strengthening young Aboriginal men: A family wellbeing pilot study. *Australian Social Work*, 69(2), 241–252.
- Dr Leigh-ann Onnis** is a researcher for the Indigenous Education and Research Centre and The Cairns Institute at James Cook University. She has a research focus on management and leadership, remote health workforces, social and emotional wellbeing and public health in the remote contexts.
- Dr Helen Klieve** is a researcher for the school of Education and Professional Studies at Griffith University whose research interest focuses on research applications addressing issue of disadvantage and social inclusion, with a particular interest in the development and use of applications that contribute to the capacity to contribute to effective information for decision making.
- Professor Komla Tsey** is Tropical Leader and Research Professor in Education for Social Sustainability within the College of Arts, Society & Education and The Cairns Institute at James Cook University. He provides leadership and works collaboratively to: undertake developmental research in the field of education for social sustainability; build a longer-term education for social sustainability collaborative research agenda; and mentor and support emerging researchers to become independent competitive researchers.