



THE UNIVERSITY OF  
MELBOURNE

# How to approach evaluation in health professions education research

Dr Louise Allen - [louise.allen@unimelb.edu.au](mailto:louise.allen@unimelb.edu.au)

19 August 2024

MCSHE Professional Development Seminar

# ACKNOWLEDGEMENT OF COUNTRY





# Who am I?

# What I Hope to Achieve



1

**Get you thinking more broadly about impacts/outcomes of educational interventions**

2

**Have you gain some insight into the limitations of outcome evaluation models – specifically the Kirkpatrick model**

3

**Get you thinking about how we can improve our evaluation processes**



THE UNIVERSITY OF  
MELBOURNE

# Outcomes or Impacts?


# What are Outcomes or Impacts?



**DEFINITION:** The multiple and varied short, medium and long-term effects that occur as a result of educational interventions



## Categorising the broad impacts of continuing professional development: a scoping review

Louise M Allen,<sup>1</sup>  Claire Palermo,<sup>2</sup> Elizabeth Armstrong<sup>3</sup> & Margaret Hay<sup>1</sup>

**CONTEXT** A number of systematic reviews have evaluated the impacts of continuing professional development (CPD). These reviews, due to their focused nature, may fail to capture the full range of impacts of CPD. This scoping review aims to explore the broader impacts of CPD with the intention of developing a categorisation of the types of impact of CPD.

**METHODS** The authors searched MEDLINE, CINAHL and ERIC databases for studies published between 2007 and 2017 that looked at the impacts of formal CPD programmes for all health professionals. Studies were independently screened for eligibility; one reviewer charted data for all included studies, a sample of 10% was reviewed by a second reviewer. The charted data were analysed using both qualitative and quantitative content analysis.

**RESULTS** The search returned 2750 manuscripts; 192 manuscripts describing 191 studies were included in this review. Most articles were from the USA (78 studies, 41%) and included medical doctors in the population (105 studies, 55%). Twelve

categories of impact were generated through conventional content analysis: knowledge, practice change, skill, confidence, attitudes, career development, networking, user outcomes, intention to change, organisational change, personal change and scholarly accomplishments. Knowledge was most commonly measured (103 studies, 54%), whereas measurement of scholarly accomplishments was the least common (10 studies, 5%).

**CONCLUSIONS** Existing evidence takes a narrow view when assessing the impacts of CPD. Emphasis on measuring impacts as knowledge, behaviour, confidence, skills and attitudes may be due to the widely accepted four levels of evaluation from the Kirkpatrick Model or because the majority of studies used quantitative methods. The categories proposed in this review may be used to capture a broader view of the impacts of CPD programmes, contributing to the evidence base for their value and translating into CPD programmes that truly transform health professionals, their careers and their practice.



SCAN ME

Knowledge

Practice change

Skill, ability,  
competence and  
performance

Confidence

Attitudes

Career development

Networking,  
collaboration and  
relationships

User outcomes

Intention to change

Organisational  
change

Personal change

Scholarly  
accomplishments



Photo by Steven Wright on Unsplash





THE UNIVERSITY OF  
MELBOURNE

# Evaluation in HPE



### What is your level of experience with evaluation?



1 I have never been involved with evaluation 10% 2



2 I am not currently involved in evaluation, but I have been previously 14% 3

3 I am newly involved in evaluation 33% 7

4 I have been involved in evaluation for several years 43% 9

Go to [wooclap.com](https://wooclap.com) and use the code **HIHNRT**



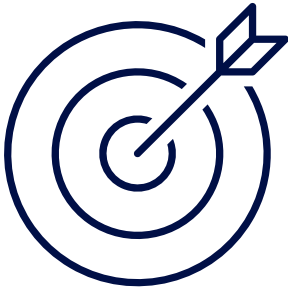
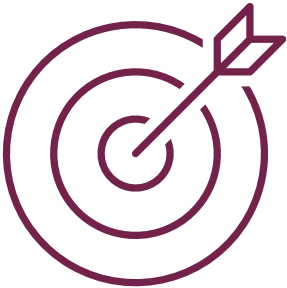
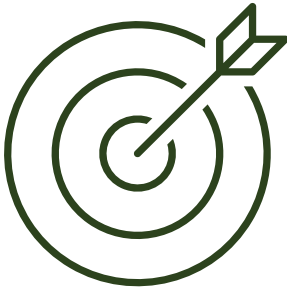
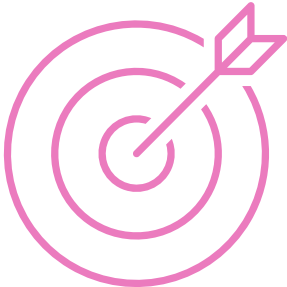
What evaluation frameworks or models do you use (or if you haven't been involved in evaluation have you heard of, or might you use)?



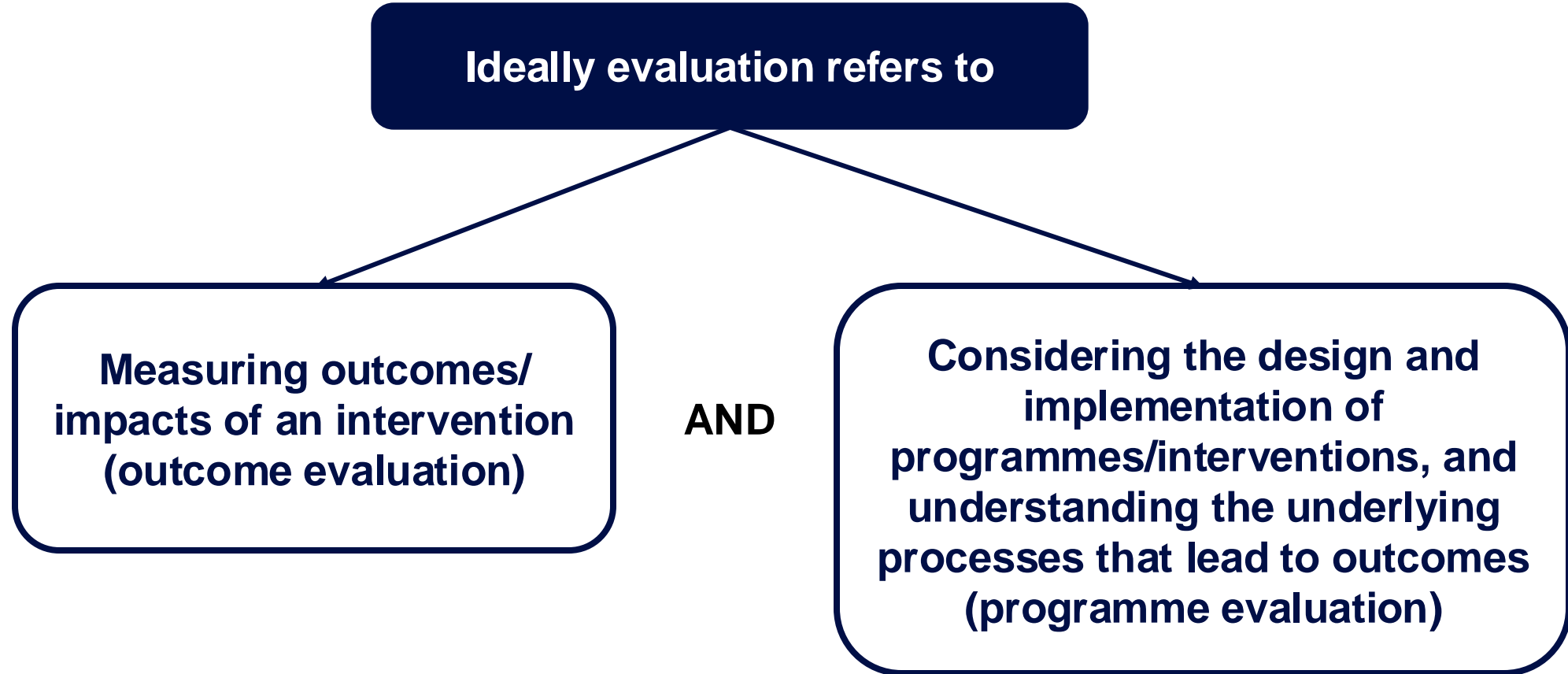
Kirkpatrick's  
**RE-AIM** **Kirkpatrick** CiPp Nothing  
New world Kirkpatrick 4-Ps CIPP Model Proctor Kitkpatrick's  
Kirkpatrick's (sorry )



# Why Evaluate?



# What is Evaluation?



# Evaluation in HPE



A range of evaluation models are used in health professions education




## HOWEVER

↑ use of the Kirkpatrick Model in health professions education despite:

- ✓ Published cautions on the use of the Kirkpatrick Model in the complex HPE environment
- ✓ Calls to move beyond focusing on outcomes to how and why programmes worked and what else happened
- ✓ A number of criticisms of the Kirkpatrick Model

STATE OF SCIENCE  
SITUATION

## Evaluation in health professions education—Is measuring outcomes enough?

Louise M. Allen<sup>1</sup>  | Margaret Hay<sup>1</sup>  | Claire Palermo<sup>2</sup> 

<sup>1</sup>Monash Centre for Professional Development and Monash Online Education, Monash University, Clayton, Victoria, Australia

<sup>2</sup>Monash Centre for Scholarship in Health Education, Monash University, Clayton, Victoria, Australia

**Correspondence**

Louise M. Allen, Monash Centre for Professional Development and Monash Online Education, Monash University, 29 Ancora Imparo Way, G.22, Clayton Campus, Clayton, VIC 3800, Australia.  
Email: louise.allen@monash.edu

**Funding information**

Australian Government Research Training Program (RTP) Scholarship

**Abstract**

**Introduction:** In an effort to increase the rigour of evaluation in health professions education (HPE), a range of evaluation approaches are used. These largely focus on outcome evaluation as opposed to programme evaluation. We aim to review and critique the use of outcome evaluation models, using the Kirkpatrick Model as an example given its wide acceptance and use, and advocate for the use of programme evaluation models that help us understand how and why outcomes are occurring.

**Methods:** We systematically searched OVID medline, Scopus, CINAHL and Pubmed, and hand searched six leading HPE journals to provide an overview of the use of the Kirkpatrick Model as well as a range of programme evaluation models in HPE. In addition to this, we synthesised the existing critiques of the Kirkpatrick Model as an example of outcome evaluation, to highlight the limitations of such models.

**Results:** The use of the Kirkpatrick Model in HPE is widespread and increasing; however, studies focus on categorising outcomes, rather than explaining how and why they occur. The main criticisms of the model are as follows: it is outcomes focused and fails to consider factors that can impact training outcomes; it assumes positive casual linkages between the levels; there is an assumption that the higher-level outcomes are more important; and unintended impacts are not considered. The use of the Kirkpatrick Model by the MERSQI, BEME and WHO contribute to the myth that the Kirkpatrick Model is the gold standard for programme evaluation.

**Discussion:** Moving forward, evaluations of HPE interventions must shift from focusing largely on measuring outcomes of interventions with little consideration for how and why these outcomes are occurring to programme evaluation that investigates what contributes to these outcomes. Other models that facilitate the evaluation of the complex processes that occur in HPE should be used instead of Kirkpatrick's.



THE UNIVERSITY OF  
MELBOURNE



SCAN ME

# The Kirkpatrick Model

## What is it?



Kirkpatrick Model	
<b>Level 1 – Reaction</b>	Measures how those who participate in the programme react to it.
<b>Level 2 – Learning</b>	Measures the extent to which participants change attitudes, improve knowledge, and/or increase skill as a result of attending the programme.
<b>Level 3 – Behaviour</b>	Measures the extent to which change in behaviour has occurred because the participant attended the training programme.
<b>Level 4 – Results</b>	Measures the final results that occurred because the participants attended the programme.

# Programme Evaluation Models

## What are They?



### ON THE OTHER HAND, PROGRAMME EVALUATION MODELS:

- ✓ Go beyond answering 'Did it work?' to 'How and why did it work?'
- ✓ Concerned with the design and implementation of the programme, in addition to the programme outcomes

**GOAL:** 'to identify the sources of variation in program outcomes both from within and outside the program, while determining whether these sources of variation or even the outcome itself are desirable or undesirable' (Frye & Hemmer, 2021, p. 289).

# Programme Evaluation Models – Realist Evaluation



Seeks to explain:

“

*“what kinds of outcomes, for whom, in what respects, in what contexts, how, why and when not, why not?” (Westthorp et al., 2018, p. 47)*

”

C

M

O

Context Mechanism Outcome

# Programme Evaluation Models – CIPP Evaluation



## Context

Assess needs, problems, assets and opportunities + relevant contextual conditions and dynamics

## Input

Examining alternative approaches and assessing feasibility and potential cost-effectiveness

## Process

Monitors and assesses implementation of the intervention

## Product

Assesses costs and outcomes

# Programme Evaluation Models – Theory-Driven Evaluation



Seeks to understand:

“

*What outcomes are occurring and the “hows and whys of programme success or failure.” (Chen, 2015, p. 232)*

”

Change Model

Action Model

# Programme Evaluation Models – Contribution Analysis



Seeks to understand:

*How much of a difference (or contribution) has the intervention or programme made to the observed outcomes? (Mayne, 2011)*

Theory of  
change

Proximal  
Outcomes

System wide  
impacts

# Programme Evaluation – Layered Analysis



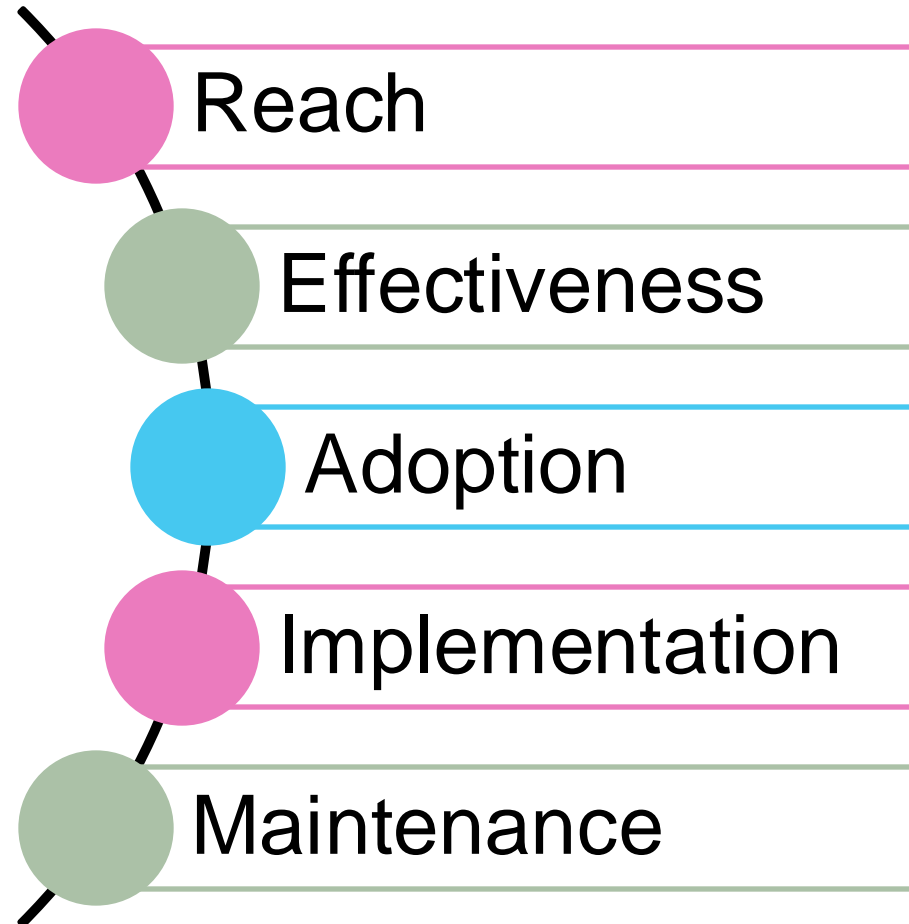
Seeks to understand:

“

*“What is the intended intervention” and “Did that intervention, in fact, occur?” (Cianciolo & Regehr, 2019, p. 789)*

”

# Programme Evaluation Models – Re-AIM Framework



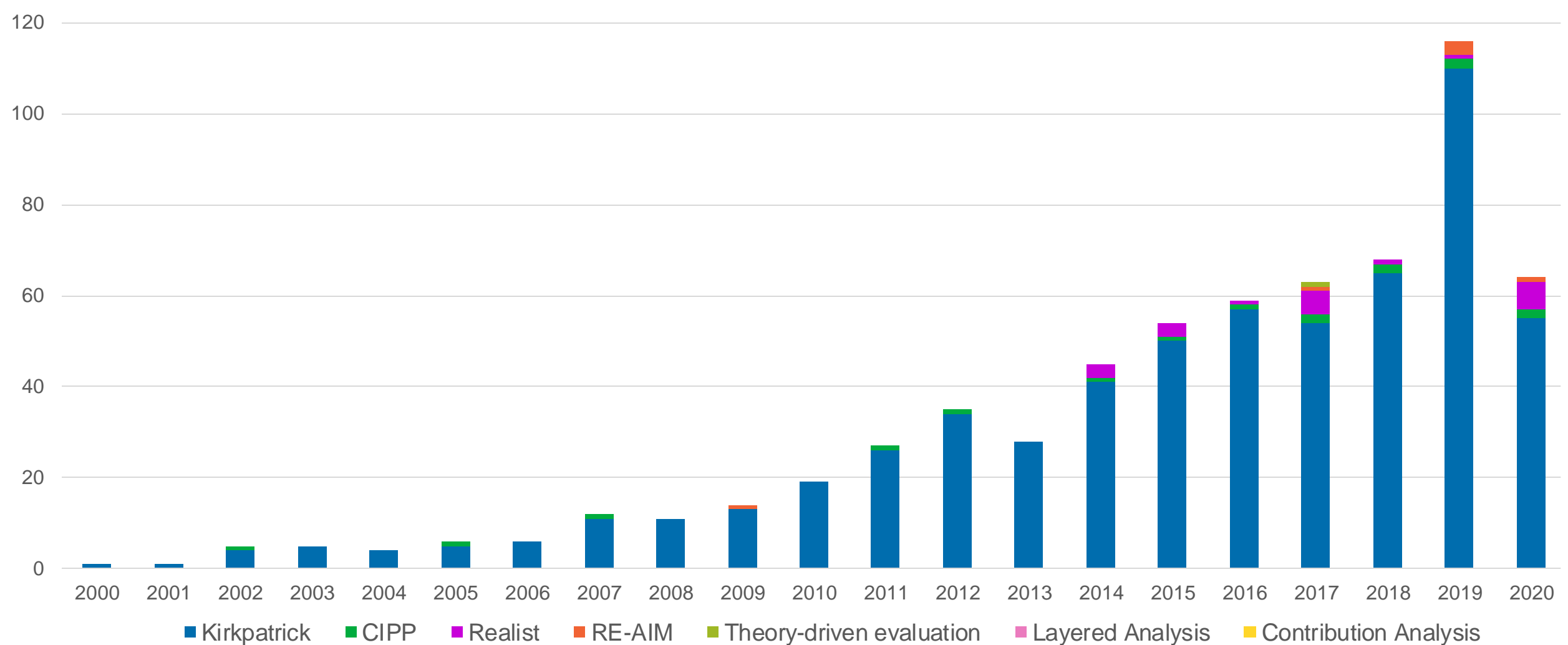


THE UNIVERSITY OF  
MELBOURNE

# What We Found

# Use of Evaluation Models in HPE

## Kirkpatrick Model compared to Programme Evaluation Models



# The Kirkpatrick Model – Summary of Critiques



- ✓ It is outcomes focused failing to consider how and why outcomes occur
- ✓ The outcomes of focus are short-term, quantifiable outcomes
- ✓ It assume positive causal linkages between the levels
- ✓ Unintended outcomes are not considered

“

*“Was outcome X achieved as intended, or not?” rather than “What were the outcomes of this intervention?” A clinical parallel would be a clinical trial that measured only the intended effects of a new drug and not its side-effects (Yardley et al., 2010, p. 100).*

”

# Use of the Kirkpatrick Model in HPE



1. Seldom reported that it is used in the development of programmes
2. Inconsistent and inaccurate terminology
3. Medical Education Research Study Quality Instrument (MERSQI)
4. Best Evidence Medical Education (BEME)
5. World Health Organisation (WHO)



THE UNIVERSITY OF  
MELBOURNE

# Moving Forward

# Moving Beyond Outcome Evaluation – Towards Programme Evaluation



Health professions education interventions are complex,  
and they do not occur in a bubble

OUTCOME EVALUATION



PROGRAMME EVALUATION

# Improving Evaluation

## What Should We Move Towards?



OUTCOME EVALUATION



PROGRAMME EVALUATION

PREPARATION AND PLANNING

Discussing and recording how and why the intervention is expected to work

Including ways to capture unintended outcomes

Evaluating at multiple time points including during interventions and longer term follow up

Multiple methods – including qualitative methods

Multiple perspectives – learners, colleagues, patients etc.

Consideration of various stakeholders and their differing goals

# Considerations: How to Choose an Evaluation Approach



1. What is the purpose of your evaluation?

- Scholarly Teaching
- Scholarly Innovation
- Research
- Accreditation

2. What resources do you have?

- Time
- People (Expertise)
- Money
- Power

3. Are there competing demands?

Accreditation vs. Improvement

# Final Thoughts



Evaluation is more than just measuring outcomes; it is about understanding:

- Why our educational interventions are intended to work,
- If they are being delivered as intended, and
- The broader context within which our educational interventions occur

All of which help us better understand how and why the outcomes are occurring.

Hence, if using Kirkpatrick Model – incorporate as part of a broader evaluation strategy and look beyond predetermined outcomes

# Thank you!



THE UNIVERSITY OF  

---

MELBOURNE

# References



Most of the references are in the 'Is measuring outcomes enough' paper, but I referred to a couple of others that I list here:

1. Choi T, Sarkar M, Bonham M, Brock T, Brooks IA, Diug B, Ilic D, Kumar A, Lau WM, Lindley J, Morphet J, Simmons M, Volders E, White PJ, Wright C, Palermo C. Using contribution analysis to evaluate health professions and health sciences programs. *Front Med (Lausanne)*. 2023;10:1146832. doi: 10.3389/fmed.2023.1146832.
2. Varpio L, Sherbino J. Demonstrating causality, bestowing honours, and contributing to the arms race: Threats to the sustainability of HPE research. *Med Educ*. 2024;58(1):157-163. doi: 10.1111/medu.15148.
3. Onyura B, Fisher AJ, Wu Q, Rajkumar S, Chapagain S, Nassuna J, Rojas D, Nirula L. To prove or improve? Examining how paradoxical tensions shape evaluation practices in accreditation contexts. *Med Educ*. 2024;58(3):354-362. doi: 10.1111/medu.15218.