REVIEW ARTICLE

Proposal of a linear rather than hierarchical evaluation of educational initiatives: the 7Is framework

Damian Roland[1,2]*

[1]Sapphire Group, Department of Health Sciences, University of Leicester; [2]Paediatric Emergency Medicine Leicester Academic Group, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

Abstract

Extensive resources are expended attempting to change clinical practice; however, determining the effects of these interventions can be challenging. Traditionally, frameworks to examine the impact of educational interventions have been hierarchical in their approach. In this article, existing frameworks to examine medical education initiatives are reviewed and a novel '7Is framework' discussed. This framework contains seven linearly sequenced domains: interaction, interface, instruction, ideation, integration, implementation, and improvement. The 7Is framework enables the conceptualization of the various effects of an intervention, promoting the development of a set of valid and specific outcome measures, ultimately leading to more robust evaluation.

Key Words: *Medical education; Health expenditures; Health resources; Outcome assessment*

INTRODUCTION

Medical professionals are always learning, whether in a formal sense via continuing medical education (CME) or as a result of their regular interaction with clinical cases [1-3]. The variable of ongoing informal learning can be a confounding factor that complicates determining the effectiveness of CME because it is difficult to control for in statistical analysis. Despite this, frameworks through which to explore the impact of educational interventions exist and are consistently being developed and refined. In this review article, I would like to propose a model of learning evaluation named the '7Is framework' based on a new paradigm of evaluation using a linear approach rather than a vertical or hierarchical approach. This framework was developed on the basis of the recommendations identified from a literature review. Using the new paradigm to better understand the learning process and to inform

educational interventions may enable better management of patients by physicians.

MODELS OF LEARNING EVALUATION

A frequently cited model of learning evaluation is that developed by Kirkpatrick [4-6]. The model, a four-stage approach to evaluation, was originally developed for business training directors. It comprises a hierarchy described in Table 1, starting with 'reaction' and progressing through 'learning', 'behaviour', and 'results'. Kirkpatrick himself did not originally use the term 'levels' although links between levels are supported by Cook and West [7] as a means to understanding the bridge between the original intervention and the overall outcome (Fig. 1).

The use of the levels concept implies that the Kirkpatrick framework is about a product, that is, the aspect of a particular learning outcome of greatest interest, rather than a process, which would be the quality of the instruction needed to achieve that outcome [8]. Conceptual frameworks are often built on previous theories and revisions of original models. One of the most commonly cited variations of the Kirkpatrick approach

Table 1. The original domains of Kirkpatrick

Level	Domain	Detail
1	Reaction	How well did the participants like the training?
2	Learning	What facts and knowledge were gained from the training?
3	Behaviour	Was the learning from the training utilised in the workplace?
4	Results	Did the training produce the overall intended benefits to the organisation?

Table 2. Modification of Kirkpatrick's domains by Barr et al.

Level	Kirkpatrick's domain	Barr's modification
1	Reaction	No change
2	Learning	2a: Modification of attitudes/perceptions 2b: Acquisition of knowledge/skills
3	Behaviour	No change
4	Results	4a: Change in organisational practice 4b: Benefits to patients/clients

Adapted from Barr et al. Evaluations of interprofessional education: a United Kingdom review for health and social care [Internet]. London: Centre for the Advancement of Interprofessional Education; 2000 [cited 2015 Apr 24]. Available from: http://caipe.org.uk/silo/files/evaluations-of-interprofessional-education.pdf [9].
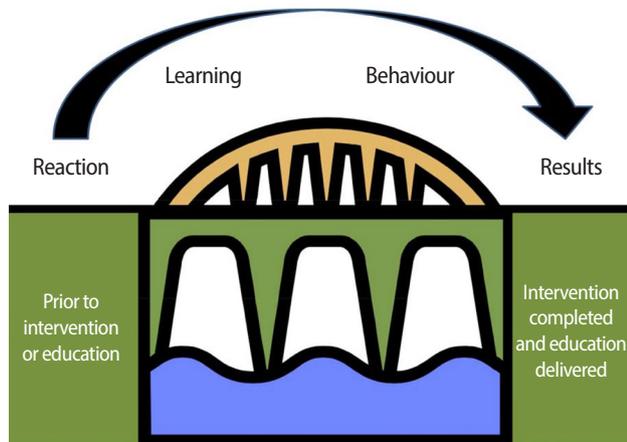

Fig. 1. The bridge of levels from reaction to results. This shows the process of understanding as the 'bridge' between the original intervention and the overall outcome [7]. Diagram was drawn by Damian Roland based on licence-free clip art.

was suggested by Barr et al. [9]. In a review of inter-professional education, subtypes of levels 2 and 4 of the Kirkpatrick framework were added (Table 2). This approach added more detail to the levels and provided researchers with a focused area of study. Although this approach is pragmatic, it has not been validated by either the original paper or subsequent work. The Best Evidence in Medical Education group [10,11] follows this framework; however, no formal study has compared a traditional Kirkpatrick level 2 with Barr's version.

Another approach has been to focus purely on the learner, that is, the health care professionals themselves. Acknowledging Kirkpatrick as a source, Belfield et al. [12] proposed a five-level hierarchy: healthcare outcomes; healthcare professionals' behaviour, performance or practice; learning or knowledge; reaction or satisfaction of participants; and participation or completion. Belfield et al. [12] highlighted the difficulty of outcomes-based research in a medical education setting, in particular, the problem that patient outcomes may only become apparent over a protracted period of time due to the time needed for the learner to acquire and implement new skills. They emphasised that levels of effectiveness, the outcome measures used, should be clearly reported and justified

in study reports. This is of relevance to the utilization of Kirkpatrick, as much terminology in medical education such as 'appraisal,' 'assessment,' 'evaluation,' and 'competency' may carry different meanings across disciplines and specialties [12]. This lack of terminology standardization may mean that despite the fact that an evaluative framework such as Kirkpatrick's is in place, the measures or metrics used to assess outcomes in each of the domains may not be reliable between, and even within, studies.

A model based on Kirkpatrick's framework but also incorporating theories specific to medical education has been devised by Moore et al. [13] (Table 3). They clearly point out that their approach is a conceptual framework rather than a fully validated model. The incorporation of Miller's pyramid [14], a theoretical approach to learning in which the learners move from 'knows' and 'knows how' to 'shows how' and 'does' is synergistic with the learning and performance domains. Similar to the approach adopted by Barr et al. [9], the essential Kirkpatrick framework is not changed but the detail provided at each level is greatly expanded. Again, although not directly validated, this approach has received endorsement by the Royal College of Physicians and Surgeons of Canada [15,16], where the knowledge and behaviour domains are separated into self-reported and observed elements.

Ultimately, in all the models discussed above, the hierarchal nature of the Kirkpatrick levels is present, and the flow from reaction to the intervention, followed by learning and implementation into behaviour and then a translation into patient benefit is still present. Given the specific lack of validation of these models, it is difficult to know whether they are enhanced versions of the Kirkpatrick model with tighter definitions of the domains or novel concepts in their own right. In an attempt to avoid the hierarchical structure inherent in Kirkpatrick's framework, Hakkennes and Green [17] used five categories in no particular order to measure three separate outcome domains (Table 4). The Hakkennes model provides clear met-

Table 3. Moore's expanded outcomes framework

Original (expanded) CME framework	Miller's framework	Description	Source of data
Participation (level 1)		The number of physicians and others who participated in the CME activity	Attendance records
Satisfaction (level 2)		The degree to which the expectations of the participants about the setting and delivery of the CME activity were met	Questionnaires completed by attendees after a CME activity
Learning (declarative knowledge level 3a)	Knows	The degree to which participants state what the CME activity intended them to know	Objective: pre- and post-tests of knowledge Subjective: self-report of knowledge gain
Learning (procedural knowledge level 3b)	Knows how	The degree to which participants state how to do what the CME activity intended them to know how to do	Objective: pre- and post-tests of knowledge Subjective: self-report of knowledge gain
Learning (competence level 4)	Shows how	The degree to which participants show in an educational setting how to do what the CME activity intended them to be able to do	Objective: observation in educational setting Subjective: self-report of competence; intention to change
Performance (level 5)	Does	The degree to which participants do what the CME activity intended them to be able to do in their practices	Objective: observation of performance in patient care setting; patient charts; administrative databases Subjective: self-report of performance
Patient health (level 6)		The degree to which the health status of patients improves due to changes in the practice behaviour of participants	Objective: health status measures recorded in patient charts or administrative databases Subjective: patient self-report of health status
Community health (level 7)		The degree to which the health status of a community of patients changes due to changes in the practice behaviour of participants	Objective: epidemiological data and reports Subjective: community self-report

Reproduced from Moore et al. J Contin Educ Health Prof. 2009;29:1-15, with permission of Wiley [13].
CME, continuing medical education.

Table 4. Hakkennes' domains of evaluation

Domains	Categories
Patient	Measurements of actual change in health status of the patient, i.e., pain, depression, mortality, and quality of life (A1) Surrogate measures of A1, i.e., patient compliance, length of stay, and patient attitudes (A2)
Health practitioner	Measurements of actual change in health practice, i.e., compliance with guidelines, changes in prescribing rates (B1) Surrogate measures of B1, such as health practitioner knowledge and attitudes (B2)
Organisational or process level	Measurements of change in the health system (i.e., waiting lists), change in policy, costs, and usability and/or extent of the intervention (C)

Data from Hakkennes S, Green S. Implement Sci. 2006;1:29 [17].

rics, with examples, in which to measure the effectiveness of the intervention. Uniquely to this model, the concept of actual versus surrogate measures is utilized. The Kirkpatrick model allows the end-user to decide the outcome measure at each particular level. Hakkennes and Green [17] provide a framework for strengthening validity by differentiating between actual or gold-standard outcomes and those that are perceived or potentially confounded.

Increasing the level of detail in each domain or level raises a challenge to the epistemological approach used by developers of outcomes-based systems. In the business-oriented model put forward by Kirkpatrick, a clear connection between learning and outcome is inherent in the structure, despite the absence of objective descriptors of each level. Conversely, the increasing detail provided by Hakkennes and Green [17] implies a more dissociated structure where information must be col-

lated and analysed at each level. These approaches represent complex methodological debates and form part of a likely spectrum in paradigms for evaluating medical interventions. On one side is a discrete hierarchal approach, as demonstrated by Lemmens et al. [18], who describe an evaluation model for determining the effectiveness of disease management programs. This model has a sequence of knowledge acquisition, leading to behaviour intention and then change—ultimately leading to improved clinical outcomes. A case study supports their proposition with neither Kirkpatrick nor Barr cited as sources or alternatives to this evaluative approach. On the other side of the spectrum is a less prescribed approach espoused by Yardley and Dornan [19] in which the researcher assesses the complexity of the outcomes and applies an appropriate framework dependent on the processes that may occur from intervention to outcome (Fig. 2).
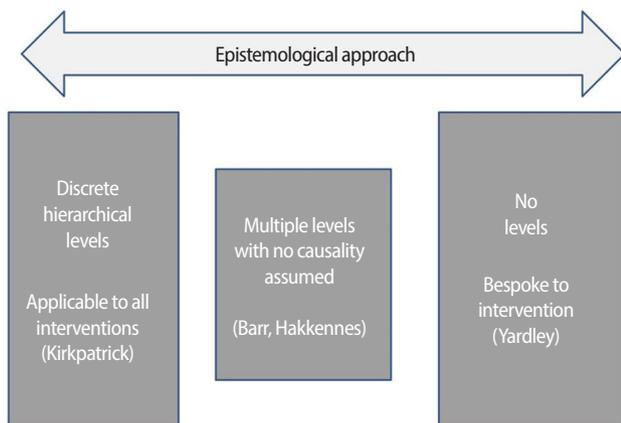
Fig. 2. The range of epistemological approaches to evaluation. The researcher assesses the complexity of the outcomes and applies an appropriate framework dependent on the processes that may occur from intervention to outcome [19]. Diagram was drawn by Damian Roland.
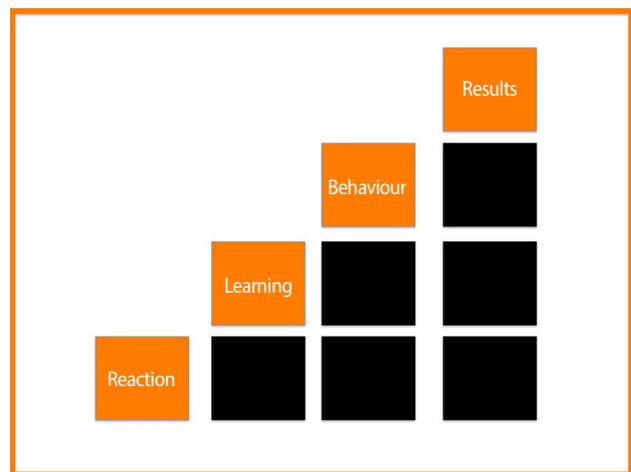

Fig. 3. The hierarchal nature of the Kirkpatrick evaluation framework where level-four outcomes are of greater importance than those of level one [21]. Diagram was drawn by Damian Roland.

However as Yardley and Dornan [19] argues, it is important to ask not, "Did a specific outcome occur?" but "What were the outcomes of this intervention?" as the latter may aid in the determination of the underlying cause of the outcomes. In this respect, the Kirkpatrick level-based approach does aid reasons for success or failure of a particular project to be elucidated. For example, in a large study of the effect of an evidence-based care model to aid management of the febrile child, a number of quality measures, including laboratory testing, admission rates, length of stay and costs, demonstrated significant improvements [20]. However, as there was no evaluation of individual performance, it is difficult to know whether it was the processes in place that resulted in change or learning at an individual level. This may have consequences for implementation at other sites and makes it difficult to determine the most influential shifts in practice that will require validation in future studies.

PROBLEMS WITH KIRKPATRICK'S DESIGN

Commentaries on concerns with Kirkpatrick's design predominantly relate to industries outside of health care. Bates [21] raised three concerns with the Kirkpatrick approach: It is an incomplete model, there is an unproven assumption of causality, and the higher outcome levels do not necessarily imply higher levels of information. Despite Kirkpatrick not initially using the term "level" himself, his design implicitly suggests that level-four outcomes are of greater importance than those of level one (Fig. 3). The hierarchy of organizational outcomes, or return on investment, having more 'value' than initial participant reactions has been challenged in the business community for some time [22]. Despite Kirkpatrick believing, in

respect of evaluating level two (knowledge change), that "No change in behaviour can be expected unless one or more of these learning objectives have been accomplished" [6], research in the business sector has failed to confirm the hierarchical relationship the Kirkpatrick model suggests [23,24]. These concerns surrounding the Kirkpatrick model have introduced uncertainty in the medical education academic community. In 2012, Yardley and Dornan published a systematic analysis of the use of Kirkpatrick in the context of medical education. The work analysed articles utilising a Kirkpatrick evaluative approach and also used a case-control journal review to determine methodologies the Kirkpatrick system may miss [19]. They also identified a collection of commentaries highlighting how the very hierarchical nature of Kirkpatrick could bias the outcomes the evaluation was aiming to examine. Alliger et al. [23] and Abernathy [25] have previously argued from a business perspective that the very notion of levels of evaluation influences the academic approach the evaluator may take. Yardley and Dornan [19] suggested that medical education is a much more complex system than business, as the stakeholders include patients and doctors but also the patients' families, the health systems and the communities around the health systems.

A NEW CONCEPTUAL MODEL FOR EVALUATION

The model put forward by Kirkpatrick has not gained universal traction throughout the academic medical educational community. Yardley and Dornan [19] felt that following Kirkpatrick blindly would be akin to "performing research on a clinical drug and not assessing the potential side effects." Those utilising the Kirkpatrick Framework in the healthcare setting

are assuming a linear sequencing of acquisition and subsequent use of knowledge, skills and attitudes. Medical interventions do not have a binary transactional nature and have unpredictable outcomes. In the same way, educational or other practice-changing interventions are also heterogeneous: Learners modifies some aspects of their behaviour but not others in response to the intervention. A framework acknowledging this effect would be beneficial. Furthermore, the continuing review of evidence, rather than the improvement of it, has been persistent in medical education evaluative literature. Buckley et al. [26] felt a level-based system of evaluation was neither intuitive nor theoretically sound, citing the need for a continuum of change rather than discrete levels. Despite this systematic review, many continue to endorse the Kirkpatrick framework and researchers continue to build upon the Kirkpatrick model [27]. Models that aim to enhance Kirkpatrick add to the levels, rather than redefining the concept entirely. The linear nature of individuals' learning, leading to behavioural change and finally patient outcomes is implicit in all such systems.

DERIVING A NEW MODEL

As demonstrated by the variety of models discussed and described, many approaches to evaluation exist. It is likely that no one system is best for all situations; rather, the best approach varies according to factors such as the intent of the evaluation, the nature of key stakeholders, and available resources [28]. It is for this reason that evaluative frameworks are created. A framework has been defined as a set of steps to operationalise a particular evaluation model or approach. However, often the terms 'approach', 'framework', and 'model' are used interchangeably. The use of frameworks as conceptual models is a pragmatic way of simplifying potential complex methodological approaches. However, there is a risk that the designs in themselves lack epistemological rigor. It is argued from an information systems perspective that this could lead to problems either regarding applicability of outcomes in certain contexts or the actual feasibility in delivering certain evaluations [29].

In order to create a valid framework, the literature on the work of Kirkpatrick and others was utilised to derive core needs as follows: First, measures should be obtained from a variety of sources: Kirkpatrick is an incomplete model and a new framework needs to allow a broad range of evaluative methodologies. Second, the domains (or measures making up those domains) are independent of each other: The assumption of causality is a core feature of the original Kirkpatrick model, and this assumption has not been proven. Third, the framework does not explicitly suggest a hierarchy: A hierarchical approach gives an undue weight to certain parts of the framework. In clinical medicine, this may well be justified – the primary clini-

cal outcome to the patient being a clear and transparent end point. In education, however, although patient benefit is of importance, not all educational intervention is necessarily related to patient experience or outcomes. Given the significant confounding influences on determining how participants may learn and apply their knowledge, concentrating on patient outcome may leave out information that determines the effectiveness of the intervention. The framework brings together a chain of methods which in totality deliver an overall effect rather than relying on just one component. Fourth, patient outcomes are considered: Although patient outcomes may not be the highest rung of outcome, they are a vital part of the evaluation process and must have internal validity with respect to the intervention used; that is, the intervention would plausibly result in the outcome being measured. Fifth, engagement with the intervention itself is considered: Previous models have only measured satisfaction with the model, not the actual uptake or utilisation itself.

The underlying tenet of the proposed structure was to demonstrate a process of evaluation. The model aimed to include some of the missing components of Kirkpatrick and other frameworks in creating domains encapsulating the interaction with the intervention, the intervention itself and the outcomes of the intervention. The purpose of including the intervention itself is to ingrain the principle that the evaluation requires information taken before, during and after the intervention.

The basic constructs of assessing learning and behaviour remain as in Kirkpatrick but are considered in parallel rather than on top of each other. The model proposed here uses the concepts of 'ideation' (what you think you have learned) and 'integration' (what you have shown you have learned) to ensure these two evaluative approaches are included. This brings
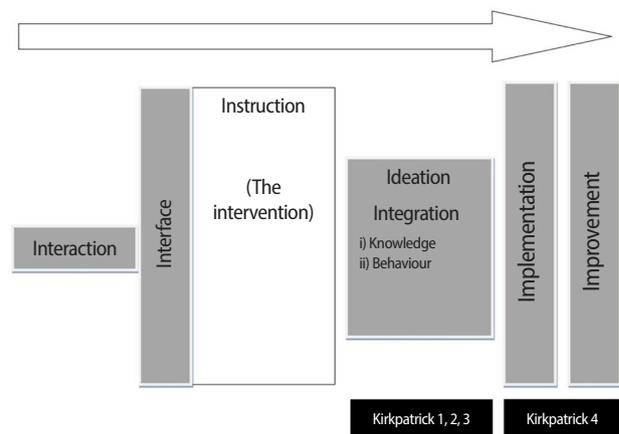

Fig. 4. A theoretical schema for evaluating outcomes of practice-changing interventions – The 7Is framework. Diagram was drawn by Damian Roland.

Table 5. Description of the 7Is framework domain headings

7I Domain	Summary
Interaction	The degree to which participants engage with and are satisfied with the instruction
Interface	The degree to which participants are able to access the instruction
Instruction	The details of the intervention itself
Ideation	The perception of improvement following the instruction
Integration	The change, in both knowledge and behaviours, as a result of the instruction
Implementation	Whether change across individuals i.e., departments or organisations following the instruction has been demonstrated
Improvement	Whether the instruction has resulted in improvements in patient care and experience

Table 6. Steps to validate the original domains of the 7Is framework

Domain	Study area
Interaction	A review of randomized control trials in medical education specifically looking at the concept of 'intention to learn' would further validate this domain. A before and after study to demonstrate effectiveness by ensuring post-learning testing was undertaken by those not completing the intervention should be performed. Although an enforced post-learning element would introduce a level of bias, differences in the outcomes would suggest interaction analysis must be a fundamental part of evaluation.
Interface	The development of software, especially in light of e-learning studies, to examine the precise nature of how participants are able to, or are blocked from, accessing all modalities of a teaching package, would allow richer data in this domain to be examined.
Instruction	The development of taxonomy of medical education, and practice changing intervention, studies to allow valid comparisons between studies via the 7Is Framework.
Ideation	Further qualitative research into exploring junior doctors' understanding of competence and confidence and safety is required. It would be beneficial to repeat the meta-planning exercise on a different clinical issue (i.e., not in the field of pediatrics). If individual discriminatory concepts making up each of the terms could be validated, this would allow the creation of a questionnaire to assess and measure initial ideation. This would then allow a more detailed exploration of the proposed matrix linking the terms together and assess its practical use in a patient safety context.
Integration (knowledge)	If it is a case of premature ventricular contraction for example, a selection of purposefully designed disease for determining gold standard quality should be collated. A qualitative study in conjunction with this benchmarking exercise should take place to capture participants' decision making processes. This process would aim to improve the assessment of disease but may also guide future telemedicine studies to create minimum quality standards.
Integration (behaviour)	An observational study comparing case note review with observed interaction with patients would further validate the Rolma matrix.
Implementation and improvement	The results from the given study can inform effect sizes and a power calculation needed for a randomized control trial of the intervention. This would allow for an understanding of the relationship between implementation and improvement to be described.

together the learning and behaviour sections, not as a hierarchy but as a common domain. In recognition of the complexity of healthcare, the concept of 'results,' which correspond to level four of Kirkpatrick's framework, becomes 'implementation' and 'improvement'. This enables patient benefit to be encapsulated by direct clinical effects but also by experience outcome measures as well. The conceptual framework, provisionally entitled 'the 7Is framework', is shown in Fig. 4 and Table 5. The alliteration is coincidental but may aid memorability of the framework.

This framework has been recently utilised in a study examining outcomes of an e-learning package [30]. An audio-visual intervention in paediatric fever was designed, delivered and tested against the new system. Interaction with the intervention was variable; only 28.7% completed the post-learning section, and issues were identified with accessing the video cases. Although measures of ideation significantly increased and there was a trend towards behaviour change, full implementation of

the guidance did not occur and overall admission rates increased. This work demonstrated that the 7Is framework allows the various effects of an intervention to be conceptualised, promoting the development of a set of valid and specific outcome measures, ultimately leading to more robust evaluation. The next steps are to validate the original domains further (Table 6).

CONCLUSION

Traditionally, evaluative frameworks have concentrated on knowledge, behaviour or system change as discreet entities. The 7Is framework may contribute a new paradigm to the literature on evaluating practice changing interventions: the recognition of the importance of interacting and interfacing with the intervention and removing a hierarchal structure. The 7Is framework can be used to improve understanding of why interventions are effective and will hopefully promote the development of improved outcome measures.

ORCID: Damian Roland: http://orcid.org/0000-0001-9334-5144

CONFLICT OF INTEREST

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Audio recording of abstract.

REFERENCES

1. Nylenna M, Aasland OG. Doctors' learning habits: CME activities among Norwegian physicians over the last decade. BMC Med Educ. 2007;7:10. http://dx.doi.org/10.1186/1472-6920-7-10

2. Roberts T. Learning responsibility?: exploring doctors' transitions to new levels of performance. ESRC End of Project Report. 2009.

3. Graham D, Thomson A, Gregory S. Liberating learning [Internet]. Milton Keynes: Norfolk House East; 2009 [cited 2015 Apr 24]. Available from: http://www.nact.org.uk/getfile/2173

4. Kirkpatrick D. Evaluation. In: Craig RL, Bittel LR, editors. Training and development handbook. American Society for Training and Development. New York: McGraw-Hill; 1967.

5. Kirkpatrick D. Evaluation of training. In: Craig RL, editor. Training and development handbook: a guide to human resource development. New York: McGraw-Hill; 1976. p.317.

6. Kirkpatrick D, Kirkpatrick J. Evaluating training programs: the four levels. 3rd ed. San Francisco: Berrett-Koehler Publishers Inc.; 2006.

7. Cook DA, West CP. Perspective: reconsidering the focus on "outcomes research" in medical education: a cautionary note. Acad Med. 2013;88:162-167. http://dx.doi.org/10.1097/ACM.0b013e31827c3d78

8. Morrison J. ABC of learning and teaching in medicine: evaluation. BMJ. 2003;326:385-387. http://dx.doi.org/10.1136/bmj.326.7385.385

9. Barr H, Freeth D, Hammick M, Koppel I, Reeves S. Evaluations of interprofessional education: a United Kingdom review for health and social care [Internet]. London: Centre for the Advancement of Interprofessional Education; 2000 [cited 2015 Apr 24]. Available from: http://caipe.org.uk/silo/files/evaluations-of-interprofessional-education.pdf

10. Association for Medical Education in Europe. Best Evidence Medical Education [Internet]. Dundee: Association for Medical Education in Europe; 2013 [cited 2015 Apr 24]. Available from: http://www.bemecollaboration.org/

11. Association for Medical Education in Europe. BEME coding sheet [Internet]. Dundee: Association for Medical Education in Europe; 2005 [cited 2015 Apr 24]. Available from: http://www.bemecollaboration.org/downloads/749/beme4_appx1.pdf

12. Belfield C, Thomas H, Bullock A, Eynon R, Wall D. Measuring effectiveness for best evidence medical education: a discussion. Med Teach. 2001;23:164-170. http://dx.doi.org/10.1080/0142150020031084

13. Moore DE Jr, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. J Contin Educ Health Prof. 2009;29:1-15. http://dx.doi.org/10.1002/chp.20001

14. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 Suppl):S63-S67.

15. Horsley T, Grimshaw J, Campbell C. Maintaining the competence of Europe's workforce. BMJ. 2010;341:c4687. http://dx.doi.org/10.1136/bmj.c4687

16. Campbell C, Silver I, Sherbino J, Cate OT, Holmboe ES. Competency-based continuing professional development. Med Teach. 2010;32:657-662. http://dx.doi.org/10.3109/0142159X.2010.500708

17. Hakkennes S, Green S. Measures for assessing practice change in medical practitioners. Implement Sci. 2006;1:29. http://dx.doi.org/10.1186/1748-5908-1-29

18. Lemmens KM, Nieboer AP, van Schayck CP, Asin JD, Huijsman R. A model to evaluate quality and effectiveness of disease management. Qual Saf Health Care. 2008;17:447-453. http://dx.doi.org/10.1136/qshc.2006.021865

19. Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence'. Med Educ. 2012;46:97-106. http://dx.doi.org/10.1111/j.1365-2923.2011.04076.x

20. Byington CL, Reynolds CC, Korgenski K, Sheng X, Valentine KJ, Nelson RE, Daly JA, Osguthorpe RJ, James B, Savitz L, Pavia AT, Clark EB. Costs and infant outcomes after implementation of a care process model for febrile infants. Pediatrics. 2012;130:e16-e24. http://dx.doi.org/10.1542/peds.2012-0127

21. Bates R. A critical analysis of evaluation practice: the kirkpatrick model and the principle of beneficence. Eval Program Plann. 2004;27:341-347. http://dx.doi.org/10.1016/j.evalprogplan.2004.04.011

22. Holton EF. The flawed four level evaluation model. Human Resour Dev Q. 1996;7:5-21. http://dx.doi.org/10.1002/hrdq.3920070103

23. Alliger GM, Tannenbaum SI, Bennett W, Trave H, Shotland A. A meta-analysis of the relations among training criteria. Person Psychol. 1997;50:341-358. http://dx.doi.org/10.1111/j.1744-6570.1997.tb00911.x

24. Alliger GM, Janak EA. Kirkpatricks's level of training criteria: thirty years later. Person Psychol. 1989;42:331-342. http://dx.doi.org/10.1111/j.1744-6570.1989.tb00661.x

25. Abernathy D. Thinking outside the evaluation box. Train Dev. 1999;53:19-23.

26. Buckley LL, Goering P, Parikh SV, Butterill D, Foo EK. Applying a 'stages of change' model to enhance a traditional evaluation of a research transfer course. J Eval Clin Pract. 2003;9:385-390. http://dx.doi.org/10.1046/j.1365-2753.2003.00407.x

27. Tian J, Atkinson N, Portnoy B, Gold R. A systematic review of evaluation in formal continuing medical education. J Contin Educ Health Prof. 2007;27:16-27. http://dx.doi.org/10.1002/chp.89

28. Kahen B. Excerpts from review of evaluation frameworks. Regina: Saskatchewan Ministry of Education; 2008.

29. Recker J. Conceptual model evaluation: towards more paradigmatic rigor. In: Castro J, Teniente E, editors. Proceedings of the CAISE'05 Workshops; 2005 Jun 13-17; Porto, Portugal. Porto: Faculdade de Engenharia da Universidade do Porto; 2005.

30. Roland D, Charadva C, Coats T, Matheson D. Determining the effectiveness of educational interventions in paediatric emergency care. Emerg Med J. 2014;31:787-788. http://dx.doi.org/10.1136/emermed-2014-204221.25