



Review

Testing the participatory education evaluation concept in a national context

Ville Pietiläinen*

The Finnish Education Evaluation Council, University of Jyväskylä, Finland¹

ARTICLE INFO

Article history:

Received 19 October 2011
 Received in revised form 10 February 2012
 Accepted 28 February 2012
 Available online 23 March 2012

Keywords:

Evaluation utilisation
 Evaluation methods
 Participatory evaluation
 National education context

ABSTRACT

The article focuses on the realisation of participatory evaluation (PE) in national educational evaluation activity. The realisation of PE is examined by adapting the [Daigneault and Jacob model \(2009\)](#); originally [Cousins & Whitmore, 1998](#)) to five national-level educational evaluations carried out in Finland. According to the chosen frame of reference, self-evaluation regarding the realisation of these five evaluations as well as the analysis of reports that are yielded from the evaluations is used as the research method. PE was realised at a general level in the Finnish national level educational evaluations, but during the process there were considerable differences between the separate dimensions of PE. The article ends in two conclusions, which are presented as alternative scenarios.

© 2012 Elsevier Ltd. All rights reserved.

Contents

Theoretical framework for PE	10
The national educational evaluation system in Finland	11
Realisation of PE in Finnish national level educational evaluations	11
Discussion	12
Acknowledgements	13
References	13

The participation of nonevaluative stakeholders has risen to a significant position in the realisation of national educational evaluations (e.g. [Mathison, 2005](#)). The aim of participatory evaluation (PE) is twofold: on the one hand to increase the usefulness of evaluation from the viewpoints of nonevaluative stakeholders, and on the other to seek fairness in evaluation. In this thinking it is believed that for the separate parties the usefulness of evaluations increases when the parties participate in carrying out the separate stages of evaluation. Inherent to participation is a support of fairness – a notion in accordance with democracy and with which people can affect the society in which they live ([Wallace & Alkin, 2008](#), p. 192). During the last few years, PE has been examined from several different points of view, such as the researcher's interest as a concept of participation and its significance ([King, 2007](#)), the separate forms of PE ([Cousins &](#)

[Whitmore, 1998](#)), the participation process ([Weaver & Cousins, 2006](#)), the dynamics of the nonevaluative stakeholder groups ([Bryson, Patton, & Bowman, 2011](#)) and the better utilisation of evaluation results with the help of the participation process ([Smits & Champagne, 2008](#)).

In spite of the interest, conceptualising PE and adapting it to practical evaluation activity have proved to be challenging. The critical questions have been, among others, which nonevaluative stakeholder groups should participate in evaluation, in which stages of evaluation should nonevaluative stakeholder groups participate, what are their roles and how is profound participation sought (e.g. [King & Ehlert, 2008](#)). Correspondingly one can ask from the viewpoint of value choices – whose values are focused on in participation – which weight values the different target groups receive and how realistic and possible it is for different target groups to bring their own values to the discussion.

This article examines the realisation of PE in national educational evaluation activity. The examination focuses on a Finnish educational evaluation activity in which PE is strongly emphasised. In the [Government Decree on Evaluation \(1061/2009\)](#) the various needs of the separate parties in the evaluation are clearly up for discussion. According to the second section of the

* Correspondence address: The Finnish Education Evaluation Council, University of Jyväskylä, P.O. Box 35, 40014 University of Jyväskylä, Finland.
 Tel.: +358 14 260 4686; fax: +358 14 260 4691; gsm: +358 400 248 095.
 E-mail address: ville.pietilainen@eval.jyu.fi.

¹ <http://www.edev.fi>.

Degree, the objective of national educational evaluations is to serve the institutes of political and local decision-making and education. Consequently, the participation of all the above-mentioned parties in the national educational evaluations can be considered justified.

In spite of the evaluation policy, which emphasises the different parties' needs, an analysis of the separate parties' participation in national educational evaluations has not yet been carried out in Finland. In order to examine the success of the evaluation policy, information about the realisation of PE in practical evaluation activity is needed. On the other hand, the critical examination of the evaluation policy, which emphasises participation, is also considered significant in the article because there are not always positive consequences from participation. In several cases the participation of too many nonevaluative stakeholders or too much independence of these stakeholders has led to anarchy and endangerment of the evaluation task (Gastic, Irby, & Zdanis, 2008; King & Ehlert, 2008; Patton, 1997). Because there can be positive or negative consequences from participation, information is needed regarding what level of participation can be considered desirable from the viewpoint of the realisation of the evaluation task.

The two needs described in the previous piece embody the purpose of this article, in which conclusions are made from the realisation of PE in national educational evaluation activity and from the target participation level. This article concentrates on the question.

Can the PE level represented by national educational evaluation activity be considered desirable from the viewpoint of the evaluation task realisation?

At the beginning of the article, a model of the constitutive dimensions of PE will be constructed. The model by Daigneault and Jacob model (2009; originally Cousins & Whitmore, 1998), considered the most thorough for the conceptualisation and measurement of PE, is used as a theoretical frame of reference to examine the way in which PE is realised in national educational evaluation activity. The five national evaluations of vocational training carried out by the Finnish Education Evaluation Council (FEEC) during the years 2008–2011 serve as the foundation for testing the model. FEEC is a central organisation in the carrying-out of national evaluations. At the end of the article, conclusions will be drawn regarding whether the level of PE represented by national educational evaluations can be considered desirable from the viewpoint of the evaluation task realisation.

The central target groups of the article are the persons who decide on the evaluation policy and carry out the national educational evaluations. The results of the article can be used in the development of both evaluation policy and concrete evaluation activity. To the researcher public the article yields an empiric test of a theoretical PE model from one point of view. To the participants of the evaluations, the article offers information about the logic of national evaluations and about the connection to the local activities.

Theoretical framework for PE

In Cousins and Whitmore's (1998) original model, PE is examined from the viewpoint of two traditions, the first of which emphasises the promotion of practical decision-making and problem-solving (P-PE) and the second emphasises social entitlement, authorisation and participation of weaker groups in the evaluation (T-PE). Taking both traditions into consideration, three central dimensions of PE are separated in the model: Control of the Evaluation Process, diversity of the interest groups and depth of participation (see also Cousins, Donohue, & Bloom, 1996; Cousins & Earl, 1992). According to Daigneault and Jacob (2009, p. 377), Cousins and Whitmore's (1998) model captures the central contents of PE, but its measurability is in need of further

development. Daigneault and Jacob (2009) have developed the central dimensions of PE naming and determining the measurability of the model as follows:

Extent of Involvement: Reflects the number of tasks of nonevaluative stakeholder groups at different stages of the evaluation process. The central stages of the evaluation process are evaluation questions and issues in the definition and methodological design, data collection and analysis, judgment and formulations of recommendations and report and dissemination of evaluation findings. The assumption of this dimension is that participation in evaluation can be improved by increasing the number of tasks for nonevaluative stakeholders at different stages of the evaluation process.

Diversity: Reflects the number of nonevaluative stakeholder types which are involved in evaluation. Nonevaluative stakeholders can be divided into four groups: policy-makers and decision-makers, implementers and deliverers, target populations (end-users of evaluation) and civil society and citizens (e.g. interest groups, unions, researchers). The assumption of this dimension is that evaluation participation can be improved by increasing the number of each type of nonevaluative stakeholder in the evaluation.

Control of the Evaluation Process: Reflects the participants' authority to make decisions or other resources in addition to the authority that participants can mobilise in an evaluation process. These other resources can include, for example, participants' methodological expertise, money or premises. The term 'participant' can encompass nonevaluative stakeholders that are neither evaluators nor sponsors of the evaluation. The assumption of this dimension is that the participation of evaluation can be improved by increasing the authority or other resources of participants.

The first two dimensions of PE are clearly quantitatively measured. As well, the third dimension of PE also contains qualitative properties because the control can receive several different forms at separate stages of the evaluation process. Daigneault and Jacob (2009, p. 343; see also Themessl-Huber & Grutsch, 2003) recommend a more intuitive and more flexible coding of the control dimension compared to the first two dimensions. According to them, the coding of the control dimension should be based on informed decision instead of a mechanical rule.

Some subjects of development can be distinguished in spite of the development of PE measurability from the model described above. According to Daigneault and Jacob (2009, p. 345), the model is in need of more validity testing and development of the scoring of different dimensions. In particular the neglect of participation intensity is emphasised as a deficiency. The participation is quite technically dealt with in the model by examining the number of nonevaluative stakeholders and their tasks during the evaluation process. Therefore participation intensity has been considered as a fourth dimension of PE as follows:

Volume of Interest: Reflects the amount and direction of the participants' interest in the evaluation process. In a technocratic evaluation, the exclusive interest in evaluation is held by the evaluators or the evaluation sponsors. In PE the participants have exclusive interest in the evaluation. The assumption of this dimension is that evaluation participation can be improved by increasing the participants' interest in evaluation. Volume of Interest contains qualitative properties similar to the third dimension of PE. Consequently, the coding of this dimension is based also on informed decision instead of a mechanical rule.

Next, the suitability and realisation of the PE model will be examined within the context of Finnish educational evaluation activity at the national level. A description of the Finnish educational evaluation system is created, after which the most suitable evaluations carried out by FEEC are chosen to be tested against the formed PE model.

The national educational evaluation system in Finland

The Finnish educational system and its evaluation went through a period of centralism that lasted until the 1990s. Centralism was considered an efficient method for the implementation of significant political education reforms and for securing educational equality. The educational evaluation system was based on inspections which were the responsibility of Provincial State Offices. In the 1990s the atmosphere changed dramatically. Centralism was abolished and the main responsibility for the organisation of educational services was transferred from the governmental to the municipal level. At the same time, the inspectorate system in education was abolished and the Provincial State Offices' tasks were gradually reduced (Eurydice, 2010a).

The external national evaluation has risen to an important position in the information control of education after the end of the inspection system. The present evaluation system for education applies to all forms of education and is based on its developments from the starting points to the 1998 legislation (e.g. *Basic Education Act, 628/1998*). According to the most recent *Government Decree on Evaluation (1061/2009)*, education and training providers have a statutory duty to evaluate their own activities and efficiency and to participate in external evaluations. The educational evaluation is implemented at a national level by the Ministry of Education and Culture, which produces a national external evaluation plan in cooperation with the evaluation actors. The focuses of the evaluation plan follow a valid government programme and the national action plan of education and research (Council of State, 2011; Ministry of Education and Culture, 2007, 2008).

At the national level, FEEC organises and coordinates evaluation processes in practice, excluding learning outcomes and higher education. FEEC serves as an expert body for educational evaluation in connection with the Ministry of Education and Culture. It is tasked with (a) assisting the Ministry of Education and Culture and supporting educational providers in matters concerning educational evaluation; (b) making an action plan for external educational evaluation in accordance with the guidelines and financial resources set by the Ministry of Education and Culture; (c) making proposals for the development of educational evaluation; and (d) participating in the international cooperation for evaluation activity (<http://www.edev.fi>; *Government Decree on Evaluation, 1061/2009*).

After the abolishment of the inspectorate system, a strong belief in PE has been dominant in the Finnish public. The interest groups have been comprehensively encouraged to join the planning and realisation of evaluations. As participation is emphasised in Finland, it is surmised that the PE model as described in this article is also being emphasised in Finnish educational evaluation. Next the validity of this assumption will be tested.

Realisation of PE in Finnish national level educational evaluations

The five national evaluations of vocational education and training (VET) carried out by FEEC were chosen to test the PE model between the years 2008 and 2010. The evaluations chosen were:

- Being well. The implementation, practices and development of student welfare in vocational upper secondary education (published 2010)
- We have evidence! How vocational skills demonstrations have been implemented in practice (published 2010)
- Social and communicative competencies in vocational basic education (published 2009)
- Working together round the table. The regional development impact of vocational education and training (published 2008)
- Apprenticeship training in vocational education (published 2008)

The uniform frame of reference for the evaluation, which enabled comparison between evaluations, was used as the foundation. All chosen evaluations were formed on the basis of the European Common Quality Assurance Framework. Other evaluations were excluded from the testing due to their lack of the uniform frame of reference, which was considered to be a central criterion for the reliable comparison of evaluations.

In evaluating the dimensions of PE, criteria and scaling were used, which is in accordance with presentation of Daigneault and Jacob (2009) and dimensions got a uniform weight value. Published reports from the evaluations (see <http://www.edev.fi>) and the self-evaluation of FEEC from the evaluations carried out were used as the foundation for testing verification.

Extent of Involvement: In this dimension a point for participation was given to the representatives of nonevaluative stakeholder types at the individual stage of evaluation (Table 1). Additional points for participation were not given to several nonevaluative stakeholder types, as the realisation of the involvement was estimated at different stages of evaluation separately for each nonevaluative stakeholder type.

Extent of Involvement scored a 4.00 on the scale. At least one type of nonevaluative stakeholder was represented at all stages of the evaluation process. First a planning group in which the representatives of different nonevaluative stakeholder types were as widely as possible gathered was founded in evaluations. In the planning group, among others, were Ministry of Education and Culture representatives, education providers, representatives of Finnish Industries, researchers and the students' union. The planning group was responsible for setting evaluation questions and creating the evaluation plan. The realisation group was smaller than the planning group, comprising 4–6 experts; yet it better represented most of those nonevaluative stakeholder types which were included in the planning group.

Diversity: In this dimension, a point was given for each nonevaluative stakeholder type that participated in some of the stages of evaluation (Table 2). The participation of each nonevaluative stakeholder type was estimated separately for each stage of the evaluation process.

Diversity scored a 3.50 on the scale. Representatives of the Ministry of Education and Culture, education providers and researchers participated in all the stages of the evaluation process as members of both the planning and realisation groups. The students' union representative participated at the beginning of the evaluation in the general determination of the evaluation task and questions as well as at the end of the evaluation in order to comment on the results and evaluation report, but did not participate in the data acquisition, valuing process and drawing of conclusions.

Table 1
Testing the Extent of Involvement dimension.

Description of the dimension
Extent of Involvement: the number of tasks in which at least one nonevaluative stakeholder is involved
Criteria
-Evaluation questions and issues definition or methodological design
-Data collection and analysis
-Judgments and recommendations formulation
-Report and dissemination of evaluation findings
Scale
0 = No involvement
1 = Limited/weak involvement
2 = Moderate involvement
3 = Substantial/strong involvement
4 = Full involvement

Table 2
Testing the Diversity dimension.

Description of the dimension
Diversity of participants: the number of types of nonevaluative stakeholders involved
Criteria
-Policy makers and decision makers
-Implementers/deliverers
-Target populations and intended beneficiaries; indirect and injured parties (end users of evaluation)
-Civil society and citizens (people and organisations with a political interest in the evaluation of e.g. unions, private firms and researchers)
Scale
0 = No diversity
1 = Limited/weak diversity
2 = Moderate diversity
3 = Substantial/ strong diversity
4 = Full diversity

Control of the Evaluation Process: In this dimension, a point was given for the participants' ability to make decisions in evaluation tasks given to them or to implement resources other than decisions for the evaluation process (Table 3). At separate stages of the evaluation process, 0.25 points were given to every decision made by the participants or to implemented resources.

Control of the Evaluation Process scored a 1.25 on the scale. The participants had very few possibilities to influence the setting of the evaluation task and questions. The students' union representative commented on the evaluation plan, but its appropriateness was not tested on a larger group of students, educational staff or education providers (0.25 points). However, participants were able to affect the method of data acquisition. Each educational provider was allowed to independently decide what kind of group represents the organisation (0.25 points), respond to evaluation inquiry and produce the information (0.25 points). Moreover, the self-evaluation section of the evaluation inquiry remained in use within the organisation that has participated in the evaluation (0.25 points). The participants were unable to affect the form or structure of the inquiry; it was not possible for them to affect neither the valuing process of the results nor the drawing of conclusions. The participants were able to present comments on the spread and utilisation of the results in the common informative meeting arranged at the end of the evaluation (0.25 points).

Volume of Interest: In this dimension a point was given for the interest that the participants showed in evaluation (Table 4). This dimension estimated if the evaluation need emerged among the political decision-makers or participants during the evaluation. The point was given at separate stages of the evaluation process to every evidence that proved the need or interest of the participants for the performing of the evaluation.

Table 3
Testing the Control of the Evaluation Process dimension.

Description of the dimension
Control of the Evaluation Process: the way in which control is exercised at different moments of the evaluation process
Criteria
-The authority to make decisions on a given evaluation task
-The other resources in addition to authority that nonevaluative stakeholders can mobilise to influence the evaluation process (e.g. substantive or methodological expertise, money and material resources)
Scale
0 = Exclusive control by evaluator and/or nonparticipating evaluation sponsor
1 = Limited/weak control by participants
2 = Shared control between participants and evaluators and/or nonparticipating evaluation sponsor
3 = Substantial/strong control by participants
4 = Exclusive control by participants

Table 4
Testing the Volume of Interest dimension.

Description of the dimension
Volume of Interest: the direction of interest in evaluation
Criteria
The volume of the participants' interest in evaluation
Scale
0 = Exclusive interest by evaluator and/or nonparticipating evaluation sponsor, or participants' interests are not known/not clarified
1 = Limited/weak interest by participants
2 = Shared interest between participants and evaluators and/or nonparticipating evaluation sponsor
3 = Substantial/strong interest by participants
4 = Exclusive interest by participants

Volume of Interest scored a 0.00 on the scale. This was due to the fact that there were no documented evidences in practice about the participants' needs and interest in evaluations. The evaluations were politically prepared among the decision-makers and their preparation did not include the participants' outline of requirements or preliminary assessments. Thus the serving of the dominating education policy can be considered a central task of the evaluations carried out.

Discussion

Despite the interest in research and evaluation policy, the realisation of PE in practical evaluation work has received minor attention. In this article the realisation of PE in practical evaluation activity has been examined within the context of the Daigneault and Jacob (2009) model. In the model, the value 1 has been set by each PE dimension as the cut-off point for the participation. If the value of the dimension is at least 1, it represents PE. As the dimensions measure the different sectors of participation, the realisation of PE requires a value of at least 1 across all the dimensions. On the basis of the Daigneault and Jacob model (2009), Finnish evaluation culture can be considered participatory for the first three dimensions and not participatory for the fourth.

The Daigneault and Jacob model (2009) has an ideal character. On the basis of the model, conclusions cannot be directly made for the target level of participation. Next, the conclusions will be produced regarding whether the level of participation represented by national educational evaluations can be considered desirable from the viewpoint of the realisation of the evaluation task. The conclusions are presented as two scenarios, which both stem from the observation that has been done on the basis of the article results. According to the observation, the national educational evaluation policy and national evaluation activity are in conflict with each other from the point of view of participation. The conflict is as follows: In the Decree on Evaluation (1061/2009), emphasis is put on paying impartial attention to the needs and participation of the separate parties of the evaluation. However, the political decision-makers' needs become more marked in national evaluation activity.

In the first scenario, an attempt is made to solve the conflict by paying more impartial attention to the separate parties' needs and participation in national educational evaluation activity. In the second scenario, an attempt is made to change the Decree on Evaluation (1061/2009) based on how the separate parties' current needs are emphasised in national educational evaluation activity. In the conclusions, a stance on which scenario should be sought is not taken. The aim is only to highlight the different possibilities of the scenarios to solve the conflict between the national educational evaluation policy and the evaluation activity. The choice of which scenario to target has been considered in the article, as has the task of those persons who decide the evaluation policy and evaluation practices.

Scenario 1. The evaluation task is realised properly by impartially emphasising the separate parties' needs in the target values of PE.

In the *Degree on Evaluation (1061/2009)*, the targets of the political and local decision-makers and educational institutions have been set as equal. Consequently, the realisable national educational evaluations also should serve these parties impartially. The *Degree on Evaluation (1061/2009)* is in the line with Wallace and Alkin's (2008, p. 192) definition, according to which the objective of PE is fairness and usefulness from the viewpoints of nonevaluative stakeholders.

However, the results which have been presented in the article point out that in the present situation the demand for the separate parties' impartial service does not realise. In spite of the emphasis on participation in Finland, a real power of decision at the separate stages of evaluation has not really been given to participants, nor have the participants' needs and interest in evaluation been systematically studied. The participation was strong in the first two dimensions of PE, which are fairly technical by nature. Several different nonevaluative stakeholder types can be chosen for evaluations and tasks can be assigned to them at different stages of the evaluation without paying attention to the contents or affecting the values of the participation. Thus the first two dimensions of PE do not help to determine the reasons for participation in evaluation. The third and particularly fourth dimensions of PE would address this issue.

Because of uneven realisation of the dimensions of PE, the process by which both the evaluating and participating groups would jointly produce and utilise the necessary evaluation data remains inadequate. Thus the greater focus on the third and fourth dimensions can be considered as a necessity. The comprehensive realisation of PE needs more practical points of view for and applications to the knowledge co-construction (see Smits & Champagne, 2008).

In addition to the clarification of the dynamics between the evaluating and participating groups, the realisation of PE requires considerable input into the needs of nonevaluative stakeholders before carrying out the evaluation. Finland is at the first stage in this matter, despite public emphasis on PE. One benchmark for Finland can be Denmark's evaluation culture, in which every assignment contains the preliminary study that clarifies the viewpoints of the different interest groups (Eurydice, 2010b, p. 150). Several tools that serve the practice have also been presented to clarify the needs of nonevaluative stakeholders (Bryson et al., 2011).

Scenario 2. The evaluation task is properly realised by emphasising the political decision-makers' needs in the target values of PE.

The expansion of the participants' role can be critically examined. One can ask if participation is overly emphasised in the first and second dimensions of PE. In several cases the participation of too many nonevaluative stakeholders or too much independence of these stakeholders has led to anarchy and endangerment of the evaluation task (Gastic et al., 2008; King & Ehlert, 2008; Patton, 1997). Furthermore, the involvement of several nonevaluative stakeholders in the evaluation easily raises costs; participation thus often depends also on the resources of the assignment. To reach a balance, a reduction in the participation can be considered justified in the first and second dimensions if the resources can then be increased in the fourth. This way the participation would be directed towards fewer nonevaluative stakeholders than before, but at the same time would increase their significance (see Patton, 1997).

In the examination of the previous section, an attempt is not made to pay impartial attention to the needs of the separate parties of the evaluation (cf. Scenario 1). If those needs are emphasised in

different ways, different target values should also be set for the dimensions of PE. The original model by Daigneault and Jacob (2009) is not suitable directly for setting these target values. The model has an ideal character in which the brim values of PE are not realised in all respects in practice. A total value of 0 for PE would mean a completely independent evaluation from nonevaluative stakeholder groups. Correspondingly, a total value of 4 would mean an evaluation that is self-organised by nonevaluative stakeholders and in which the evaluators become unnecessary. Consequently, a brim value 0 or 4 cannot be the target value of the realistic evaluation activity, but the target value must be somewhere between these numbers. Moreover, in the original model by Daigneault and Jacob (2009), all the parties in the evaluation have been dealt with equally. Each nonevaluative stakeholder received a similar weight value for its participation. However in practice the matter is otherwise, as the separate stakeholders' participation holds a different significance. In the present article the results that describe the third and fourth dimensions of PE indicate that the political decision-makers are primary beneficiaries of the evaluation; as such, they should receive the greatest weight value in PE.

If one wants to give the political decision-makers the biggest weight in educational evaluation activity, the *Degree on Evaluation (1061/2009)* should also be emphasised in a corresponding manner. In that case it would be clearly indicated in the Decree on Evaluation that the political decision-makers' needs take precedence in the realisation of national educational evaluations. With this procedure the uniformity between the Degree on Evaluation and national educational evaluation activity would be secured.

Acknowledgements

Thank you to Mr. Risto Hietala for information regarding the evaluations that have been examined in the article. Thank you also to the Finnish Education Evaluation Council and to director Pasi Reinikainen for the opportunity to write this article.

References

- Basic Education Act [Perusopetuslaki] 628/1998. <http://www.finlex.fi/fi/laki/ajantasa/1998/19980628> Retrieved 21.10.98.
- Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, 34, 1–12.
- Council of State. (2011). *Government programme [Neuvottelutulos hallitusohjelmasta]*. <http://www.valtioneuvosto.fi/tiedostot/julkinen/hallitusneuvottelut-2011/neuvottelutulos/fi.pdf> Retrieved 17.06.11.
- Cousins, J. B., Donohue, J. J., & Bloom, G. A. (1996). Collaborative evaluation in North America: Evaluators' self-reported opinions, practices and consequences. *Evaluation Practice*, 17, 207–226.
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14, 397–418.
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 80, 5–23.
- Daigneault, P.-M., & Jacob, S. (2009). Toward accurate measurement of participation: Rethinking the conceptualization and operationalization of participatory evaluation. *American Journal of Evaluation*, 30, 330–348.
- Eurydice. (2010a). *Organisation of the education system in Finland 2009/2010*. <http://eacea.ec.europa.eu/education/eurydice/> Retrieved 31.12.08.
- Eurydice. (2010b). *Organisation of the education system in Denmark 2009/10*. <http://eacea.ec.europa.eu/education/eurydice/> Retrieved 31.12.08.
- Gastic, B., Irby, D. J., & Zdanis, M. (2008). When stakeholders rebel: Lesson from a safe schools program. *Studies in Educational Evaluation*, 34, 208–211.
- Government Decree on Evaluation [Valtioneuvoston asetus arviointista] 1061/2009. <http://www.finlex.fi/fi/laki/alkup/2009/20091061> Retrieved 10.12.09.
- King, J. A. (2007). Making sense of participatory evaluation. *New Directions for Evaluation*, 114, 83–86.
- King, J. A., & Ehlert, J. C. (2008). What we learned from three evaluations that involved stakeholders. *Studies in Educational Evaluation*, 34, 194–200.
- Mathison, S. (Ed.). (2005). *Encyclopedia of evaluation*. Thousand Oaks: Sage.
- Ministry of Education and Culture. (2007). *Education and research 2007–2012. Development plan. [Koulutus ja tutkimus 2007–2012. Kehittämissuunnitelma]*.

- <http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2008/liitteet/opm09.pdf?lang=fi> Retrieved 5.12.07.
- Ministry of Education and Culture. (2008). *Education evaluation plan 2009–2011* [Koulutuksen arviointisuunnitelma vuosille 2009–2011]. <http://www.okm.fi/export/sites/default/OPM/Julkaisut/2008/liitteet/opm38.pdf?lang=fi> Retrieved 20.10.08.
- Patton, M. Q. (1997). *Utilization-focused evaluation*. Thousand Oaks: Sage.
- Smits, P. A., & Champagne, F. (2008). An assessment of the theoretical underpinnings of practical participatory evaluation. *American Journal of Evaluation*, 29, 427–442.
- Themessl-Huber, M. T., & Grutsch, M. A. (2003). The shifting locus of control in participatory evaluations. *Evaluation*, 9, 92–111.
- Wallace, T. L., & Alkin, M. C. (2008). Process of evaluation: Focus on stakeholders. *Studies in Educational Evaluation*, 34, 192–193.
- Weaver, L., & Cousins, J. B. (2006). Unpacking the participatory process [Electronic Version]. *Journal of MultiDisciplinary Evaluation (JMDE)*, 1, 19–40. http://www.as-tir.it/teseo/Download/Libanora/Unpacking%20Participatory%20Process_2002.pdf Retrieved 11.04.06.

Ville Pietiläinen, D.Sc. (Admin.), is Chief Planning Officer at The Finnish Education Evaluation Council (FEEC), University of Jyväskylä, Finland.