

A Review of Theoretical Frameworks to Evaluate School-University Partnerships that Improve Teaching and Learning Practices

***Edna Milena Sarmiento- Márquez¹, Luis P. Prieto², and Gerti Pishtari³**

¹ PhD student at the School of Educational Sciences, Junior research fellow at the School of Digital Technologies in Tallinn University, Estonia.

² Senior research fellow at the School of Educational Sciences in Tallinn University. Estonia

³ Junior research fellow at the School of Digital Technologies in Tallinn University. Estonia.

*Corresponding author

Abstract

School-University Partnerships (SUPs) are long-term collaborations that solve problems of educational practice, especially those related to teaching and learning practices (TLPs). Monitoring and evaluation (M&E) are key processes to understand partnership evolution and its degree of impact (to what extent it solves the target problem of practice) while allowing for a continuous adaptation of practices, thus increasing the SUP's chances of sustainability and success. Despite their importance, SUPs M&E are still challenging for practitioners. Due to the long timespan and evolving nature of SUPs, the intended changes in TLPs are difficult to measure. Moreover, practitioners do not have clear guidelines about what to evaluate, when, and how. In this paper, we reviewed seven frameworks related to the evaluation of partnerships in education to understand how they are useful in the M&E of SUPs that improve TLPs. The revision shows that the frameworks only partially address the evaluation of SUPs that promote TLPs. They focus on specific contexts (i.e., pre-service teachers), phases of development (early stages), or processes (i.e., collaboration) within a partnership. There is thus a need for a holistic, more complete framework that guides practitioners to evaluate both the collaboration and the outcomes of SUPs throughout their lifecycle (i.e., since SUPs initiation until their sustainability). We try to fill this gap by proposing a rationale for an evaluation framework to guide the evaluation of SUPs that improve TLPs. Future research should focus on developing tools that can be widely used, allowing for the transferability of practices and knowledge independently of the context.

Keywords: Educational innovations; monitoring and evaluation; evaluation framework; evaluation of teaching and learning practices; School-University Partnerships.

1. Introduction

School-University Partnerships (SUPs) are a prominent strategy to foster educational innovations (Clark, 1988; Coburn et al., 2013). They involve multiple stakeholders in long-term collaborations to solve problems of educational practice (Coburn, et al., 2013). The

3rd International Conference on Innovative Research in EDUCATION

23-25 July 2021

Stockholm, Sweden

multi-stakeholder nature of SUPs provides benefits to each partner. While teachers are provided with professional development, students get enhanced instruction that can improve their learning, and researchers develop further in the field (Stahmer et al., 2015; Imants & van de Ven, 2011). In general, SUPs have the potential for educational and social innovations as evidence-based practices are used to adapt processes during SUP's lifetime (Mclaughlin & Black-Hawkins, 2007; Handscomb et al., 2014). While SUPs can have multiple aims (e.g., teacher professional development, integration of technologies, etc.), we will focus on those SUPs that focus on improving teaching and learning practices (TLP) because they directly impact the objective of the education endeavor (Mclaughlin & Black-Hawkins, 2007).

Despite their benefits, SUPs face some challenges. For instance, the long-term commitment of stakeholders requires sustained efforts to maintain a healthy relationship among them (Mclaughlin & Black-Hawkins, 2007). Moreover, the sustainability of enhanced TLPs can fade over time due to structural aspects related to a lack of support or funding (Seels et al., 2003). Additionally, the complexity of educational change makes it difficult to track and evaluate the outcomes of SUP in terms of TLPs (Desimone et al., 2016). Moreover, the evaluation of SUPs over time is particularly challenging, as changes might take time to be evident (Desimone et al., 2016). The factors affecting the collaboration and thereby the TLP outcomes might play a different role at different stages of the SUP, for instance, building trust might be more important at the beginning of the SUP than at the end of it (e.g., Ervin et al., 2007). Furthermore, the factors to be evaluated in an SUP are not well-defined, which makes it difficult to develop measurement instruments (Desimone et al., 2016; Martinez et al., 2014)). Meanwhile, the existing instruments to evaluate SUPs are not widely used, since they are usually developed ad-hoc to suit the purposes and context of specific SUPs (Fynn et al., 2020; Desimone et al., 2016). Also, as a consequence of the limited use of existing instruments to evaluate SUPs, there is a lack of common metrics, making it difficult to assess the impact of SUPs in different contexts (Clifford & Millar, 2008).

Problems in the evaluation of SUPs pose several challenges to SUP practice. SUPs monitoring and evaluation (M&E) are needed for making decisions about changes in the planning and implementation. Yet, it is difficult to adapt processes (i.e., following data analysis for improvement) without a systematic feedback loop (e.g., results from the evaluation are to the stakeholders to make decisions based on that, resulting in maintaining -positive feedback- or dampening -negative feedback- TLPs). Moreover, putting into practice research-based changes becomes challenging when practitioners are not aware of the impact of their practices on student learning, as they cannot modify their practices accordingly (Durlak, 2015).

By analyzing several existing frameworks, regarding the evaluation of partnerships, we would have a greater understanding on how to guide practitioners in monitoring and evaluating their SUPs for TLP. To the best of our knowledge, there is no systematic assessment of existing frameworks in the context of SUPs and educational innovations. Hence, the goal of this paper is to analyze seven frameworks related to the evaluation of partnerships and to provide a rationale for their synthesis. Our aim is to understand how have different frameworks approached the evaluation of partnerships and educational innovations. In the next section, we overview the notion of SUPs and their lifecycle. Later, section 3 reports on the methods used for the assessment of existing evaluation frameworks. Section 4

describes the results of our assessment, while section 5 discusses its implications and presents possible future research avenues in the area of SUPs evaluation.

2. School-University Partnerships that improve teaching and learning practices, and their evaluation

2.1. School-University Partnerships to improve teaching and learning practices

SUPs have been defined as long-term collaborations between research and K-12 educational institutions (Clark, 1988; Coburn et al., 2013) that evolve from initial ideas to changes in TLPs. While there is a lot of research concerning teacher professional development, the aspects related to TLPs remain challenging to study due to both the specificity of each SUP (i.e., goals and context) and difficulties defining what is a successful SUP in terms of TLP (Stirman et al., 2012, Dhillon, 2013, Markiewicz & Patrick, 2016). SUP lifecycle has been described as multifaceted. This paper builds on two frameworks (the SURF framework by Hauth et al., 2019 and the GUSP -Growing University-School Partnerships- framework by Jones et al., 2016) to considers SUP as evolving through three major phases, Exploration, Implementation, and Sustainability.

As SUPs are evolving in nature (they entail different processes at different phases during their development), needing stakeholders to constantly adapt, as individuals and as an organization to new dynamics (i.e., applying research in the classroom as in Imants & van de Ven, 2011). As such, M&E contribute to SUP evolution and continuous adaptation by providing evidence that supports stakeholders' decision-making (Durlak, 2015). For instance, as data is gathered and analyzed, stakeholders can decide whether the actual course of action is to be maintained or changed to meet specific goals (e.g., Imants & van de Ven, 2011).

2.2. Evaluation of School-University Partnerships to improve teaching and learning practices

M&E have been described as key processes for the effective development of educational initiatives, as they attest to the overall value derived from them while informing a program planning and design (Markiewicz & Patrick, 2016, Durlak, 2015). While monitoring has been defined as the ongoing process of gathering information as a basis for management decisions, evaluation is a systematic process of gathering information about specific aspects of a specific educational system (Scheerens et al., 2003). In the context of SUPs that improve TLPs, monitoring has been done through regular meetings between the evaluation team to make decisions based on the results from the periodical evaluations (e.g., Ervin et al., 2007).

M&E are important processes to assess the extent to which an SUP actually solves the target problem of practice (Durlak, 2015). Thereby, a valid and reliable evaluation, allows the SUP community to improve TLPs and establish successful education systems (Markiewicz & Patrick, 2016). Moreover, having evaluation tools (i.e., frameworks, instruments, checklists, etc.) that can be widely used by the community, allows us to understand the level of advancement in the field, as we share common indicators, and to transfer good practices to different contexts (Clifford & Millar, 2008; Scherer, 2009). The availability of M&E frameworks supports practitioners' accountability and guides organizational learning in the

SUP (Markiewicz & Patrick, 2016). Despite the benefits and importance of M&E for SUP development, SUPs still struggle to evaluate their processes throughout their lifecycle, and more research is needed on this matter (Desimone et al., 2016). For instance, research in SUP M&E has focused very often on implementation (Walsh & Backe, 2013). In consequence, processes happening in the earlier exploration phase remain unmonitored and cannot explain implementation dynamics (Walsh & Backe, 2013). Similarly, the lack of sustainability studies prevents practitioners from understanding the long-term impacts of their SUPs on TLPs (Walsh & Backe, 2013; Teitel, 2001). Moreover, there is a greater need for research about the impact (i.e., outcome) evaluation of SUPs that improve TLPs (Walsh & Backe, 2013). Considering how important are the continuous M&E processes to SUPs (Durlak, 2015, Panirsilvam, 2017), we assessed several partnership evaluation frameworks, to identify how they overlap, and to discuss how they could be synthesized in the future.

2. Methodology

3.1. Selected frameworks from literature review

We performed a systematic literature review in October 2019, that aimed to understand the evaluation of SUPs that improve TLPs. We first selected six relevant databases in educational sciences: ERIC, EBSCO, SpringerLink, ScienceDirect, Scopus, and SAGE. To assemble a consistent corpus of papers, the best-resulting query was: (("Living labs" OR "University-school Partnerships" OR "School-University Partnerships" OR "Partnerships in Education") AND "Education") AND ("Adoption" OR "Implementation" OR "Sustainability"). From an initial pool of 5001 papers, that we screened and filtered through an iterative process (based on predefined inclusion criteria like papers in English that addressed any phase OF A sup that improves TLPs), we finally selected 100 papers for in-depth review. The whole process and results are described in Sarmiento-Marquez et al. (under review). From the papers left in the related work because they didn't fully comply with the criteria to be analyzed in-depth (i.e., they did not present an empirical evaluation of SUPs), we selected four of the frameworks considered in this paper for further analysis (Henrick et al., 2017; Jones et al., 2016; Ley et al., 2021; Teitel, 2001). We have chosen frameworks addressing any aspect of M&E (i.e., factors to be evaluated, how to evaluate, research instruments) in settings related to educational innovations or partnerships. Moreover, three additional frameworks (Mattessich & Johnson, 2018; Diaz et al., 2015; Jacobs, 2000) were identified from the reference lists of papers in the review. Hence, seven frameworks to evaluate partnerships have been elicited, and are analyzed in the following sections:

- Ley et al. (2021) evaluated their Knowledge Appropriation Model (KAM), to evaluate co-creation processes in SUP, through a five-point Likert scale questionnaire. The model is divided into three dimensions that explain how teachers Appropriate, Mature, and Scaffold knowledge in SUPs. Additionally, the model includes teachers' use of the improved TLPs with the "Adoption" dimension.
- The Wilder Collaboration Factors Inventory (Mattessich & Johnson, 2018) describes the factors (grouped in six dimensions) that influence multi-stakeholder collaborations. This framework was designed for all kinds of multi-stakeholder collaboration contexts and

includes dimensions such as Environment, Membership Characteristics, Process and Structure, Communication, Purpose, and Resources.

- Henrick et al., (2017) developed the framework “Five dimensions of effectiveness” to assess research-practice partnership (RPP) effectiveness by considering five dimensions. The authors considered the partnership relationships, the quality of the research to inform action, the support to achieve goals, the contribution to educational improvement, and the capacity of different partners to engage in the collaborative work.
- The GUSP framework (Jones et al., 2016), addresses SUP development and activities that support SUP evaluation in the context of pre-service teacher professional development. The framework includes components for initiating and sustaining successful partnerships across three phases of SUP and provides “Action Planning Tools” to support partners’ discussion and negotiation of the phases and components.
- The evaluative framework for assessing multi-institutional academic partnership sustainability (Diaz et al., 2015), focuses on the evaluation of factors (grouped in three dimensions) influencing the initial phases of partnerships (in general, not only SUPs).
- Teitel (2001) presents a conceptual framework for assessing Professional Development School (PDS) impacts. Such framework groups the expected PDS outcomes into four categories, suggesting data sources that can guide their evaluation.
- Jacobs (2000) presents a framework to guide the evaluation of educational innovations through a 10-step flow chart that indicates which stakeholders are concerned and which aspects should be considered.

A comparison of these frameworks is presented as part of section 4.1 below (Table 1).

3.2. Assessment procedure and criteria

We first compared general aspects of the frameworks such as their purpose, context of development and use, composition and main factors described, the total number of factors (also known as indicators, constructs, components or steps in some of the frameworks, from now on called *factors*), their visual representation, and the structure of the evaluation they suggest. An overview of such basic descriptive aspects is presented in Table 1 (Section 4.1).

The second step in the assessment consisted of the in-depth analysis of frameworks based on more specific criteria (described below). These criteria emerged in the field of social and educational innovations and has been used in education to analyze frameworks (e.g., Kirss et al., 2020). We adapted the assessment criteria to the purposes of this paper from a literature review of evaluation frameworks (Fynn et al., 2020). We also considered the criteria of comprehensiveness from indicator development theory in the area of social policy (Noll 2002), as such standards have been a useful tool for policymakers and school leaders. Moreover, we integrated criteria from the S.M.A.R.T. (Specific, Measurable, Assignable, Realistic, Time-bound) approach for indicator development (Doran, 1981) as it outlines critical aspects of monitoring tools to be functional and effective for policy (Kirss et al., 2020). For our analysis, we considered the full set of factors instead of each factor separately. Thereby, our assessment criteria can be defined as:

- Comprehensiveness A: The extent to which the analytical framework approach SUP systematically so that *the factors outlined on it, cover all different phases.*

- Comprehensiveness B: The extent to which the analytical frameworks approach partnerships holistically so that *the factors include SUP processes and SUP outcomes*.
- Comprehensiveness C: The extent to which the analytical frameworks approach partnerships so that *the factors outline clear subcategories under each category of inputs, processes, and outputs?*
- Specificity: The extent to which the factors from the frameworks *target a specific topic (i.e., organizational learning, collaboration aspects, etc.)*. To what extent the framework presents a definition of these factors instead of an abstract description.
- Measurability: The extent to which the proposed factors are *measurable or quantifiable to make their evaluation possible*. The extent to which it is possible to *collect data for such factors*.
- Relevance: To what extent *are the factors included in the framework relevant for the evaluation of SUPs that improve TLPs*.
- Time-bound: Is it specified *by when the indicators should be reached*.

Two of the authors of the paper assessed independently the frameworks by identifying whether each criterion has been met (+), partially met (+/-); not met at all (-). Once the coding finished, and the independent results were compared, disagreements were discussed and resolved by consensus, resulting in the table displayed in section 4.2.

3. Assessment of evaluation frameworks

4.1. Descriptive overview of general aspects

In general, the *purpose* of the frameworks is related to the description of specific processes (i.e., collaboration, learning, professional development, etc.) to inform planning and decision-making. Regarding the *frameworks intended use*, we identified that most of them are meant to provide guidance related to specific partnership elements (i.e., teacher professional development). Other frameworks include structured guidance to facilitate both evaluation and planning (e.g., GUSP). Some frameworks provided more flexible guidance (e.g., suggestions to guide the specific processes). For instance, Teitel and Jacob's frameworks help guide practitioners in the assessment of the exploration phase and the evaluation of educational innovations, respectively.

Regarding the *context*, we found that, despite some frameworks emerged in the context of SUPs (or RPPs), they focused on specific purposes (i.e., professional development -PD-), or specific actors (e.g., pre-service teachers). Other frameworks (Mattessich & Johnson 2018; Diaz et al., 2015; Jacobs 2000) address a broader range of collaborative initiatives (i.e., multi-stakeholder partnerships). Furthermore, we noted that most of the frameworks were derived from a literature review and/or interviews with different stakeholders from real cases. All the frameworks are multidimensional. Most of them are *composed* of factors aligned under different dimensions. Moreover, the GUSP goes one step further and provides additional tools for decision-making, such as a set of questions to help practitioners negotiate different ideas about the SUP.

Regarding the *framework representation* (format), we found that most of the frameworks provided a list of elements to be considered, organized as a table. We also found graphical

representations such as Venn diagrams (e.g., Ley et al., 2021) and flowcharts (e.g., Jacobs, 2000). Furthermore, some frameworks (KAM, Mattessich & Johnson, 2018 and GUSP) also developed their corresponding evaluation instruments (e.g., questionnaires, checklists). Table 1 presents an overview of the general aspects of the frameworks.

4.2. Detailed assessment of evaluation frameworks

The frameworks under review do not comply with the *comprehensiveness* criterion, approaching the evaluation of partnerships partially. The assessed frameworks focus on the collaboration or the outcome, but seldom both. An exemplary framework that considers both is GUSP (Jones et al., 2016). Moreover, none of the frameworks guides the evaluation considering the full SUP lifecycle, from exploration to sustainability. Indeed, Diaz et al. (2015) assess the initial phases of SUPs while anticipating factors that can support sustainability. Additionally, some frameworks described what would be the expected

Table 1. Overview of frameworks for the evaluation of partnerships

General aspects	Ley et al., 2021	Mattessich & Johnson, 2018	Henrick et al., 2017	Jones et al., 2016	Diaz et al., 2015	Teitel, 2001	Jacobs, 2000
Framework	Knowledge appropriation model (KAM)	The Wilder collaboration factors inventory	5 dimensions of effectiveness	GUSP	Evaluative framework for assessing multi-institutional academic partnership sustainability	A conceptual framework for assessing PDS impacts	Stages in the process of evaluating educational innovation
Purpose	To describe, analyze and evaluate organizational learning	To identify strengths and weaknesses in collaborative settings	To describe, organize, guide the evaluation of RPPs	To describe, analyze, and evaluate PD	To evaluate partnership readiness	To organize and conceptualize PD	To guide the evaluation of educational innovations
Context of development	Interviews for workplace learning and application in educational living labs	Systematic literature review and interviews regarding factors that influence successful collaboration	Literature review on RPPs and experts interviews	Interviews with stakeholders in the context of a SUP for pre-service teachers	Multi-stakeholder academic partnerships	PDS	Evaluation of educational innovations
Use	Analytical guidance	Analytical guidance for decision-	Analytical guidance	Structured guidance for	Analytical guidance	Flexible guidance/specific evaluation	Flexible guidance

3rd International Conference on Innovative Research in EDUCATION

23-25 July 2021

Stockholm, Sweden

		making		decision-making		element	
Composition	Dimensions and constructs	Dimensions and factors	Dimensions and indicators	3 components	Factors, indicators, and elements	Areas of outcomes and data sources	10 steps
Main factors	3	6	5	5	3	4	10
# of factors	13	22	25	15	33	11+	Not specified
Visual representation	Venn diagram	Table with domains and factors	Table with indicators	Tables presenting each domain in 3 phases	Explanatory diagram	Table expected outcomes	Flow chart
Structure of the evaluation	Set of factors + Likert scale questionnaire	Set of factors + Likert scale questionnaire	Set of indicators	Set of factors, set of questions + toolbox	Set of indicators	Set of orientative topics	Set of steps

outcomes regarding each one of the factors, providing a more comprehensive guide to practitioners.

The *specificity* criteria was met or partially met by all the frameworks. All of them described and defined the constructs they intended to measure, some of them being more specific and presenting indicators. Regarding the *measurability*, the criteria was partially fulfilled by most of the frameworks as they proposed instruments to evaluate their factors. For instance, Mattessich & Johnson (2018) uses a Likert scale questionnaire to evaluate the factors influencing the collaboration between partners. Teitel (2001) provides less structured guidelines about different sources of data that would help to evaluate each one of the factors in the framework. The assessment of the *relevance* showed that despite all the frameworks assumed that their factors have an influence on successful partnerships, most of them did not consider the factors directly related with the ability of the SUPs to solve a problem of educational practice. Indeed, this can be due to the fact several frameworks were not specifically developed in the context of SUPs.

Most of the frameworks do not establish a *time-boundary* for their factors. It was not possible to identify whether all the factors manifest similarly in all phases of the partnerships, for instance in Mattessich & Johnson's framework, we do not know if the environment dimension considers the same factors in all the SUP phases. Nevertheless, Ley et al. (2021) established that social factors might take some time to happen and hence should be measured at the middle of an SUP or at the end in order to find significant changes. Moreover, Jones et al. (2016) described how each of the five domains manifests differently at each SUP phase. Table 2 presents an overview of the assessment.

Table 2. Comparative evaluation of frameworks for the evaluation of partnerships

Evaluation criteria and frameworks	Ley et al., 2021	Mattessich & Johnson, 2018	Henrick et al., 2017	Jones et al., 2016	Diaz et al., 2015	Teitel, 2001	Jacobs, 2000
Comprehensiveness A	+	-	-	-	-	-	-
Comprehensiveness B	+/-	-	-	+	-	+/-	-
Comprehensiveness C	-	-	-	+/-	+	+/-	-
Specificity	+	+	+	+/-	+	+/-	+
Measurability	+	+/-	-	+/-	+	+/-	-
Relevance	-	-	+	-	-	+/-	+
Time-bound	+	-	-	+	-	-	-

Note: +: criterion has been met; +/-: criterion has been partially met; -: criterion has not been met, Ns: Unspecified

This assessment illustrates that the frameworks under review mostly work as broad maps of factors influencing partnerships at different levels, which can be used complementary. As none of the frameworks fulfilled all the criteria but rather seem to complement each other, elements from each framework could eventually be extracted and synthesized into a single, more holistic framework to evaluate SUPs that improve TLPs.

4. Discussion

We have analyzed and compared seven frameworks related to the evaluation of partnerships or educational innovations. Results illustrate that the assessed frameworks only partially fulfill our criteria. This assessment highlighted areas of overlap, strengths and limitations of the selected frameworks. For instance, the inclusion of specific evaluation components (e.g. the context of the partner organizations, the strategic planning, the relevance of the innovations, co-creation, etc.) in all the frameworks highlights the importance of these aspects for the M&E of partnerships. On the other hand, the frameworks seem to complement each other. While frameworks GUSP and Teitel’s framework focus on PDS, Mattessich & Johnson, consider the collaborative aspects, KAM focuses on learning, Jacobs guide practitioners through the evaluation process and Henrick et al. (2017) consider the orientation towards a problem of educational practice. We also noticed a wide linguistic diversity in the terms used to describe similar processes. For instance, while KAM uses the term “adapt” to refer to adaptation processes based on evidence, Mattessich & Johnson (2018) name it as “Evaluation and continuous learning”. The field of SUPs can promote greater consensus on how terms are defined within the community, as having an agreed common language not only would ease the accumulation of knowledge, but also would lead stakeholders (i.e., teachers, researchers, school leaders) working on SUP evaluation to achieve a better understanding of the applicability of the different frameworks to their particular contexts (Fynn, 2020).

We found limited guidance on how to evaluate the educational impact. Only the framework proposed by Jacobs (2000) provides a step-by-step guide on how to plan the evaluation (but not monitoring). The framework also includes partnership elements that can be related to the

3rd International Conference on Innovative Research in EDUCATION

23-25 July 2021

Stockholm, Sweden

educational impact of SUPS that improve TLPs (e.g., combination of summative and formative assessment, collaborative definition of indicators and outcomes, etc.). These components represent important aspects to be included when designing an evaluation framework for SUPs. Some frameworks that intended to support both planning and evaluation provided insufficient details about how these facilitated M&E. Moreover, it was not specified if the frameworks are intended to be used by researcher-led or practitioner-led evaluation, which can influence their applicability. We also identified certain gaps in most of the assessed frameworks. For instance, aspects related to the *measurability* of the constructs are ill-defined, while specific guidelines about how to evaluate, when to evaluate, and how to perform the dissemination of the evaluation results, often are non-explicit. Such aspects should be considered in the design of M&E frameworks, as they guide practitioners to make adjustments in their partnerships (Markiewicz & Patrick, 2016).

How can these frameworks be helpful for the evaluation of SUPs that improve TLPs? While the importance of an extensive and holistic evaluation of educational programs has been discussed (OECD, 2013), some researchers advocate for a more pragmatical evaluation focusing on more meaningful factors influencing SUP impact (McIntyre, 2005). Although an exhaustive evaluation of an SUP would allow tracking all the variables potentially influencing TLPs, this is challenging to execute in practice, as there might be too many research instruments to apply (or those instruments could be too long), hence overwhelming the participants (Coombe & Davidson, 2015). On the other hand, monitoring and evaluating only the most influential factors for that specific SUP case may allow practitioners to analyze them more in depth (Markiewicz & Patrick, 2016). A midpoint would be an approach that monitors several factors but evaluate fewer of them (Markiewicz & Patrick, 2016). Due to the ad-hoc nature of this approach, however, it would be difficult to generalize findings from one case to the others, impacting the transferability of practices within the field of SUP (Clifford & Millar, 2008; Scherer, 2009).

Despite the existence of several approaches to develop M&E frameworks in education (e.g., program theory-driven, learning-driven, use-driven, system-driven, etc.), there is a growing interest in developing eclectic approaches (Markiewicz & Patrick, 2016). Hence, we consider that M&E frameworks for SUPs that promote TLPs should probably consider both SUPs (collaborative) processes and their outcomes. Based on the elements elicited by the assessed frameworks in this paper, we consider that the evaluation of collaborative processes can include factors such as the *context* of the partner institutions (e.g., school characteristics and needs, national policies, etc.), *the relationships between the partners* (e.g., communication, mutualism, etc.), and *organizational aspects* (e.g., capacity building, social capital etc.). Regarding the evaluation of SUPs impact in TLPs, future frameworks can include the different kinds of *learning processes* (e.g., stakeholder learning, student learning, organizational learning) and the *SUP orientation towards problems of educational practice* (e.g., to what extent the problem that originated the SUP has been solved so far).

The limitations of this work relate to bias in the assessment of the frameworks, the framework selection and the criteria considered for the evaluation, focusing on SUPs that improve TLPs. Other limitations include the fact that the literature review did not target specifically frameworks (i.e., the word framework was not in the query to search for the papers), but rather used a strategy similar to a snowball approach, using particular SUPs and

their evaluations as the starting point. Hence, a wider review could be done in the future as well.

5. Conclusion

SUPs are rewarding collaborations, for the involved stakeholders, that promote educational innovations. Nevertheless, they have been evaluated very differently so far and thereby, the resulting knowledge is scatterly accumulated. Through the assessment of the seven existing evaluation frameworks, we have come to an understanding of overlapping and complementary areas that can be useful in the future for the development of frameworks that guide the evaluation of SUPs that improve TLPs.

Existing frameworks seem to partially evaluate partnerships, focusing on specific aspects of SUPs, stakeholders or contexts. While future frameworks can focus on factors addressing both the partnership collaboration process and its outcome (i.e., the potential impact on TLPs), they could also build upon the frameworks assessed here to synthesize them and adapt their factors and guidelines to achieve a more complete view of the success and sustainability of SUPs. Finally, the development of a framework to evaluate SUPs that improve TLPs, can be nourished by different approaches, considering summative and formative assessment, constructivist, organizational and learning approaches to provide a rich but deep M&E tool. Once such a framework is in place, practitioners of SUPs will be able to adapt their practices to increase their chances of success and sustainability over time, while the field will gain knowledge about important factors across contexts.

Acknowledgment

The authors would like to thank Katrin Poom-Valickis for her feedback for this paper. This project was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No. 669074. This project has received funding from the European Union's Erasmus Plus program, grant agreement 2019-1-NO01-KA203-060280. The research on which the article is based has been supported by the European Union from the European Regional Development Fund.

References

- Clark, R.W. (1988). School–university relationships: An interpretive review. In K.A. Sirotnik & J.I. Goodlad (Eds.), *School-university partnerships in action: Concepts, cases, and concerns* (pp. 32-65). Teachers College Press.
- Clifford, M., & Millar, S. B. (2008). K-20 partnerships: Literature review and recommendations for research. WCER working paper No. 2008-3. Wisconsin Center for Education Research (NJ3).
- Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). *Research-practice partnerships: A strategy for leveraging research for educational improvement in school districts*. William T. Grant Foundation.
- Coombe, C., & Davidson, P. (2015). Constructing questionnaires. In J. D. Brown & C. Coombe (Eds.) *The Cambridge guide to research in language teaching and learning*, (1st ed., pp. 217-223). Cambridge University Press.

- Desimone, L. M., Wolford, T., & Hill, K. L. (2016). Research-practice: A practical conceptual framework. *AERA Open*, 2(4). <https://doi.org/10.1177/2332858416679599>
- Dhillon, J. K. (2013). Senior managers' perspectives of leading and managing effective, sustainable and successful partnerships. *Educational Management Administration & Leadership*, 41(6), 736-750.
- Diaz, J. M., Stallings, K. D., KC, B., & Seekamp, E. (2015). Evaluating multi-institutional partnership sustainability: a case study of collaborative workforce development in renewable energy assessment. *Educational Research and Evaluation*, 21(5-6), 466-484.
- Doran, George T. (1981) "There's a S.M.A.R.T. way to write managements's goals and objectives". *Management Review* 70 (11), 35.
- Durlak, J. A. (2015). What everyone should know about implementation. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 395–405). The Guilford Press.
- Ervin, R. A., Schaughency, E., Matthews, A., Goodman, S. D., & McGlinchey, M. T. (2007). Primary and secondary prevention of behavior difficulties: Developing a data-informed problem-solving model to guide decision making at a school-wide level. *Psychology in the Schools*, 44(1), 7–18. <https://doi.org/10.1002/pits.20201>
- Fynn, J. F., Hardeman, W., Milton, K., & Jones, A. P. (2020). A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health*, 20(1), 1-16.
- Handscomb, G., Gu, Q., & Varley, M. (2014). School-university partnerships: Fulfilling the potential. Literature review. Research Councils UK and National Co-ordinating Centre for Public Engagement.
- Hauth, C., Cuenca-Carlino, Y., Mills, S., Allen-Bronaugh, D., & Thompson, C. C. (2019). Learning to SURF: Teachers and researchers creating partnerships for success. *Teacher Educators' Journal*, 12, 73–92.
- Henrick, E. C., Cobb, P., Penuel, W. R., Jackson, K., & Clark, T. (2017). Assessing research-practice partnerships: Five dimensions of effectiveness. New York, NY: William T. Grant Foundation
- Imants, J., & van de Ven, P.-H. (2011). Practice-based research on the development of activating instruction and self-directed student learning: Dutch writing instruction. *Journal of Curriculum Studies*, 43(3), 333–355. <https://doi.org/10.1080/00220272.2010.516073>
- Jacobs, C. (2000). The evaluation of educational innovation. *Evaluation*, 6(3), 261-280.
- Jensen, B., & Iannone, R. L. (2018). Innovative approaches to continuous professional development (CPD) in early childhood education and care (ECEC) in Europe: Findings from a comparative review. *European Journal of Education*, 53(1), 23-33.
- Jones, M., Hobbs, L., Kenny, J., Campbell, C., Chittleborough, G., Gilbert, A., ... & Redman, C. (2016). Successful university-school partnerships: An interpretive framework to inform partnership practice. *Teaching and Teacher Education*, 60, 108-120.
- Kirss, L., Säälik, Ü., Pedaste, M., & Leijen, Ä. (2020). A Synthesis of Theoretical Frameworks on Multilingual Education for School Leaders. *Trames*, 24 (74/69), 1-27.
- Ley, T., Tammets, K., Sarmiento-Márquez, E. M., Leoste, J., Hallik, M., & Poom-Valickis, K. (2021). Adopting technology in schools: modelling, measuring and supporting knowledge appropriation. *European Journal of Teacher Education*, 1-24.
- Markiewicz, A., & Patrick, I. (2015). Developing monitoring and evaluation frameworks. Sage Publications.

3rd International Conference on Innovative Research in EDUCATION



23-25 July 2021

Stockholm, Sweden

- Martinez, R. G., Lewis, C. C., & Weiner, B. J. (2014). Instrumentation issues in implementation science. *Implementation Science*, 9(1), 1-9.
- McIntyre, D. (2005). Bridging the gap between research and practice. *Cambridge Journal of Education*, 35(3), 357–382. <https://doi.org/10.1080/03057640500319065>
- McLaughlin, C., & Black-Hawkins, K. (2007). School-university partnerships for educational research: Distinctions, dilemmas, and challenges. *Curriculum Journal*, 18 (327–341).
- Mattessich, P., W. & Johnson, K. M. (2018) *Collaboration: What Makes It Work* (3d ed.). Fieldstone alliance.
- Noll, Heinz-Herbert (2002) “Towards a European system of social indicators: theoretical framework and system architecture”. *Social Indicators Research*, 58 (1–3), 47–87.
- OECD (2013) Synergies for better learning. an international perspective on evaluation and assessment. *OECD Reviews of evaluation and assessment in education*. <https://doi.org/10.1787/9789264190658-en>
- Panirsilvam, S. R. (2017). Literature review: Scaling up of education innovation and their impact on students learning outcomes. Sydney, NSW, Australia: Centre for Research on Computer-supported Learning at The University of Sydney.
- Sarmiento-Márquez, E. M., Pishtari, G., Prieto, L.P. & Poom-Valickis, K. (Under review) The Evaluation of School-University Partnerships that Improve Teaching and Learning Practices: A Systematic Review.
- Scheerens, J., Glas, C. A., Thomas, S. M., & Thomas, S. (2003). Educational evaluation, assessment, and monitoring: A systemic approach. 13. Taylor & Francis.
- Scherer, J. (2009). Understanding the Role of Partnership Configuration in the NSF MSP Program. *Journal of Educational Research & Policy Studies*, 9(2), 1-21.
- Seels, B., Campbell, S., & Talsma, V. (2003). Supporting excellence in technology through communities of learners. *Educational Technology Research and Development* 51(1), 91-104.
- Stahmer, A. C., Rieth, S., Lee, E., Reisinger, E. M., Mandell, D. S., & Connell, J. E. (2015). Training teachers to use evidence-based practices for autism: Examining procedural implementation fidelity. *Psychology in the Schools*, 52(2), 181–195.
- Stirman, S. W., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. *Implementation Science*, 7(1), 17.
- Teitel, L. (2001). An assessment framework for professional development schools: Going beyond the leap of faith. *Journal of Teacher Education*, 52(1), 57–69.
- Walsh, M. E., & Backe, S. (2013). School-university partnerships: Reflections and opportunities. *Peabody Journal of Education*, 88(5), 594–607.