

Impact Evaluation of Teacher-Training Extension Program in Higher Education: A Qualitative Assessment Approach

Cherrylyn P. Labayo

cherrylyn.labayo@bicol-u.edu.ph

Bicol University Center for Policy Studies and Development

Abstract While the teacher-training program is one of the prominent extension modalities in Higher Education Institutions, a closer look at the literature has revealed several gaps in evaluating the long-term impacts of training programs. The present study addressed limitations related to the absence of baseline data and the evaluation framework by using Kirkpatrick's four levels of evaluation to assess a teacher-training extension program at a university in Bicol region, Philippines. Findings suggest that the teacher-training was successfully implemented at the reaction level but failed to assess changes in learning and behavior after the training was conducted. However, the results level was measured using the Qualitative Impact Assessment Protocol (QuIP). Causal statements from the key informants in QuIP revealed positive changes during the evaluation period but without explicit reference to the training program. Based on the evaluation results, lessons learned were documented along with the training extension program's timing, duration, and monitoring. Consequently, the study recommended the appraisal of existing practices and extension policies on teacher-training programs and other similar undertakings with evaluability criteria and standards.

Keywords: *Evaluability; Kirkpatrick model; lesson learning; qualitative assessment; training extension program*

Introduction

Teacher-training has been an important modality in education systems as a strategic tool to achieve quality education of the Sustainable Development Goal 4 (SDG 4) outlined by the United Nations (Dange & Siddaraju, 2020). As a form of professional development, teacher-training aims to improve teachers quality and learning outcomes so that students develop the skills necessary to contribute to the betterment of the society. Thus, it is unsurprising that governments invest heavily on teacher-training programs. Data shows that nearly two-thirds of World Bank projects in developing countries were funneled into teacher-training (Popova et al., 2016). In the Philippines, billions of investments were allocated for human resource development and training programs to be implemented by the Department of Education (DepEd) and the Commission on Higher Education (CHED) from 2017 to 2022 (National Economic and Development Authority [NEDA], 2018).

In higher education, the CHED defines extension as the act of transferring knowledge and technology generated by Higher Education Institutions (HEIs) to address developmental concerns (CMO No. 52, s. 2016). In national development and SDG 4, teacher-training extensions can be classified under the CHED platform of Education for STEAM (Science, Technology, Engineering, Agriculture-Fisheries, and Mathematics), which calls for extension programs to focus on lifelong learning and impact assessment on student learnings on STEAM.

Consequently, evaluating teacher-training extension programs is extremely necessary to carefully

assess the program's merit and deal with issues concerning accountability, organizational learning, and future policy formation (Rogers et al., 2015). However, measuring training outcomes has always been frustrating for institutions because of data limitations and the absence of rigorous evaluation frameworks. A review of the literature reveals that existing approaches to teacher-training were mostly unsubstantiated by high-quality evidence of their impact despite substantial spending in this area (Popova et al., 2016). Past research had indicated little consistent evidence on the impact of teacher-training programs on teacher and student outcomes (Thurlings & Brok, 2017; Zhang et al., 2013).

In developing countries, teacher-training programs evaluated using credible methodologies were scant (Popova et al., 2022; Schaffner et al., 2021). Several impact evaluation studies focused on improving numeracy and literacy skills in primary and secondary education. In Kenya, for instance, the impact of teacher-training literacy skills was measured using an experimental design (Jukes et al., 2017; Lucas et al., 2014). In a similar study in the Philippines, teacher-training impacts on reading skills were measured using a randomized experimental approach (Abeberese et al., 2014). Other comparable studies were conducted by Fuje and Tandon (2018) in Mongolia and Albornoz, and colleagues (2019) in Argentina. Results showed that for short-term teacher-training programs, such as the case of the Philippine and Mongolia studies, student performance reported small positive impacts. According to Gulamhussein (2013), it takes a minimum exposure of 50 to 80 hours to training and post-training support before any outcomes become apparent. The case of the Argentina study supports this finding, wherein results reported significant effects.

A key question in any impact assessment of programs points to attribution. However, in instances involving small "n" cases, experimental or quasi-experimental designs are not

practical and not feasible to permit statistical inferences on the magnitude of impacts created by the intervention (White & Phillips, 2012). Contribution analysis thus emerged as an alternative tool to document plausible conclusions about the impacts contributed by the intervention (Mayne, 2008). Causality is then inferred from the narrative statements of program beneficiaries, taking into consideration the multiple sources that might have influenced the changes reported.

An example of an impact evaluation approach that draws on contribution analysis is the Qualitative Impact Assessment Protocol (QuIP) (Remnant & Avard, 2016). The QuIP involves four steps. The first step is selecting the domains of change based on the outcomes of interest underpinning a project. These domains are then incorporated in the questionnaire, which is implemented in select sample cases. Finally, QuIP analysis can be done either inductively by identifying patterns or deductively using predetermined themes.

Another practical framework commonly used by evaluators in HEIs is Kirkpatrick's four levels of evaluation (Cahapay, 2021). For Kirkpatrick, evaluation is a four sequential process for capacity development (Japan International Cooperation Agency [JICA], n.d.): Reaction, Learning, Behavior, and Results. The first level assesses participant reactions to the training program in the same way as customer satisfaction. Questionnaires are the most common tool used to evaluate this level. The types of information asked of participants usually include topics concerning the content, trainer, relevance, time, facilities, overall evaluation, and suggestions. On the second level, changes in participants' attitudes, knowledge improvement, or skills improvement are being assessed. The pre-and post-test is the commonly used tool to measure this level. Level three focuses on participants' changes in behavior attributed to attending the training program. In the final evaluation

stage, final outcomes attributed to the training program are assessed using an impact survey.

The study proposed a systematic and practical approach to evaluating the impact of the teacher-training program implemented by a university considered as a regional state higher education institution in Bicol Region, Philippines. Specifically, the training program was assessed using Kirkpatrick's four levels of evaluation, wherein the results level was assessed using the QuIP. Recognizing the role of science in producing students capable of facing techno-scientific and social issues (Jocz et al., 2014), the training programs focused on topics necessary to help the schools establish a learning environment that promotes a healthy and productive life for students through science education. The training components focused on the following topics: (1) medicinal flora, (2) edible gardening, and (3) antimicrobial resistance. In addition, the study analyzed the differentiated narrative causal statements in the three training components. Finally, the study documented lessons learned in assessing efforts on future teacher-training programs. The implementation period was between June 2013 and March 2015, making the program suitable for an impact study.

The impact study contributes to the scant literature in two ways. First, an insufficient formal evaluation is published in reputable journals addressing the impact of teacher-training programs implemented in small "n" cases. Second, the evaluation specifically focused on topics relevant to science education. The study results will be helpful for teachers and other professionals planning to implement similar training programs, particularly on how training programs should be designed so that meaningful impact can be measured using a credible approach.

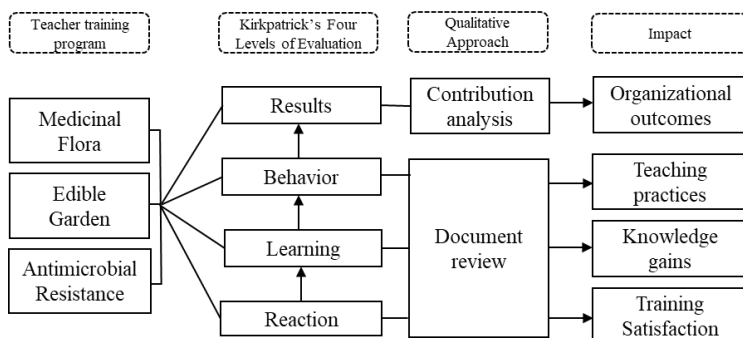
Methodology

Research Design

This study employed a qualitative approach to investigate the impact of the teacher-training program initiated by a university, in Bicol, Philippines. The teacher-training program started in 2013 and officially ended in 2015. Because of the absence of baseline data, it was impossible to select a true counterfactual. Given this data limitation, the study employed a non-experimental design. This analysis is statistically much weaker than Randomized Control Trial (RCT) and quasi-experimental designs. However, the lessons learned from the evaluation may be used to appraise future teacher-training programs.

Figure 1

Research Framework for the Teacher-Training Program Impact Study



The study adopted Kirkpatrick's framework to assess the program's immediate, short-term, and long-term impacts. The first three levels relied on existing data provided by the program implementers. The reviewed documents included the program proposal, monitoring reports, and accomplishment

reports for the three training components: medicinal flora, edible garden, and antimicrobial resistance. For the results level, contribution analysis was used to determine the organizational outcomes of the teacher-training program in partner schools (Figure 1).

Case Selection

The study utilized the University of Bath's Qualitative Impact Protocol (QuIP) Guidelines for Field Use (Copestake & Remnant, 2014). The QuIP approach to sampling was through purposive sampling. Respondents for the Key Informant Interview (KII) were 23 teachers who were the direct beneficiaries of the teacher-training and have handled science subjects for at least five years in the partner schools under the Legazpi Port District II of the DepEd, Legazpi City Division. Because these teachers have gained experience in the target schools during the evaluation period, they were the most qualified to provide information about the organizational outcomes that the research sought to investigate.

Data Collection

Data collection for the first three levels of Kirkpatrick's framework involved the review of existing documents provided by the extension implementers. The objective of the document review was to determine the immediate and short-term impact of the teacher-training. Documents requested included the training evaluations to assess training satisfaction (Reaction). Any changes in knowledge (Learning) and teaching practices (Behavior) were assessed by looking into monitoring and accomplishment reports.

The long-term impact (Results) was assessed using QuIP. The QuIP questionnaire included three open-ended and three closed questions concerning medicinal flora, edible gardening, and the promotion of antibiotic

resistance. Open questions refer to a general question asking the science teachers to narrate any changes concerning the school's performance in maintaining school gardens and promoting antimicrobial resistance from 2015 to 2019 if any changes have occurred. Each open question was followed by supplementary questions, which asked the reasons for the changes mentioned and any new activities undertaken by the school during the evaluation period. After which, respondent validation or member check was performed so that the key informants had the opportunity to add more information or suggest corrections to the statements captured during the interview.

The QuIP method employs double blindfolding to avoid confirmation bias on key informants. However, the study employed partial blinding, which the QuIP also allows, as long as questioning during the interview does not prompt study participants to respond to questions with prior understanding and interest in the teacher-training program being evaluated. In other words, the key informants were not briefed on the program being evaluated by asking them questions about the general changes they observed without directly mentioning the teacher-training program. Blindfolding is considered ethically acceptable since blindfolding is considered temporary rather than a permanent state. After the interview, the key informants were fully debriefed and acknowledged the necessity and usefulness of blindfolding during the interview process.

Other ethical considerations were also observed during the interview. Informed consent, emphasizing the anonymity and confidentiality of responses, was sought from each key informant. Furthermore, it was made clear that their participation was voluntary and that their information would remain confidential if they chose to end the survey at any time.

Data analysis

The study employed thematic analysis to analyze qualitative data from key informant interviews, which involved qualitative coding of drivers, outcomes, and attribution that emerged during the interviews. Cause-and-effect statements were coded and categorized into four types of statements: (1) Expl = changes with explicit attribution to the training program or named training program; (2) Impl = changes referring to outcomes of interest by which the training program aims to achieve, but with no explicit attribution to the training program or named training program; (3) Inci = change attributed to other activities not related to the training program's outcome of interests; and (4) Unat = change not attributed to any specific cause. These statements were further categorized as positive or negative.

Results

This section is organized into two parts. The first part presents the results for the reaction level of assessment based on the document review. The second part reports the findings of the results level of assessment based on the qualitative data generated from the QuIP.

Reaction Level of Assessment

Based on the document review, the average level of satisfaction ranged from 4.38 to 4.70 for medicinal flora training components, and 4.62 to 4.77 for antimicrobial resistance components (Table 1). The findings suggest that the teachers found the quality of the training to be very satisfactory based on the criteria listed. However, data limitations for the edible garden component made it difficult to say much about trainees' satisfaction at the reaction level.

Table 1.

Summary of Training Evaluation Surveys on Teacher-Training on Medicinal Flora and Antimicrobial Resistance

Training conducted^a	Date of implementation	Number of Participants	Average Rating^b
Medicinal Flora			
Medicinal Plants and Ointment Preparation Workshop	July 30, 2014	16	4.38
	Aug. 31, 2014	66	4.38
	Sept. 19, 2014	17	4.61
Medicinal plants gardening	Feb. 26, 2015	66	4.38
	July 31, 2015	66	4.61
Medicinal plants training: Preparation of Akapulko Ointment	Sept. 25, 2015	71	4.64
Antimicrobial Resistance			
Antimicrobial susceptibility testing on substances	Feb. 15, 2014	37	4.77
Lecture forum on water and foodborne diseases and antimicrobial resistance	Aug. 6-7, 2014	36	4.63
Lecture on antimicrobial resistance	Sept. 17, 2014	28	4.62
Antimicrobial resistance training and seminar forum on health-related topics	Feb. 13, 2015	36	4.66

^aNo data available for the Edible Garden Component

^b1 – Poor, 2 – Fair, 3 – Good, 4 – Very Good, 5 - Excellent

The absence of monitoring data provided insufficient data to analyze the outcomes of the teacher-training on learning and behavior levels. This limitation in data indicates the absence of a mechanism in the university that required program managers to use rigorous evaluation frameworks. However, the study assessed the results level to determine the final outcomes of the training program five years after the partner schools received the intervention.

Results Level of Assessment

Key informants supplied the perceived changes from 2015 to 2019 in partner schools using recall (Table 2). Based on the data captured by QuIP, positive changes were associated with the ability to maintain the medicinal garden (9 cases), the ability to maintain an edible garden (17 cases), and the ability to promote the rational use of antibiotics (5 cases).

Table 2.

Responses to Closed Questions in QuIP

Responses	Ability to maintain a medicinal garden n=23	Ability to maintain an edible garden n=23	Ability to promote the rational use of antibiotics n=23
Better	9	17	5
No Change	6	0	4
Worse	1	1	1
Not Sure	0	0	3
No response	7	5	10

As a form of juxtaposition, responses to closed questions were analyzed alongside the causal statements provided by the key informants. There were bare to no explicit positive statements on the three domains (ability to maintain medicinal and edible garden, and ability to promote rational use of antibiotics) based on the causal statements generated from the narrative data of study participants (Table 3), an indication that the training program was weak in terms of establishing the benefits at the institutional level. On a positive note, there were numerous references to training related to medicinal flora and edible gardening that confirmed the increased ability of schools to maintain medicinal and edible gardens in their area.

Table 3.

Frequency of Causal Statement of Respondents From KII

Indicators	POSITIVE				NEGATIVE			
	Expl ^a	Impl ^b	Inci ^c	Unat ^d	Expl	Impl	Inci	Unat
Ability to maintain a medicinal garden	0	6	1	4	0	2	2	1
Ability to maintain an edible garden	0	13	5	0	0	0	1	0
Ability to promote the rational use of antibiotics	1	5	0	0	0	0	0	8

^aExpl = changes with explicit attribution to the training program or named training program; ^bImpl = changes referring to outcomes of interest by which the training program aims to achieve, but with no explicit attribution to the training program or named training program; ^cInci = change attributed to other activities not related to the training program's outcome of interests; ^dUnat = change not attributed to any specific cause

In other words, when asked about the reasons for changes along with the three domains, the key informants cited positive outcomes but without mentioning the name of the training program. For example, one of the key informants reported saying:

“Dito sa amin, patuloy ang pagpapanatili ng medicinal garden, may iba’t ibang mga halamang gamot na may tamang label. Sa tulong ng PTA at YES-O, nabigyan ng tamang lugar ang medicinal garden.” (Key Informant 3) (The school had sustained medicinal garden, with different herbs properly labeled. Appropriate locations were identified with the help of PTA and YES-O.)

Interestingly, one key informant stated:

“Simula 2015-2019, malaki na ang ipinagbago ng school garden namin. Nung sumali nga kami sa regional contest, kami ang nanalo ng best-implementing school ng Gulayan sa Paaralan Program noong 2018.” (Key Informant 10) (Our

school's ability to maintain edible gardens changed a lot from 2015-2019. In fact, we participated in the regional contest and won the best-implementing school of Gulayan sa Paaralan Program in 2018.)

Another key informant pointed out the practical and educational usefulness of having maintained medicinal and edible gardens:

“Sa aking klase sa science at EPP, ginagamit ko ang school garden para mas maintindihan ng mga estudyante ko ang mga bagong konsepto.” (Key Informant 14) (In my science and EPP class, I am using the school garden to introduce new concepts to my class.)

Some respondents identified positive drivers of change incidental to the training program's outcome of interest—those relating to increased vegetable consumption and improved health often linked to DepEd programs. As an example, one key informant stated:

“Nakatulong talaga ‘yung Gulayan sa Paaralan Program na mapanatili ang school garden dito. Sa katunayan nga, ang ilan sa mga gulay na inani ay nagamit para sa aming feeding program. Dinadala din minsan ng mga magulang at estudyante ang mga sobra na gulay.” (Key Informant 9) (The Gulayan sa Paaralan Program helped maintain the school's edible garden. In fact, some of the vegetables harvested were used for our feeding program and even brought home by parents and students.)

Meanwhile, the negative drivers of change not attributed to any specific cause were highest in the antimicrobial resistance training component. Most key informants simply stated that their schools do not promote information drives on antimicrobial resistance.

To avoid bias from the narrative quotations that support specific points, data analysis for drivers of change for each training component mentioned by at least two respondents was inductively grouped for each training component (Table 4). The findings reveal that similar drivers were repeatedly mentioned in different impact domains. For example, in the case of medicinal flora and edible gardening, school-based organizations were widely cited as improving the ability of schools to maintain medicinal and edible gardens.

Table 4.

Positive and Negative Drivers of Change Widely Cited by Key Informants

Training Outcome	Positive	Negative
Ability to maintain a medicinal garden	School-based organization support (Parent Teacher Association (PTA), Youth for Environment in Schools Organization (YES-O)) (3)	Typhoons (2)
	Government memo/order (Department of Education); Subject curriculum (4)	Lack of time, resources, staffing (3)
Ability to maintain an edible garden	Government support (7) School and stakeholder support (parents, teachers, student organizations) (11)	
Ability to promote the rational use of antibiotics	Wash and win program on sanitation and good hygiene (2)	No emphasis on health subjects; Lack of dissemination on prudent use of antibiotics (8)
	Access to information found on the internet (2)	

Participants identified several facilitating factors and various barriers to maintaining medicinal and edible gardens and promoting rational consumption of antibiotics. Difficulties in staffing, time, and resources were identified

as negative drivers in maintaining school gardens. One key informant even stated:

“Napakarami kasing paperworks. Sabay-sabay din madalas ang activities kaya hindi na nabibigyan ng panahon ang pag maintain ng school garden dito.” (Key Informant 11) (Too much paperwork and overlapping activities decreased the school’s effort to maintain the school garden.)

In the case of the antimicrobial resistance component, the major difficulty identified was applying the topic to health-related subjects.

Discussion

The analysis used Kirkpatrick’s framework to assess the immediate (reaction and learning), short-term (behavior), and long-term (results) impacts of the teacher-training program. A review of existing documents from the program implementers revealed that the teacher-training program was only assessed immediately after the training. Thus, because of insufficient data, no evidence was established about the program’s learning and behavior impacts. Although the results suggest that the training program met the expectations of most of the participants, evaluation at reaction level does not measure what participants have learned but only gauges participants’ interest, motivation, and attention levels (Smidt et al., 2009).

Consistent with this finding, a recent study by Wallo and colleagues (2020) concluded that training programs were seldom evaluated for their impacts and were evaluated only at Kirkpatrick’s Level 1. Articles examined reveal inhibiting factors that preclude institutions from evaluating the final outcomes of training programs, such as the absence of assessment systems and processes and lack of assessment

methods and tools. Nonetheless, on teacher knowledge and practices, the literature suggests that professional development initiatives create impacts. Specific to science education, a review of empirical studies showed that science teachers' participation in professional development efforts enhanced their disciplinary content knowledge and pedagogical content knowledge (Dogan et al., 2016). Changes in teacher practices, on the other hand, imply that participation in professional development programs resulted in improvements in science teaching practices, which included the application of student-centered approaches through facilitation and scaffolding of inquiry-based teaching (Diaconu et al., 2012).

An important implication of this finding for teacher-training implementers is how data should be collected, analyzed, and reported. Existing regulation by CHED outlines the strategic thrusts to equity, relevance, and advancement in research, innovation, and extension in Philippine higher education (CMO No. 52, s. 2016). The guideline explicitly requires extension proposals to have a monitoring and evaluation plan with clear and measurable indicators and measurement methods. Hence, the university needs to revise and align its existing policies governing the systems, processes, and practices in designing teacher-training programs and other extension programs with CHED guidelines so that desired outcomes are determined during the design stage. With an appropriate evaluation framework in place, this should increase the likelihood that data on all levels of Kirkpatrick's framework are collected (Rooij et al., 2015).

Assessing the outcomes of training programs at reaction level is not enough. It is also important to determine if learning resulted in relevant organizational outcomes, which is only possible if levels three and four are not neglected (Throgmorton et al., 2016). However, the complexity lies in

multiple factors influencing how the training program has been translated into organizational outcomes (Grossman & Salas, 2011). Using Kirkpatrick's framework alone cannot probe the factors that inhibit or facilitate the transfer of learning. Thus, the analysis looked into the narrative accounts to identify barriers and facilitators that influenced the final outcomes of the teacher-training program, which can be viewed as a contribution analysis (Copestake & Remnant, 2014).

Analyzing the qualitative data from QuIP reveal several patterns and themes related to key factors that influenced the final outcomes of the teacher-training program in partner schools. Corresponding with previous research, results from the impact study found that stakeholder support played a central role in maintaining school gardens (Schreinemachers et al., 2019). The impact study also supports the findings of past studies that pointed out several barriers to maintaining school gardens and their use for experiential learning (Burt et al., 2018; Huys et al., 2017). For example, Loftus and colleagues (2017) identified funding, staff, and volunteer support as major barriers in Illinois public schools. In Nepal, barriers associated with using school gardens in science education included time constraints, gaps in the science curriculum, and overlapping engagements of parents (Acharya et al., 2020). This is an important finding that future teacher-training providers can build upon.

A review of the impact of school gardens on student performance revealed consistent positive outcomes. Studies showed that maintaining a school garden resulted in positive health impacts like nutritional benefits (Landry et al., 2021) and well-being impacts derived from satisfaction and pride from nurturing the plants (Ohly et al., 2016). However, a divide exists between those teachers who are willing to use the school garden in their lessons and those who are not. As Gozalbo and colleagues (2020) pointed out, there is a gap

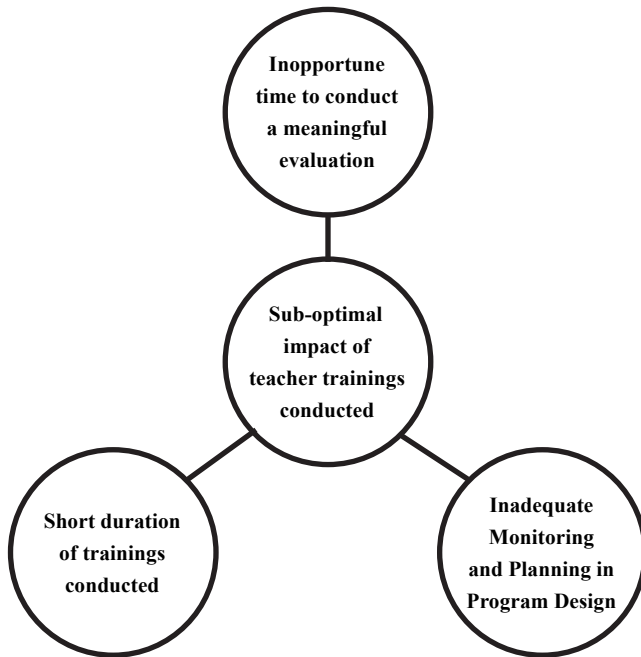
between the science curriculum and daily life. Numerous policies, such as the DepEd Memorandum No. 293 s. 2007 (Gulayan sa Paaralan Program) and the DepEd Memorandum No. 187 s. 2018 (Pilot Implementation of the School Inside a Garden Program) was issued by the DepEd to strengthen the integration of environmental education in the elementary and high school curriculum. Key informants in the study area expressed familiarity with these policies. However, only a few applied them in practice. As emphasized by the key informants during the interview, administrative works preempted them from exploring outdoor spaces, like school gardens, as contextual scaffolding in teaching science since maintaining gardens requires time, labor, and resources. The result of the impact study indicates the need to explore approaches to policy support that will strengthen and promote the transfer of learning in schools.

Another scientific knowledge that has a direct impact on learners is antimicrobial resistance. Most research findings supported the idea that information campaigns on antibiotics consumption changed attitudes and behaviors (Azevedo et al., 2013; Lee et al., 2015; Thong et al., 2021). Recognizing the threat of antimicrobial resistance to global health and food security, the DepEd has been championing antimicrobial stewardship through its policies, such as during World Antibiotic Awareness Week (WAAW). Although some key informants had a general sense of the concept and its risks, many were indifferent about its promotion in schools. Comparable studies in developing countries also revealed that students and teachers demonstrated poor knowledge and perception of antibiotics (Askarian & Maharlouie, 2012; Kotwani et al., 2016). The results suggest that the previous and current efforts to promote antimicrobial stewardship are still far from being materialized. To this end, Marvasi and colleagues (2021) suggested the “One Health” approach to arresting the surge of antimicrobial resistance in the post-antibiotic era.

The impact study findings draw lessons for future evaluations of teacher-training programs in higher education. Experiences at the time of the evaluation reveal three key barriers that resulted in the program's sub-optimal achievement of impact (Figure 2).

Figure 2

The Framework of Lessons Learned for the Teacher-Training Program



Timing

The timing of the evaluation was inopportune to conduct the final evaluation of the program. With the COVID-19 pandemic, several restrictions were encountered during the data collection. This kind of extraordinary situation made it challenging to get information through recall. Thus, the

evaluation period did not include the perceived changes in 2020. When evaluating a program, it is crucial to have an environment that is facilitative to evaluation. When shocks arise (e.g., pandemic and economic recessions), UNDP (2021) suggests that evaluations need to be flexible in their methodology and data-collection approach or wait until the environment becomes conducive for evaluation settings.

Duration

It is hardly possible for short-duration training to establish long-term impacts. For instance, research suggests that a short-course continuing professional development (CPD) can achieve short-term outcomes by improving the self-efficacy of teachers and teaching assistants but remains equivocal about their lasting impact (Makopoulou et al., 2021). The literature supports the argument that a longer duration of exposure, combined with meaningful follow-ups and ample time for practice, intensifies the impact of professional development interventions, especially for science teachers (Conradty & Bogner, 2020; Li et al., 2021).

Monitoring

The absence of follow-up data made it difficult to document the short-term impacts of the training program. To inform decision-making, particularly when the training program is delivered to beneficiaries with different contents, quality, or duration, timely information is essential (Gertler et al., 2016). This involves tracking inputs, activities, outputs, and outcomes, which allows identifying how outcomes have changed over time and the effects of all factors changing over time (White & Raitzer, 2017).

It is important to note that this lesson learning framework is not comprehensive. However, the lessons learned captured and documented in this evaluation can be

applied to develop a culture of continuous improvement in implementing extension programs in HEIs in the Philippines in the long run.

Conclusion

The study assessed the impact of the teacher-training program implemented at a university in the Bicol Region at four levels. Existing documents from teacher-training program implementers were used to assess reaction, learning, and behavior levels. However, limitations in data precluded analysis at learning and behavior levels. On the other hand, contribution analysis was used to evaluate the results level.

Findings from document review reveal several gaps in the university's existing monitoring and evaluation practices. In the case of the teacher-training program, the training impact was insufficiently evaluated using satisfaction surveys that only addressed the reaction level of assessment based on Kirkpatrick's model. However, final outcomes were assessed five years after the program's implementation. Results showed that positive organizational outcomes occurred during the evaluation period but without explicit reference to the teacher-training program provided by the university. Furthermore, several facilitating and inhibiting factors were reported during the interview. The impact study sheds light on the relevance of teacher support that will complement teacher-training programs. Although the training was provided and policies supporting school gardens and antimicrobial stewardship were in place, resources that will materialize these policies into practice were lacking. While these factors can be a possible reason why final outcomes turned out to be below expectations, the sub-optimal impacts could have been minimized if appropriate evaluation frameworks were in place before the training program was provided. Furthermore,

given the constraints, there is a need to rethink how teacher-training programs are being implemented by going beyond the traditional short-term professional development efforts that lack opportunity for practice and follow-up support. It would be practical to provide professional development programs in lower-cost forms, such as implementing a more targeted approach in selecting participants who can serve as coaches to their peer teachers.

The evaluation experience of the training program assessed in this study shows that there are still many difficulties in making an evaluation as a tool for optimizing the benefits of training extensions. These challenges encountered suggest improvements in reporting extension programs in higher education by complying with specific evaluability criteria. This ensures that evaluations at the final stage are carried out in a meaningful and effective manner.

There are a number of limitations to this study. First, there was a paucity of data to further analyze the impact of the teacher-training. Most of the teachers were working under the Work From Home (WFH) scheme during the data collection. Additionally, it was not practical to assess the impact on the part of their students during the evaluation period since teachers cannot provide reliable contact details. Hence, impact reported in the study only focused on organizational changes.

All of these limitations suggest extreme caution in providing interpretations from the findings of the impact study. The evidence identified in the study was not statistically representative of all beneficiaries. However, the narrative causal statements collected generated insights into the factors that influenced the perceived organizational outcomes investigated in the study.

Recommendations

Based on the results and lessons learned from the evaluation, this study recommends the appraisal of teacher-training programs and other similar undertakings in HEIs, with evaluability criteria and standards. The Kirkpatrick model provides a simple framework applicable to extension guidelines governing training programs. Furthermore, extension providers should incorporate a prospective evaluation framework during program design for policy advocacy. This will allow the program to establish baseline data before the implementation, generate valid counterfactuals, and plan follow-up surveys for monitoring. Finally, it is best to use qualitative evidence in conjunction with quantitative estimates of changes. In the case of QuIP, narrative accounts can complement experimental and quasi-experimental models.



References

- Abeberese, A.B., Kumler, T.J., & Linden, L.L. (2014). Improving reading skills by encouraging children to read in school: A randomized evaluation of the Sa Aklat Sisikat Reading Program in the Philippines. *Journal of Human Resources*, 49(3), 611–633. <https://doi.org/10.3368/jhr.49.3.611>
- Acharya, K. P., Budhathoki, C. B., Bjønness, B., & Bhimsen, D. (2020). School gardening activities as contextual scaffolding for learning science: Participatory action research in a community school in Nepal. *Educational Action Research*, 1–18. <https://doi.org/10.1080/09650792.2020.1850494>

- Albornoz, F., Anauati, M. V., Furman, M. G., Luzuriaga, M., Podestá, M.E., & Taylor, I. (2019). *Training to teach science: Experimental evidence from Argentina* (Policy Research Working Paper 8594). World Bank Group. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/992241537378173492/training-to-teach-science-experimental-evidence-from-argentina>
- Askarian, M., & Maharlouie, N. (2012). Irrational antibiotic use among secondary school teachers and university faculty members in Shiraz, Iran. *International Journal of Preventive Medicine*, 3(12), 839–845. <https://doi.org/10.4103/2008-7802.104854>
- Azevedo, M., Pinheiro, C., Yaphe, J., & Baltazar, F. (2013). Assessing the impact of a school intervention to promote students' knowledge and practices on correct antibiotic use. *International Journal of Environmental Research and Public Health*, 10(7), 2920–2931. <https://doi.org/10.3390/ijerph10072920>
- Burt, K. G., Luesse, H. B., Rakoff, J., Ventura, A., & Burgermaster, M. (2018). School gardens in the United States: Current barriers to integration and sustainability. *American Journal of Public Health*, 108(11), 1543–1549. <https://doi.org/10.2105/ajph.2018.304674>
- Cahapay, M. B. (2021). Kirkpatrick model: Its limitations as used in higher education evaluation. *International Journal of Assessment Tools in Education*, 8(1), 135–144. <https://doi.org/10.21449/ijate.856143>
- Commission on Higher Education. (2016). *Pathways to equity, relevance, and advancement in research, innovation, and extension in Philippine higher education* (CMO No. 52, s. 2016). <https://ched.gov.ph/wp-content/uploads/2017/10/CMO-52-s.-2016.pdf>
- Conradty, C., & Bogner, F.X. (2020). STEAM teaching professional development works: Effects on students'

creativity and motivation. *Smart Learning Environments*, 7(26). <https://doi.org/10.1186/s40561-020-00132-9>

- Copestake, J., & Remnant, F. (2014). *Assessing rural transformations: Piloting a qualitative impact protocol in Malawi and Ethiopia* (Working Paper No. 35). The University of Bath, Centre for Development Studies (CDS). <https://researchportal.bath.ac.uk/en/publications/assessing-rural-transformations-piloting-a-qualitative-impact-pro>
- Dange, J. K., & Siddaraju. (2020). Role of teacher training program in enhancing quality education. *International Journal of Education, Culture and Society*, 5(6), 137-140. <https://doi.org/10.11648/j.ijecs.20200506.14>
- Diaconu, D. V., Radigan, J., Suskavcevic, M., & Nichol, C. (2012). A multi-year study of the impact of the rice model teacher professional development on elementary science teachers. *International Journal of Science Education*, 34(6), 855-877. <https://www.tandfonline.com/doi/abs/10.1080/09500693.2011.642019>
- Dogan, S., Pringle, R., & Mesa, J. (2016). The impacts of professional learning communities on science teachers' knowledge, practice and student learning: A review. *Professional Development in Education*, 42(4), 569-588. <https://doi.org/10.1080/19415257.2015.1065899>
- Fuje, H., & Tandon, P. (2018). *When do in-service teacher training and books improve student achievement? Experimental evidence from Mongolia* (Policy Research Working Papers). World Bank Group. <https://doi.org/10.1596/1813-9450-7485>
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice* (2nd ed.). Inter-American Development

- Bank and World Bank. <https://openknowledge.world-bank.org/handle/10986/25030>
- Gozalbo, M., Aragón, L., & Cubero, I. (2020). Gardens as science learning contexts across educational stages: Learning assessment based on students' graphic representations. *Frontiers in Psychology, 11*. <https://doi.org/10.3389/fpsyg.2020.02226>
- Grossman, R., & Salas, E. (2011). The transfer of training: What really matters. *International Journal of Training and Development, 15*(2), 103–120. <https://doi.org/10.1111/j.1468-2419.2011.00373.x>
- Gulamhussein, A. (2013). *Teaching the teachers: Effective professional development in an era of high stakes accountability* (Center for Public Education Report). National School Board Association. <http://conference.ohioschoolboards.org/2017/wp-content/uploads/sites/17/2016/07/1pm111317A114Job-embedPD.pdf>
- Huys, N., De Cocker, K., De Craemer, M., Roesbeke, M., Cardon, G., & De Lepeleere, S. (2017). School gardens: A qualitative study on implementation practices. *International Journal of Environmental Research and Public Health, 14*(12). <https://doi.org/10.3390/ijerph14121454>
- Japan International Cooperation Agency. (n.d.). *Manual on training evaluation*. https://www.jica.go.jp/project/cambodia/0601331/pdf/english/5_TrainingEvaluation.pdf
- Jocz, A. J., Zhai, J., & Tan, A. L. (2014). Inquiry learning in the Singaporean context: Factors affecting student interest in school science. *International Journal of Science Education, 36*(15), 2596–2618. <https://doi.org/10.1080/09500693.2014.908327>
- Jukes, M. C., Turner, E. L., Dubeck, M. M., Halliday, K. E., Inyega, H. N., Wolf, S., Zuilkowski, S. S., & Brooker,

- S. J. (2017). Improving literacy instruction in Kenya through teacher professional development and text messages support: A cluster randomized trial. *Journal of Research on Educational Effectiveness*, 10(3), 449–481. <https://doi.org/10.1080/19345747.2016.1221487>
- Kotwani, A., Watal, C., Joshi, P. C., & Holloway, K. (2016). Knowledge and perceptions on antibiotic use and resistance among high school students and teachers in New Delhi, India: A qualitative study. *Indian Journal of Pharmacology*, 48(4), 365–371. <https://doi.org/10.4103/0253-7613.186208>
- Landry, M., Van den Berg, A., Hoelscher, D., Asigbee, F., Vandyousefi, S., Ghaddar, R., Jeans, M., Waugh, L., Nikah, K., Sharma, S., & Davis, J. (2021). Impact of a school-based gardening, cooking, nutrition intervention on diet intake and quality: The TX sprouts randomized controlled trial. *Nutrients*, 13(9), 3081. <https://doi.org/10.3390/nu13093081>
- Lee, C., Lee, J., Kang, L., Jeong, B., & Lee, S. (2015). Educational effectiveness, target, and content for prudent antibiotic use. *BioMed Research International*, 1–13. <https://doi.org/10.1155/2015/214021>
- Li, N., Liu, E., Liu, C., and Guo, S. (2021). Rethinking the factor of duration for professional development: A workshop-seminar-demonstration class model for science teachers. *EURASIA Journal of Mathematics, Science and Technology Education*, 17(12). <https://doi.org/10.29333/ejmste/11351>
- Loftus, L., Spaulding, A. D., Steffen, R., Kopsell, D., Nnawe, N. (2017). Determining barriers to use of edible school gardens in Illinois. *Journal of the American College of Nutrition*, 36(7), 507–513. <https://doi.org/10.1080/07315724.2017.1326323>

- Lucas, A. M., McEwan, P. J., Ngware, M., & Oketch, M. (2014). Improving early grade literacy in East Africa: Experimental evidence from Kenya and Uganda. *Journal of Policy Analysis and Management*, 33(4), 950–976. <https://doi.org/10.1002/pam.21782>
- Makopoulou, K., Neville, R., Ntoumanis, N., & Thomas, G. (2021). An investigation into the effects of short-course professional development on teachers' and teaching assistants' self-efficacy. *Professional Development in Education*, 47(5), 780-795. <https://doi.org/10.1080/19415257.2019.1665572>
- Marvasi, M., Casillas, L., Vassallo, A., & Purchase, D. (2021). Educational activities for students and citizens supporting the one-health approach on antimicrobial resistance. *Antibiotics*, 10(12). <https://doi.org/10.3390/antibiotics10121519>
- Mayne, J. (2008). *Contribution analysis: An approach to exploring cause and effect (ILAC Brief 16)*. Institutional Learning and Change. <https://cgspace.cgiar.org/handle/10568/70124>
- National Economic and Development Authority. (2018). *2017-2022 Public Investment Program (NEDA Report)*. <https://neda.gov.ph/wp-content/uploads/2017/01/PIP-2017-2022-as-of-10042018.pdf>
- Ohly, H., Gentry, S., Wigglesworth, R., Bethel, A., Lovell, R., & Garside, R. (2016). A systematic review of the health and well-being impacts of school gardening: Synthesis of quantitative and qualitative evidence. *BMC Public Health*, 16(286), 1-36. <https://doi.org/10.1186/s12889-016-2941-0>
- Popova, A., Evans, D.K., & Arancibia, V. (2016). *Training teachers on the job: What works and how to measure it* (Policy Research Working Paper 7834). The World Bank. <http://hdl.handle.net/10986/25150>

- Popova, A., Evans, D.K., Breeding, M., & Arancibia, V. (2022). Teacher professional development around the world: The gap between evidence and practice. *The World Bank Research Observer*, 37(1), 107-136. <https://doi.org/10.1093/wbro/lkab006>
- Remnant, F., & Avard, R. (2016). *Qualitative Impact Assessment Protocol (QUIP)*. BetterEvaluation. <http://betterevaluation.org/en/plan/approach/QUIP>
- Rogers, P., Hawkins, A., McDonald, B., Macfarlan, A., & Milne, C. (2015). *Choosing appropriate designs and methods for impact evaluation* (Technical Report). Office of the Chief Economist, Australian Government. <https://www.industry.gov.au/data-and-publications/choosing-appropriate-designs-and-methods-for-impact-evaluation>
- Rooij, V., Williams, S., & Jerusalem, M. (2015). Measuring the business impact of employee learning: A View from the professional services sector. *Human Resource Development Quarterly*, 26(3), 275–297. <https://doi.org/10.1002/hrdq.21211>
- Schaffner, J., Glewwe, P., & Sharma, U. (2021). *Evaluation of secondary school teacher training under the School Sector Development Programme in Nepal* (Impact Evaluation Report 135). International Initiative for Impact Evaluation (3ie), New Delhi. <https://doi.org/10.23846/PW3IE135>
- Schreinemachers, P., Ouedraogo, M., Diabougua, S., Thiombiano, A., Kouamé, S., Sobgui, C., Chen, H., & Yang, R. (2019). Impact of school gardens and complementary nutrition education in Burkina Faso. *Journal of Development Effectiveness*, 11(2), 132-145. <https://doi.org/10.1080/19439342.2019.1624595>
- Smidt, A., Balandin, S., Sigafos, J., & Reed, V. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. *Journal of Intellectual &*

Developmental Disability, 34(3), 266–274. <https://doi.org/10.1080/13668250903093125>

Thong, K., Chang, C., Lee, M., Lee, J., Tan, H., & Shafie, A. (2021). Impact of targeted educational intervention towards public knowledge and perception of antibiotic use and resistance in the state of Perak, Malaysia. *Antimicrobial Resistance & Infection Control*, 10(1). <https://doi.org/10.1186/s13756-021-00892-0>

Throgmorton, C., Mitchell, T., Morley, T., & Snyder, M. (2016). Evaluating a physician leadership development program – A mixed methods approach. *Journal of Health Organization and Management*, 30(3), 390–407. <https://doi.org/10.1108/JHOM-11-2014-0187>

Thurlings, M., & Brok, P. (2017). Learning outcomes of teacher professional development activities: A meta-study. *Educational Review*, 69(5), 554–576, <https://doi.org/10.1080/00131911.2017.1281226>

United Nations Development Programme. (2021). *UNDP Evaluation Guidelines (UNDP Technical Report)*. Independent Evaluation Office of UNDP. http://web.undp.org/evaluation/guideline/documents/PDF/UNDP_Evaluation_Guidelines.pdf

Wallo, A., Kock, H., Lundqvist, D., & Coetzer, A. (2020). Understanding factors that enable and inhibit assessment of outcomes of competence development. *Human Resource Development Review*, 19(4), 384–421. <https://doi.org/10.1177/1534484320943332>

White, H., & Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: Towards an integrated framework* (3ie Working Paper 15). International Initiative for Impact Evaluation. https://www.3ieimpact.org/sites/default/files/2019-01/working_paper_15.pdf

- White, H., & Raitzer, D. (2017). *Impact evaluation of development interventions: A practical guide*. Asian Development Bank. <http://dx.doi.org/10.22617/TCS179188-2>
- Zhang, L., Lai, F., Pang, X., Yi, H., & Rozelle, S. (2013). The impact of teacher training on teacher and student outcomes: Evidence from a randomised experiment in Beijing migrant schools. *Journal of Development Effectiveness*, 5(3), 339–358. <https://doi.org/10.1080/19439342.2013.807862>