

Evaluation Methodologies in University Extension: A Comparative Analysis of Central American Practices

Vivian Harberts , Myrthe Reinsberg 

Institute Technology and Education, University of Bremen, Germany.

How to cite: Harberts, V.; Reinsberg, M. (2026). Evaluation Methodologies in University Extension: A Comparative Analysis of Central American Practices. In: 12th International Conference on Higher Education Advances (HEAd'26). Valencia, 15-18 June 2026. <https://doi.org/10.4995/HEAd26.2026.21553>

Abstract

University extension programs face persistent challenges in demonstrating social impact, particularly in resource-constrained contexts. This study develops a systematic eleven-field documentation matrix for comparing evaluation practices and applies it to 24 purposively-selected university extension programs in Central America. We analyzed evaluation approaches across five dimensions: responsible agents, temporal structures, evaluation objects, methodological frameworks, and instruments employed. Three key patterns emerged: 29.2% of programs adopted hybrid models combining internal oversight with external validation; methodological approaches distributed relatively evenly across participatory integral (29.2%), operational follow-up (29.2%), systematic institutional (25%), and formative-educational frameworks (16.7%); one-third developed digital evaluation tools despite resource constraints. However, only half systematically assessed long-term impacts, and genuine participatory partnership (beyond consultation) remained rare. These findings demonstrate the framework's utility for systematic comparison while identifying capacity development needs in resource-constrained contexts.

Keywords: Central America; Evaluation methodologies; Social impact; University Extension; Vulnerable Populations

1. Introduction

Over the recent decades, the relationship between universities and society has broadened beyond unidirectional service delivery: universities are increasingly expected to engage in deeper, mutually beneficial knowledge and societal engagement rather than simply providing services to communities (Benneworth & Osborne, 2014). This evolution creates new demands for evaluation methodologies capable of capturing complex, long-term social changes rather than simply counting outputs or measuring satisfaction.

These evaluation challenges intensify in regions facing significant socioeconomic inequalities like in Central America. Universities in this region serve communities experiencing high poverty rates, limited access to basic services, and ongoing struggles with social exclusion. At the same time, these institutions often lack the extensive financial resources available to universities in wealthier contexts. Thus, the question arises: how do universities extensions develop rigorous evaluation approaches for their programs under such constraints?

Recent research suggests that resource limitations may actually drive methodological innovation. Renick et al. (2025) found that universities facing budget pressures often develop creative approaches to execution and evaluation of their programs, combining formal assessment with community-based knowledge production. Similarly, review of research impact assessment highlights that impact is a highly complex and contested concept, with considerable methodological diversity across universities, research organizations, and funding agencies. Rather than relying on universal metrics, impact assessment tools are often tailored to specific missions and contexts, suggesting that universities must adapt evaluation frameworks to their own goals and environments rather than importing standardized models (Pedersen et al. 2020; Patton 2024).

This paper makes three contributions. First, methodologically, we develop an eleven-field documentation framework enabling systematic comparison of evaluation practices across diverse institutional contexts. Second, empirically, we apply this framework to 24 purposively-selected university extension programs from 10 Central American universities, revealing patterns in evaluation responsibility structures, methodological approaches, and temporal patterns. Third, practically, we identify capacity gaps (particularly in impact assessment) and innovations (notably digital tool development) relevant to universities seeking to strengthen extension evaluation within resource constraints. The analysis reveals considerable diversity in evaluation practices while identifying recurring patterns. Findings are organized by the five evaluative process dimensions (who, when, what, how, with what). Our analysis reveals both promising innovations and persistent struggles, offering insights relevant to universities globally who seek to strengthen their extension evaluation practices.

2. Methodology

This study draws on data from the FUESE project (Strengthening Extension in the Social Ecosystem of Universities in Central America), funded through Erasmus+ CBHE. The project involves 12 Central American universities across Panama, Costa Rica, Honduras, El Salvador, and Guatemala, working alongside partner institutions from Spain and Germany.

A core project activity involved systematic documentation of university extension practices across participating institutions. Each university identified and documented reference practices according to shared criteria, creating a database of programs that could be analyzed comparatively.

2.1. Program Selection Criteria

Universities selected programs meeting six mandatory requirements:

- Service to objectively vulnerable populations (rural women, indigenous communities, or areas with low human development indices)
- Documented needs assessment justifying the intervention
- Clear objectives articulated in accessible formats
- Systematic record-keeping (participant databases, activity logs, progress reports)
- Designated staff responsible for maintaining documentation
- Designated staff responsible for coordinating evaluation activities

These criteria ensured that documented programs had sufficient structure and documentation to permit meaningful evaluation analysis, while maintaining focus on social equity concerns.

This study employs purposive sampling focused on well-documented programs rather than representative sampling. The six selection criteria prioritized programs with sufficient documentation infrastructure—specifically, designated documentation staff, designated evaluation staff, and systematic record-keeping. These requirements systematically excluded programs lacking institutional infrastructure. Consequently, these 24 programs represent reference cases demonstrating different evaluation approaches rather than a statistically representative sample of all Central American extension programs. Findings should be interpreted as patterns observed in well-documented cases rather than definitive statements about Central American university extension generally.

2.2. Analytical Framework

We analyzed programs using an eleven-field matrix organized around three analytical dimensions. The first dimension captured basic program identification (institution, program name, duration). The second dimension examined evaluative processes through five analytical dimensions: what gets evaluated, how evaluation occurs, when evaluation happens, who conducts evaluation, and what instruments they use. The third dimension documented results at three levels—immediate products, medium-term effects, and long-term impacts.

This five-dimension framework allowed systematic comparison across diverse programs while remaining flexible enough to accommodate different institutional contexts and evaluation philosophies. Table 1 illustrates the analytical dimensions that structured our comparative analysis.

Table 1. Five-Dimension Framework for Evaluation Practices (Source: Proyecto FUESE 2025b).

Dimenion	Key Question	Components
WHO? (Agents)	Who conducts the evaluation?	Internal staff; External evaluators; Beneficiary communities
WHEN? (Timing)	When does evaluation occur?	Systematic periodic cycles; Project-phase milestones; Continuous monitoring
WHAT? (Objects)	What is evaluated?	Outputs (activities, products); Outcomes (behavior, knowledge changes); Impacts (structural transformation)
HOW? (Approaches?)	What methodological approach is used?	Participatory frameworks; Mixed methods; Systematic institutional approaches
WITH WHAT? (Instruments)	What instruments are used?	Institutional forms; Quantitative-qualitative tools; Participatory techniques; Digital technologies

2.3. Sample Characteristics

Data collection occurred between January and September 2025 through a standardized protocol. Each participating university assigned a documentation team responsible for completing the eleven-field matrix for their selected programs. Teams gathered data through three sources: institutional records and databases, interviews with program coordinators and evaluation staff, and review of evaluation reports and instruments.

The final dataset included 24 programs distributed across 10 universities in five countries (Table 2). Programs spanned six thematic areas: health and nutrition (5 programs), environmental sustainability (6 programs), economic development (3 programs), education and capacity-building (4 programs), legal and social services (3 programs), and local development (3 programs).

Table 2. Sample Distribution Across Countries and Institutions. (Source: Proyecto FUESE 2025a).

Country	University	Analysed Programs (n)
Costa Rica	Universidad Nacional (UNA)	2
Costa Rica	Universidad Estatal a Distancia (UNED)	2
El Salvador	Universidad Gerardo Barrios (UGB)	2
El Salvador	Universidad de El Salvador (UES)	2
Guatemala	Universidad de San Carlos (USAC)	2
Guatemala	Universidad Regional de Guatemala (URG)	5
Honduras	Universidad Nacional Autónoma (UNAH)	3
Honduras	Universidad Pedagógica Nacional (UPNFM)	2
Panama	Universidad Especializada de las Américas (UDELAS)	2
Panama	Universidad Autónoma de Chiriquí (UNACHI)	2
Total	5 Countries, 10 Universities	24

This distribution provided geographic breadth while providing us with sufficient depth to identify patterns within individual institutions.

2.4. Analytical Approach

Qualitative content analysis was conducted following Kuckartz (2018) for systematic categorization of textual data. This method is appropriate for analyzing structured documentation across multiple cases to identify patterns and compare approaches.

The analysis involved three steps: First, data familiarization. Systematically all documented practices have been reviewed, examining how each of the 24 programs addressed the five evaluative process dimensions. This initial review provided an overview of the diversity in evaluation approaches.

Second, category development. For each dimension, programs sharing similar characteristics have been grouped. For example, within the “who evaluates” dimension, we identified programs that relied solely on internal university staff, those using external evaluators, those involving students, and those including community members. Categories emerged from the data itself rather than being predetermined. We iteratively refined categories until each one clearly distinguished between different approaches while ensuring all programs could be classified.

Third, frequency analysis and interpretation. We counted how many programs fell into each category, calculating percentages to identify dominant patterns and outliers. Also, examination has been applied whether certain approaches tended to co-occur—for instance, whether participatory methodologies typically involved beneficiaries as evaluators.

This analysis describes evaluation practices without judging their quality. It documents what institutions do but does not rate how well they do it. No scoring, weighting, or ranking was applied. The matrix served as a documentation framework, not an assessment rubric.

3. Results

The analysis of the 24 programs reveals considerable diversity in evaluation practices but also identifies recurring patterns across institutions.

Evaluation responsibility structures showed five distinct organizational models. The most common approach (29.2% of programs, $n=7$) combined internal university staff with external evaluators—Universidad Nacional conducts annual internal evaluations while commissioning external experts for quinquennial assessments. This hybrid structure combines institutional control with external validation. A quarter of programs (25%, $n=6$) assigned evaluation to specialized institutional coordination units, while 20.8% ($n=5$) placed responsibility under senior academic leadership. Faculty-student teams conducted evaluation in 16.7% of programs ($n=4$), and individual specialized evaluators were designated in 8.3% ($n=2$). Notably, only one-third of programs systematically included program beneficiaries in evaluation activities.

Also temporal evaluation patterns were identified across four approaches. Systematic periodic evaluation emerged as the dominant structure (37.5%, $n=9$), with universities establishing regular semester, annual, or quinquennial cycles. Universities establishing regular cycles treat assessment as ongoing institutional function rather than project-specific requirement. Nearly 30% (29.2%, $n=7$) aligned evaluation with project phases—diagnostic, formative, summative, and deferred impact assessments. One-fifth (20.8%, $n=5$) conducted evaluation at specific moments determined by program characteristics rather than fixed schedules, while 12.5% ($n=3$) embedded evaluation continuously within academic calendars.

We found that the evaluated programs clustered into four thematic areas reflecting Central American development priorities. Health, wellbeing, and social problems represented the largest category (33.3%, $n=8$), including maternal health, childhood nutrition, psychological services, and legal aid. Environmental management and sustainability comprised a quarter of programs (25%, $n=6$), addressing climate adaptation, reforestation, and waste management. Economic and productive development (20.8%, $n=5$) focused on value chains, market systems, and entrepreneurship. Institutional strengthening and capacity building also represented 20.8% ($n=5$), emphasizing municipal governance, professional training, and emerging technologies.

Methodological frameworks revealed four approaches with near-equal distribution. Participatory integral evaluation represented 29.2% ($n=7$), emphasizing beneficiary involvement in evaluation design and implementation—though depth of participation varied considerably. Operational follow-up evaluation also comprised 29.2% ($n=7$), focusing on monitoring implementation and measuring specific results. Systematic institutional evaluation accounted for 25% ($n=6$), characterized by formalized procedures with external validation. Formative-educational evaluation represented 16.7% ($n=4$), integrating assessment with learning processes and competency development.

We found five categories of evaluation instruments. Institutionalized administrative instruments (25%, $n=6$) included standardized progress report templates and evaluative matrices. Hybrid quantitative-qualitative tools also represented 25% ($n=6$), deliberately combining surveys with interviews and observations. Technical-operational tracking instruments (20.8%, $n=5$) emphasized productivity indicators and implementation monitoring. Participatory-qualitative tools (20.8%, $n=5$) employed community workshops, photographic records, and focus groups. Formative-academic instruments (8.3%, $n=2$) included pre/post competency assessments and certification criteria. Remarkably, eight programs (33%) incorporated digital technologies—analytical dashboards, mobile applications, electronic databases—despite resource constraints.

Three patterns could be observed. Programs serving indigenous or rural populations demonstrated stronger participatory orientation and culturally adapted instruments. Programs with hybrid internal-external responsibility structures more frequently attempted impact assessment, possibly because external evaluators brought specialized attribution expertise. Digital tool

adoption appeared across institutions of varying sizes. Eight programs (33%) incorporated digital technologies, including both larger research universities and smaller regional institutions. However, without systematic data on institutional budgets or student enrollment, definitive conclusions about relationships between institutional characteristics and technological capacity cannot be drawn.

The most significant gap concerned participatory practice versus rhetoric. While 29.2% described approaches as participatory, only two programs achieved genuine partnership where beneficiaries shaped evaluation questions and interpreted findings. Most programs limited involvement to consultation. Similarly, only half (50%) attempted systematic long-term impact assessment, focusing instead on immediate outputs or medium-term outcomes rather than structural transformation.

4. Conclusion

Central American universities are developing evaluation approaches adapted to challenging contexts while maintaining methodological seriousness. Several practical implications emerge. First, institutions should distinguish between consultation (asking for input), collaboration (working jointly), and partnership (sharing authority) rather than labeling all approaches "participatory." Second, the impact assessment gap reflects capacity constraints more than knowledge deficits—regional peer learning networks could address this more effectively than external consultants. Third, digital innovations show that sophisticated evaluation need not require expensive software, though successful adoption requires complementary training. Fourth, methodological diversity represents appropriate contextual adaptation, not inconsistency—different programs require different approaches while maintaining basic quality standards. Finally, meaningful impact assessment requires timeframes exceeding typical project cycles; universities might pool resources for longitudinal studies rather than expecting individual projects to demonstrate long-term transformation within limited durations.

This study has several important limitations that readers should consider when interpreting findings. First, the purposive sampling strategy prioritized well-documented programs over representativeness. The six mandatory selection criteria—particularly requirements for designated documentation and evaluation staff—systematically excluded programs lacking institutional infrastructure. These 24 cases likely represent higher-performing programs with stronger evaluation capacity rather than typical extension practice across Central America. Claims about patterns should be understood as observations from well-documented reference cases rather than statements about regional prevalence. Second, this document-based analysis examines reported evaluation practices rather than observed implementation. Programs may describe participatory approaches in documentation while implementing them quite differently in practice. The gap between stated participatory orientation (29.2%) and genuine partnership (two programs)

illustrates this challenge. Future research combining document analysis with observational studies and stakeholder interviews could reveal discrepancies between documented and enacted practices. Third, the small sample size ($n=24$, with subcategories as small as $n=2$) limits generalization. Fourth, cross-sectional data capture practices at one moment rather than tracking evolution over time. Evaluation systems develop gradually; understanding these developmental trajectories requires longitudinal research designs beyond this study's scope.

Finally, document-based analysis cannot assess evaluation quality or effectiveness—only reported practices. A program may employ rigorous external evaluation or implement tokenistic participation; both appear similar in documentation. Quality assessment requires examining evaluation reports, stakeholder interviews, and actual evaluation processes rather than program documentation alone. Central American universities face enormous challenges serving vulnerable populations with limited resources. The evaluation practices documented here suggest institutions are developing serious responses combining rigor with adaptation

References

- Benneworth, P., & Osborne, M. (2014). Knowledge engagement and higher education in Europe. In C. Escribas (Ed.), *Higher education in the world 5: Knowledge, engagement and higher education: Contributing to social change* (pp. 219–231). Palgrave Macmillan.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Beltz Juventa.
- Patton, M. Q., & Campbell-Patton, C. E. (2024). Using principles to guide evaluation theory and practice. In K. E. Newcomer & S. W. Mumford (Eds.), *Research handbook on program evaluation* (pp. 137–160). Edward Elgar Publishing.
- Pedersen, D. B., Grønvad, J. F., & Hvidtfeldt, R. (2020). *Methods for mapping the impact of social sciences and humanities: A literature review*. Research Evaluation, 29(1), 4–21. <https://doi.org/10.1093/reseval/rvz033>
- Proyecto FUESE (2025a). *Banco web con 28 prácticas de referencia*. Informe, Panamá.
- Proyecto FUESE (2025b). *Protocolo e instrumentos para la identificación y selección de prácticas de impacto social*. Informe, Panamá.
- Renick, J., Furco, A., Janke, E.M. et al. (2025). *More than Just Money: Strategic Allocation of Resources to Support Community-Engaged Scholarship*. Innov High Educ. <https://doi.org/10.1007/s10755-025-09831-y>
- Universidad Especializada de las Américas. (2023). *FUESE-Fortalecimiento de la Extensión en el Ecosistema Social de las universidades en Centro América*. Panamá