José Luis Escamilla Reyes
(organizador)

EDUCAÇÃO E ENSINO DE CIÊNCIAS EXATAS E NATURAIS

VOL II

EDITORA ARTEMIS
2024

José Luis Escamilla Reyes
(organizador)

EDUCAÇÃO E ENSINO DE CIÊNCIAS EXATAS E NATURAIS

VOL II

EDITORA ARTEMIS
2024

Editora Chefe	Profª Drª Antonella Carvalho de Oliveira
Editora Executiva	M.ª Viviane Carvalho Mocellin
Direção de Arte	M.ª Bruna Bejarano
Diagramação	Elisangela Abreu
Organizador	Prof. Dr. José Luis Escamilla Reyes
Imagem da Capa	ekaart/123RF
Bibliotecário	Maurício Amormino Júnior – CRB6/2422

EDITORA ARTEMIS
2024

Editora Artemis
Curitiba-PR Brasil
www.editoraartemis.com.br
e-mail:publicar@editoraartemis.com.br

Editora Artemis
Curitiba-PR Brasil
www.editoraartemis.com.br
e-mail:publicar@editoraartemis.com.br

Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)

E24 Educação e ensino de ciências exatas e naturais II [livro eletrônico] / Organizador José Luis Escamilla Reyes. – Curitiba, PR: Artemis, 2024.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
Edição bilíngue
ISBN 978-65-81701-29-1
DOI 10.37572/EdArt_311024291

1. Educação. 2. Ciências exatas e naturais – Estudo e ensino. 3. Professores – Formação. I. Reyes, José Luis Escamilla.

CDD 371.72

Elaborado por Maurício Amormino Júnior – CRB6/2422

Editora Artemis
Curitiba-PR Brasil
www.editoraartemis.com.br
e-mail:publicar@editoraartemis.com.br

PRÓLOGO

En este volumen, se presentan los resultados de varios y diversos proyectos de investigación en innovación educativa relacionados con la enseñanza de las ciencias y la ingeniería, tanto en niveles universitarios como básicos. Es así como, a través de distintas experiencias, se aborda la enseñanza de la Física, la Química Analítica y la enseñanza de temas matemáticos tales como la Aritmética y el Álgebra. También, se explora la incorporación de nuevas alternativas como la Inteligencia Artificial y sus aplicaciones en la enseñanza de las ciencias, particularmente de la Química.

Adicionalmente, en este libro se discuten los procesos de evaluación, no sólo de las actividades realizadas por los alumnos en los diferentes niveles educativos, sino de la pertinencia y adecuación del currículum en las disciplinas científicas, dentro de las que se puede mencionar a la Química Analítica y las Ciencias Exactas en general.

Por supuesto, hago la invitación a nuestros lectores para que disfruten la lectura de estos artículos de innovación educativa y, si son docentes en activo, que implementen alguna o varias de las estrategias y metodologías expuestas en este volumen con el fin de enriquecer su práctica docente y, de esta manera, contribuir en la mejora de los procesos educativos desde los niveles básicos hasta los universitarios.

Finalmente, los autores de este libro agradeceremos la retroalimentación y los comentarios propositivos que nos hagan llegar, puesto que lo más importante es asegurar que nuestros alumnos tengan una educación de calidad y que logren un aprendizaje significativo que les permita superar con éxito los problemas tanto en su formación académica como en su vida cotidiana.

Dr. José Luis Escamilla Reyes

SUMÁRIO

NUEVAS PERSPECTIVAS EN LA ENSEÑANZA DE LAS CIENCIAS E INGENIERÍA

CAPÍTULO 6

ANALYZING THE USE OF THE KIRKPATRICK MODEL IN HIGHER EDUCATION: INSIGHTS FROM AN NSF-FUNDED CHEMISTRY CURRICULUM PROJECT

Data de submissão: 10/09/2024
Data de aceite: 30/09/2024

Dr. James Lipuma
Director of Collaborative for Leadership Education, and Assessment Research
New Jersey Institute of Technology
New Jersey, USA
https://orcid.org/0000-0002-9778-3843

Dr. Cristo Leon[1]
Director of Research
College of Science and Liberal Arts
New Jersey Institute of Technology
New Jersey, USA
https://orcid.org/0000-0002-0930-0179

ABSTRACT: This article delves into the unique development of a Kirkpatrick 4-Level Model (K4LM) evaluation plan for a groundbreaking National Science Foundation (NSF) project (ID#: EFMA-2203704). This project, conducted at an R1 Polytechnic Research University in the United States, integrated social, economic, environmental, and governance (SEEG) aspects into a chemistry curriculum. The aim was to promote environmental sustainability and the circular economy, a novel approach in the field. The project focused on the graduate and doctoral programs in the Chemistry department, making a significant contribution to higher education. The article began with an introduction to the project, followed by an analysis of the Kirkpatrick Model's application to program changes in higher education, discussing its benefits and criticisms. A methods section includes a literature review. It emphasized the importance of educational evaluation for course design, curriculum redesign, and student success. The practical case study, a key highlight of this article, illustrated the 4-level model's use in gathering actionable formative data. This data guided administrative decisions in transforming chemistry education, providing a tangible example of the model's practical application. Finally, challenges and solutions for applying this model in a systemic educational context were presented.

KEYWORDS: Interdisciplinary Chemistry Education. Kirkpatrick Model. ADDIE Approach. Sustainability Education. Circular Economy. Socio-Economic Governance. Environmental Challenges. Doctoral Pedagogy. Sustainable Development Goal 4 Quality Education (SDG 4). Transdisciplinary Communication.

[1] corresponding author

ANÁLISIS DEL USO DEL MODELO KIRKPATRICK EN LA EDUCACIÓN SUPERIOR: PERSPECTIVAS DE UN PROYECTO DE CURRÍCULO DE QUÍMICA FINANCIADO POR LA NSF

RESUMEN: Este artículo explora el desarrollo único de un plan de evaluación basado en el Modelo de 4 Niveles de Kirkpatrick para un proyecto innovador de la National Science Foundation (NSF) (ID#: EFMA-2203704). Este proyecto, llevado a cabo en una Universidad Politécnica de Investigación R1 en los Estados Unidos, integró aspectos sociales, económicos, ambientales y de gobernanza en un plan de estudios de química. El objetivo era promover la sostenibilidad ambiental y la economía circular, un enfoque novedoso en el campo. El proyecto se centró en los programas de posgrado y doctorado del departamento de Química, haciendo una contribución significativa a la educación superior. El artículo comenzó con una introducción al proyecto, seguida de un análisis de la aplicación del Modelo de Kirkpatrick a los cambios de programa en la educación superior, discutiendo sus beneficios y críticas. Una sección de métodos incluye una revisión de la literatura. Se enfatizó la importancia de la evaluación educativa para el diseño de cursos, el rediseño del currículo y el éxito estudiantil. El estudio de caso práctico, un aspecto destacado de este artículo, ilustró el uso del modelo de 4 niveles para recopilar datos formativos procesables. Estos datos guiaron las decisiones administrativas en la transformación de la educación en química, proporcionando un ejemplo tangible de la aplicación práctica del modelo. Finalmente, se presentaron desafíos y soluciones para aplicar este modelo en un contexto educativo sistémico.
PALABRAS CLAVE: Educación Interdisciplinaria en Química. Modelo Kirkpatrick. Enfoque ADDIE. Educación para la Sostenibilidad. Economía Circular. Gobernanza Socioeconómica. Desafíos Ambientales, Pedagogía Doctoral. Objetivo de Desarrollo Sostenible 4: Educación de Calidad (ODS 4). Comunicación Transdisciplinaria.

1 INTRODUCTION

This article, based on experience from a National Science Foundation (NSF) project (ID#: EFMA-2203704), examines elements that integrate social, economic, environmental, and governance (SEEG) aspects into a chemistry context to provide a foundation for understanding environmental sustainability and the circular economy. Integrating these ideas within the chemistry discipline aims to develop students' abilities to formulate new research questions addressing many of society's "grand challenges," particularly in environmental sustainability, climate change, and the circular economy. A two-year pilot study investigated a comprehensive strategy that includes new multidisciplinary courses, mentoring, graduate research proposals, seminars, and dissertation research. The NSF required using Kirkpatrick's 4-level evaluation model as part of the grant to provide data on student learning and program change as part of goal attainment.

The project was designed to measure potential participants' baseline knowledge about sustainability and the circular economy in a chemistry context and evaluate the value and effectiveness of external guest presentations. The findings were used to adjust the programming and inform the development of necessary materials and additional program support.

The research topic is developing an evaluation plan based on the Kirkpatrick Model that integrates social, economic, environmental, and governance aspects into chemistry. This evaluation assessed its impact on curriculum design at an R1 Polytechnic Research University in the USA.

The survey instrument was designed using the Kirkpatrick model of educational program evaluation and the NSF Self-Assessment of Learning Gains (SALG). Both are valid and reliable tools that have been used for many years. The data collection process aimed to gather information about students' baseline knowledge and evaluate the effectiveness of guest presentations and workshops.

The article begins with an overview of the Kirkpatrick Model and some of its benefits and criticisms within scholarly literature. It then discusses the importance of educational evaluation for course design, curriculum redesign, and overall student success in higher education institutions. This is followed by an explanation of why analyzing the application of Kirkpatrick from a system change perspective provides insights into how this training evaluation tool might be more effectively applied to the goals and outcomes sought in the U.S. educational system.

Next, the paper delves into a discussion of the practical case undertaken by the author to utilize the 4-level model to yield formative data that was actionable in informing the choices of the Chemistry department administration as it worked to transform graduate and doctoral education towards a more sustainable view of chemistry and the circular economy. As part of this discussion, the authors provide specific survey questions, along with the contextual factors and thought processes that influenced the development of the evaluation tools. Finally, the article presents challenges and solutions for applying this evaluation model in situations requiring a more systemic and holistic view of education, going beyond training or single courses to identify the interconnected nature of curriculum and instructional design (CID) needed to inform administrative decision-making and programmatic change initiatives. The article concludes with thoughts on the next steps that might be considered.

2 BRIEF OVERVIEW OF THE KIRKPATRICK MODEL

The Kirkpatrick Model first appeared in the 1950s and has been refined and reimagined several times, most recently in 2010. It systematically measures training in corporate settings to connect quantitative data to institutional goals and metrics. Over the years, it has been used in many domains, including education and program evaluation. It comprises four levels of investigation designed to gather data about different aspects of program implementation and efficacy to make systemic changes.[2]

The four levels look at varying time frames but are typically tied to an experience delivered to a target population.

- **Level 1 - Reaction**: This level surveys participants through observation and self-reported surveys to determine factors that can be adjusted to improve content delivery and increase knowledge transfer and internalization. In education, this is often seen as course evaluations. Still, Kirkpatrick training surveys are much more granular and intended to provide immediate and detailed feedback on ongoing training to understand its effectiveness and how it might be improved.
- **Level 2 - Learning**: This level assesses the knowledge gained by participants, akin to testing in an educational context.
- **Level 3 - Behavior**: This level examines how the experience changes participants' actions or work execution. In education, this can be seen in skills development, the transfer of training and knowledge to new contexts or scenarios, or the ability to tackle more advanced work based on prior learning. This level often requires more time, data collection, and a clear understanding of environmental and contextual factors that can impact behavioral change. Level three assessments are challenging to translate into specific classes as they typically do not follow students but result in instructional design, content delivery, or curriculum changes at the program level.
- **Level 4 - Results**: This level measures institutional or programmatic outcomes to see whether impacts meet metrics. The specific experiences are no longer as vital as the impacts shown by key performance indicators, metrics of success, or other larger institutional goals. These are typically measured in aggregate and not tied to a single data point or survey response.

Though presented from Level 1 to Level 4, the Kirkpatrick Model can also work backward, similar to what was described in *Understanding by Design* (Wiggins, 2005).

[2] Source: Yale University. (2019, January 10). Kirkpatrick Model. Poorvu Center for Teaching and Learning. https://poorvucenter.yale.edu/Kirkpatrick

In this approach, the outcomes are specified, metrics for the organization or program of study are identified, and milestones are built backward to drill down to specific tests of knowledge in the form of assessments and formative feedback based on classroom experiences and observations. This method is more often seen in K-12 education and not as much in higher education in the USA. Still, it works toward more authentic assessments to provide data as part of a portfolio of information about teaching and learning.

3 METHODS

The qualitative research approach utilized by the author while serving as the internal evaluator for the NSF project was Action Research for Informed Decision Making (Frey, 2018, p. 37). It was also informed by the ideas of Reflective Practice (Langer & Ghaye, 2012). This method allows him to reflect on his processes, actions, and products to provide insights without disclosing the project's data or other proprietary information that has not been reported publicly. This mixed-method approach leans heavily on qualitative methods for recording and analyzing meetings and deliverables from the project. These are paired with quantitative performance measures and the project outcomes for which the author was engaged as the evaluator.

The author drew heavily from three seminal texts in evaluations and assessment research: "Evaluation: A Systematic Approach" (Rossi et al., 2003), "The Practice of Social Science Research" (Babbie, 2016), and "Research Design: Qualitative and Quantitative Approaches" (Creswell & Creswell, 2018). These texts, along with many others in the author's career, allowed for the design and refinement of the K4LM used in the grant project and this critical analysis of that work to provide the sample instruments, insights, and suggestions below.

The analysis focused on the effectiveness of the curriculum redesign in a chemistry program. A systematic literature review was conducted to contextualize the methodology. The central question for the review was:

How has the application of the Kirkpatrick Model in higher education evolved, and what are its demonstrated strengths and limitations in assessing the effectiveness of educational programs compared to other evaluation frameworks?

3.1 SYSTEMATIC LITERATURE REVIEW

The application of the Kirkpatrick Model in higher education has significantly evolved, expanding its use across diverse educational contexts to assess the effectiveness of various programs. Initially developed for training evaluation in the corporate sector,

the Kirkpatrick Model has been adapted to meet the unique needs of higher education, demonstrating both strengths and limitations in this new context. This section reviews the evolution, strengths, and limitations of the Kirkpatrick Model in higher education compared to other evaluation frameworks.

3.2 METHODOLOGY

This systematic literature review (SLR) involved a comprehensive search of relevant databases using specific search terms related to the Kirkpatrick Model and higher education. The search strategy included PubMed, Scopus, and Google Scholar databases. The inclusion criteria focused on studies published in peer-reviewed journals from 2020 to 2024 that applied the Kirkpatrick Model in higher education settings. Exclusion criteria were studies outside this date range or those not directly relevant to higher education. Data extraction involved identifying the selected studies' key themes, strengths, and limitations.

The search string used is presented in Table 1:

Table 1. Bibliometric Search String.

("Kirkpatrick Model" OR "Kirkpatrick Evaluation Model" OR "Kirkpatrick Framework") AND ("higher education" OR "university" OR "college") AND ("evaluation" OR "assessment" OR "effectiveness") AND ("strengths" OR "advantages" OR "benefits") AND ("limitations" OR "challenges" OR "weaknesses" OR "drawbacks")

Results and Study Selection: The search strategy yielded numerous studies, of which a subset met the inclusion criteria. Fundamental studies included those by Paull, Whitsed, and Girardi (2020), Chang and Chen (2014), Akbar, Darungan, and Rahma (2024), Reio et al. (2017), and Matondang, Sitompul, and Wijaya (2023).

Study Characteristics: The selected studies applied the Kirkpatrick Model to various educational interventions, from curriculum frameworks to online learning environments. These studies spanned multiple disciplines, demonstrating the model's adaptability.

3.3 SYNTHESIS OF RESULTS

Several studies illustrate the Kirkpatrick Model's evolution in higher education. Paull, Whitsed, and Girardi (2020) used the model to evaluate an "interaction for learning framework" curriculum intervention, showing its utility across multiple levels. Chang and Chen (2014) applied the model to assess an online information literacy class's learning effectiveness, demonstrating its applicability in online settings. Akbar, Darungan, and

Rahma (2024) emphasized the model's comprehensive evaluation across four levels: reaction, learning, behavior, and results, highlighting its structured approach to assessing educational interventions.

3.4 STRENGTHS OF THE KIRKPATRICK MODEL

The primary strength of the Kirkpatrick Model lies in its multi-level evaluation framework. By examining reaction, learning, behavior, and results, the model provides a holistic view of educational programs' impact (Akbar et al., 2024). This approach captures immediate feedback and long-term outcomes, facilitating iterative program improvements. Paull et al. (2020) highlighted the model's ability to provide structured and quantifiable data, which is essential for rigorous evaluation and stakeholder reporting, such as to the NSF.

3.5 LIMITATIONS OF THE KIRKPATRICK MODEL

Despite its strengths, the Kirkpatrick Model has limitations in higher education. Reio et al. (2017) critiqued the model for its focus on lower levels (reaction and learning), which can neglect the more complex levels of behavior and results, leading to an incomplete picture of long-term program effectiveness. Chang and Chen (2014) noted that while the model evaluates learning quality effectively, it may lack the depth of analysis provided by other frameworks, which better address educational outcomes' nuances. Matondang, Sitompul, and Wijaya (2023) highlighted the challenge of applying the model to measure long-term outcomes, emphasizing the need for more rigorous research designs and standardized evaluation tools.

3.6 COMPARISON WITH OTHER EVALUATION FRAMEWORKS

The Kirkpatrick Model stands out for its structured and tiered approach compared to other evaluation frameworks. Alhassan (2022) emphasized the model's ability to assess the holistic impact of educational programs, making it valuable for training initiative evaluations. However, its rigidity and quantitative focus can overlook nuanced educational evaluation aspects that other frameworks might capture more effectively.

3.7 DISCUSSION

The Kirkpatrick Model's application in higher education has evolved to provide a comprehensive evaluation framework. Its strengths include a multi-level approach and the ability to generate quantifiable data, which is essential for program assessment and

improvement. However, limitations such as focusing on lower levels and challenges in measuring long-term outcomes suggest that integrating other evaluation frameworks might offer a richer understanding of educational program effectiveness.

3.8 IMPLICATIONS FOR PRACTICE AND RESEARCH

While the Kirkpatrick Model benefits from immediate and structured evaluations, a mixed-method approach incorporating qualitative data could enhance the depth of analysis. Future research should explore combining the Kirkpatrick Model with other frameworks to address its limitations and provide a more comprehensive evaluation of educational programs.

The Kirkpatrick Model has proven to be a valuable tool for evaluating educational programs in higher education, offering structured and quantifiable insights. However, a mixed-method approach that integrates other evaluation frameworks is recommended to achieve a more nuanced understanding of program effectiveness.

4 PRACTICAL CASE STUDY

4.1 SPONSOR AND PROJECT AIMS

The NSF Directorate for Engineering GERMINATION program aims to foster the development of pedagogical frameworks, platforms, and environments to enable participants to formulate research questions and ideas with potentially transformative outcomes (Giarratano, 2022).

The extraordinary response of the STEM research community to the COVID-19 pandemic, exemplified by the record-breaking speed of novel vaccine development, highlights the outstanding capabilities at all levels of the research enterprise. The GERMINATION program seeks to harness the immense capacities of academic researchers to similarly address other critical global challenges by supporting the development of new pedagogical approaches that train researchers to formulate and develop key research questions.[3]

Based on the premise that graduate chemistry education is traditionally highly disciplinary and often disconnected from society's critical challenges, the grant project sought a multifaceted approach to incorporate various perspectives into the doctoral chemistry curriculum. This was to be attained by incorporating guest speakers, presenters, and experts from a wide range of disciplines into the new courses, including a team-taught

[3] Source: GERMINATION: Germination of Research Questions for Addressing Critical Societal Challenges https://new.nsf.gov/funding/opportunities/germination-germination-research-questions

course that involves SEEG dimensions, a course in Green Chemistry, and changes to required dissertation research to include substantial SEEG components. The overarching aim was to increase consideration of and appreciation for institutional contexts, policy processes, and innovation opportunities to enable students to use their chemical knowledge more effectively for societal change, particularly with respect to sustainability and the circular economy (Clift et al., 2022). At the heart of this was the need to teach critical thinking skills and ask questions beyond those typically addressed in Chemistry.

This project challenged the traditional, disciplinary-focused pedagogical paradigm by piloting a more holistic approach to education and professional development that is purposefully aligned with the heavily integrated and interdependent world in which we live. Success in this exploratory study could lay the foundations for expansion to disciplines and institutions beyond the pilot base. The project required data on the interventions and formative data about how the implemented elements worked so changes could be made to optimize the work. This led to initial survey instruments focused on levels one and two of the 4KLM being created in conjunction with the researchers to identify the areas of learning and the critical content to be judged. Once in place within the classes and tested for two semesters, the project focused on the identified need for training in critical thinking and asking the right questions. These two efforts led to the series of survey prompts described below in this section. Throughout this process, informal discussions and work sessions were held between the author as an evaluator and the research team to refine and adjust the instruments within the K4LM to allow them to be debugged and optimized for ongoing use in the project.

The significant data gathering and incremental improvement came through arranging and holding the guest speakers and the knowledge baseline survey. New Jersey Institute of Technology's Institutional Review Board (IRB) Protocol number 2208024268 was reviewed and given exempt status. Excerpts of those instruments are provided, but as an ongoing study, the data and tools have not been reported to the sponsor NSF nor made public.

4.2 LEVEL 1 AND 2

The project's initial focus was establishing a baseline in two areas requiring different assessment and evaluation instruments. The first would establish student and faculty reactions to guest speakers focused on sustainability and the circular economy. The second would determine students' level of knowledge and sophistication about these content areas.

For the guest speakers, the initial data showed positive attitudes from students in response to the content and delivery. Knowledge was also shown to be effectively transferred. In addition, the speakers and the faculty reported highly positive results, though this was anecdotal. This led to the incorporation of guest-speaker evaluation into the two new courses. To provide more weight and relevance to the tool as it was integrated into the learning management system Canvas, a free-response item was added to request a summary of the content covered. Otherwise, the instrument matched the level-1-type survey items of the K4LM.

The instrument had a digital consent and disclosure waiver prior to the questions provided below. To determine the degree of reaction, a five-point Likert scale, using the standard range from 'Strongly Agree' to 'Strongly Disagree', was used. After the K4LM-Level-one-type questions, the open-ended content question was given (See Instrument 1).

Instrument 1.- Speaker Feedback Instrument

1. Engagement: The presenter did a good job of generating participant interaction.
2. Application: I will be able to apply what I learned during this session in other parts of the program.
3. Relevance: The material was relevant to my needs regarding SEEG and the circular economy in Chemical Sciences.
4. Learning: My learning was enhanced by the knowledge of the presenter.
5. Delivery: The presenter effectively delivered the program material.
6. Free Response: In your own words, please provide a brief summary (3-4 sentences) of the main points of the session. Then explain what the most important social, economic, environmental, and governance (SEEG)-related things tied to Chemistry and the circular economy you learned by attending the presentation.

As seen from the above items, the questions start with the reaction to the experience being examined. These quantitative items provide a sense of the overall value of the experience. In addition, they allow the separation of issues related to the presenter from other aspects tied to content or context. The free-response item was used to verify the speaker's presence and provide insight to the researchers about the content retained and valued by the participants.

For level 2 learning, the researchers started by taking a general baseline. The tool was distributed widely to gather a broad understanding of what type of content was understood and where gaps remained. Over time, the tools were converted into a pre-

posttest of content knowledge and incorporated into classes. Building on the work in year one, the researchers integrated the baseline evaluation into the two courses that were developed and began to gather data over time, both within a single semester and between semesters for students who completed both courses. The instrument designed for the Pre-Post evaluation of knowledge included the following open-ended questions (See Instrument 2):

Instrument 2.- Open-Ended Questions Instrument

1. What is Sustainability?
2. What is Green Chemistry?
3. What is the difference between Environmental Chemistry and Green Chemistry?

These content-based questions most closely matched typical assessments within the classroom. However, since they were ungraded and left open, several students made multiple attempts at the pretest as their knowledge grew. The data provided by these types of evaluations allow for course content and additional experiences to be planned and provided to attain the project's goals fully. Once the researchers were confident in the content and context taught to the students, the focus shifted to more Level-3-type concerns. Critical thinking skills, an open mindset, and asking questions could not be assessed with simple tests or evaluated easily with level one or two types of questions. As a result, the project contracted an expert consultant to host a series of workshops concluding in a symposium where examples and best practices would be shared.

4.3 LEVELS 3 AND 4

These sessions provided an opportunity to gather more data about level-one reactions from faculty, staff, and students, along with some level-two data on learning. It also allowed the author to observe the participants in his role as evaluator. These data reinforce and enrich the K4LM data gathered in the isolated context. These experiences also provide longitudinal views of the participants' work, allowing for higher-level Kirkpatrick analysis. Examples of the questions for level three and four instruments are supplied with some sample responses. The Right Question Institute ran a series of workshops and a symposium at NJIT on campus. From February to May 8, 2024, the Right Question Institute (RQI) provided training through a series of workshops for doctoral students and faculty:

- Session #1: All-February 2, 2024.
- Session #2: Faculty and students separated.

- Session #3: Faculty and students separated.
- Symposium: Full group of faculty, staff, administrators, and graduate students.

Ph.D. students were introduced to the Question Formulation Technique as a thinking process (Session 1). In Sessions 2 and 3, they gained expertise in using the Question Formulation Technique for Research (QFT-R) for their research work. Faculty were also introduced to the QFT in Session 1. Then, they had additional training in Sessions 2 and 3 on how to design their use of the Question Formulation Technique (QFT) and facilitate its use in their courses to enhance students' learning. The symposium allowed faculty and students to share their work and concepts related to the training. K4LM questions at level three are used to prompt presentations and gather feedback and input:

- How have you adopted the QFT and QFT for Research?
- How have faculty and students been using the QFT and QFT for Research?
- What value do students and faculty recognize both in learning/teaching and research?
- What changes have you seen in attitude/behavior?
- How has the quality of questions improved?

These series of workshops with a culminating experience allow for the higher level of Kirkpatrick's model to be utilized to produce both quantitative and qualitative data. Beyond that, the data and experiences provide a pathway for the level four reflection and evaluation of attainment by the researchers and program providers.

As an extension of the study evaluation, the team from RQI was surveyed to gather information about their experience with the NJIT team and the workshop participants. Excerpts from their response are provided below:

"The overall purpose of the symposium is to learn from practice and discuss the next steps to encourage further adoption at the institutional level and promote continuous use of QFT/QFT for Research. We have not had this culminating event from a partnership at the university level. We have convened separate RQI events, bringing together people to share what they have done and learned. But, there have not been any evaluations of those sessions."

This feedback from the outside consultant is used to provide evidence to NSF, the sponsor, to support the ongoing work. As the project moves towards a more stable and regular set of experiences, the leads and institutional leaders also need to be investigated. These are essential to completing the picture sought by K4LM.

The final set of prompts was an instrument for the research team. These were meant to gather level-four-type data to determine the degree of attainment. The following items are intended to stimulate reflection and help provide some feedback on the project to help make improvements and adjustments. These questions are meant to be high-level reviews of the work to help generate best practices and program improvement. A more open-ended style was chosen as a formative tool during the project. If the project was at completion or this was meant as a reflective activity, the more quantitative-style questions would be used with the associated Likert scale similar to that used for the speaker's reaction prompts.

1. Generally, what are your impressions of the progress of the project? What milestones or accomplishments are you most satisfied with, and which do you feel still need the most work?
2. Please consider the 'Course Development' aspect of the project. What is your opinion of the courses? What aspects do you feel are highlights to be showcased and/or shared? What aspects of the courses still need work to develop more fully and improve in order to meet the goals of the project?
3. Please consider the 'Expert Guest Speaker' aspect of the project. What is your opinion of the use and integration of expert speakers? What aspects do you feel are highlights to be showcased and/or shared? What aspects of the courses still need work to develop more fully and improve in order to meet the goals of the project?
4. Please consider the 'Question Formulation Technique (QFT) Workshop Series' hosted by the paid consultant. What is your overall opinion of the value provided by this series? Do you feel the project received a good return on its investment? Which aspects, if any, do you feel should be integrated into the program? Are there specific items you feel were of value or that were not worth the investment of time and resources?
5. Overall, what would you like to share about the project so far? What new actions are needed to move the work forward? What resources or additions would improve the outcomes and impacts of the project?

5 RESULTS

The most significant result is that the Kirkpatrick 4-Level Model (K4LM) served as a strong foundation for gathering quantitative data. When paired with qualitative data and direct observation, this data provides a compelling picture of the program for

evaluation reporting to the NSF and formative data for iterative curriculum design. The application of the Kirkpatrick Model in higher education has evolved to assess various levels of educational programs' effectiveness, including reaction, learning, behavior, and impact (Alsalamah & Callinan, 2021; Liang, 2023; Matondang et al., 2023). This model has demonstrated strengths in providing a comprehensive evaluation framework that considers immediate reactions and long-term impacts on students' skills and competencies. However, limitations exist, such as challenges in measuring specific intangible outcomes and needing more rigorous research designs and standardized evaluation tools (Smidt et al., 2009).

Compared to other evaluation frameworks, the Kirkpatrick Model stands out for its ability to assess the holistic impact of educational programs on learners, making it a valuable tool for evaluating training initiatives in higher education (Alhassan, 2022). K4LM has become a widely utilized framework for training evaluation across various fields. Numerous evaluators have applied this model to assess the effectiveness of training programs. Aljawharah Alsalamah (2022) presents a comprehensive bibliometric analysis of the Kirkpatrick Model, examining its utility and effectiveness in meeting the needs of training evaluations. This analysis highlights the model's significance by tracking the growth of studies focused on its applications across different contexts and settings. However, a limited number of studies use the K4LM in curriculum redesign projects. An apparent strength is the quantitative data it provides about the learning experience and associated assessments of participant learning related to that experience. This allows for iterative optimization and feedback on the value and effectiveness of the experience. Ludmila Pavlova (2010) provides an in-depth examination of the Kirkpatrick Model's application in assessing educational effectiveness, offering valuable feedback to higher education institutions. Praslova's work underscores the importance of using the model to enhance academic outcomes and institutional practices.

When focusing on the instructional design aspects of teaching within classroom settings, K4LM works well for the reaction to the experience and assessment of learning evaluation tasks. Mosquera et al. (2023) conducted a study involving an experimental group (EG) and a control group (CG) of students, each completing 15 programming tasks to measure the level of knowledge acquired. The evaluation utilized pretest knowledge, grade assessments, activity time, and post-test learning outcomes to determine the effectiveness of the interventions. In this case, K4LM lower-level evaluations were useful.

In the above case study, technology facilitated the use of K4LM tools. Students are accustomed to completing Google surveys, answering tests on computer-mediated platforms, and providing a wealth of personal data and opinions in online forums.

However, this familiarity does not mean the survey instruments or overall evaluation plan are effective. The authors found the resources of the website 'Ed Tech' provided helpful information about utilizing educational technology as part of the Kirkpatrick level one and two type evaluations within classes. It emphasized that learning management systems used to deliver class materials and conduct courses in online formats can facilitate the surveying and assessments designed with K4LM. However, since this model is more rigid and meant for training, just drawing the parallel between training in business and education in higher educational classrooms is not clear and direct. Each use case must be judged, and the desired outcomes and metrics must be matched with the experiences and activities being designed to determine how well Kirkpatrick-style levels of evaluation match the CID of the intended program.

Michael Cahapay (2021, p. 140) notes that the Kirkpatrick Model has limitations in higher education. He points out that the model "presents a propensity towards the use of the lower levels only, rigidity which leaves out other essential aspects of the evaluand, and paucity of evidence on the causal chains among the levels." If the quantitative focus of Kirkpatrick is not paired with other mixed-method approaches, its usefulness may be limited to reporting or benchmarking practices at the higher levels.

Overall, the K4LM was adequate for the redesign project because it was seen as a starting point from which a practical CID project could be evaluated. It was meant as summative measures of attainment nor proof of work being accomplished as part of compliance. As with any evaluation system, the K4LM is an evolving tool to provide data on progress toward desired outcomes and measure participant satisfaction and value. In that role, this model was advantageous to the project and the author as its evaluator.

6 CONCLUSION

Educational evaluation is essential for effective teaching and learning. With the advent of technology, data gathering and analysis have become more accessible, but this can also lead to difficulties in understanding the implications of what the data shows. The Kirkpatrick model provides a planning and implementation structure that can leverage technology to ensure an ethical and inclusive picture of the curriculum and instructional design being examined.

This article provides a foundation for evaluators and researchers working to understand and improve CID in higher education. The Model K4LM offers a structure for planning and implementing evaluation but has limitations. K4LM helps create an evaluation plan and ensures clear quantitative measures tied to overall aims and success metrics. However, from literature research and practical experience, a mixed-method approach

applied across in-class and out-of-class experiences provides a more accurate and richer perspective. The closer to the interface between learner, educator, and content, the more easily K4LM can be applied, especially at levels one and two. The more broadly education is examined, such as at the curriculum or program level, the more evaluations must be multifaceted and gather data through a range of means to tell a clearer story of what is happening over time as the participants move through a series of experiences aimed at the overarching goal of that program.

SOURCES

Akbar, S., Darungan, T. S., & Rahma, D. S. (2024). Instrument Development for Study Programs Excellence Evaluation Based on Kirkpatrick Evaluation Model. *Jurnal Pendidikan Kedokteran Indonesia: The Indonesian Journal of Medical Education*, *13*(2), Article 2. https://doi.org/10.22146/jpki.91960

Alhassan, A. I. (2022). Implementing Faculty Development Programs in Medical Education Utilizing Kirkpatrick's Model. *Advances in Medical Education and Practice*, *13*, 945–954. https://doi.org/10.2147/AMEP.S372652

Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's Four-Level Model of Training Criteria to Evaluate Training Programmes for Head Teachers. *Education Sciences*, *11*(3), 116. https://doi.org/10.3390/educsci11030116

Babbie, E. R. (2016). *The Practice of Social Research, 14th Edition* (14th ed.). Cengage Learning.

Cahapay, M. (2021). Kirkpatrick Model: Its Limitations as Used in Higher Education Evaluation. *International Journal of Assessment Tools in Education*, *8*(1), Article 1.

Chang, N., & Chen, L. (2014). Evaluating the Learning Effectiveness of an Online Information Literacy Class Based on the Kirkpatrick Framework. *Libri*, *64*(3), 211–223. https://doi.org/10.1515/libri-2014-0016

Clift, R., Martin, G., & Mair, S. (2022). Chapter 3 – Sustainability and the circular economy. In C. Teodosiu, S. Fiore, & A. Hospido (Eds.), *Assessing Progress Towards Sustainability* (pp. 35–56). Elsevier. https://doi.org/10.1016/B978-0-323-85851-9.00001-8

Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, Fifth Edition (5th ed.). SAGE Publications, Inc.

Frey, B. B. (2018). *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation, First Edition* (1ra Edición). SAGE Publications, Inc. https://doi.org/10.4135/9781506326139

Giarratano, J. (2022, July 6). Interdisciplinary Advancement: New Center for Urban Transformations to Bring Innovations in Research and Implementation. *Next Generation Program*. https://nextgen.gsu.edu/2022/07/06/center-for-urban-transformations/

Langer, A., & Ghaye, T. (2012). *Reflective Practice*. *13*. https://www.tandfonline.com/journals/crep20

Liang, H. (2023). Using Kirkpatrick Four Level Evaluation Model to Assess English for Medical Purposes Courses – A Case Study at a Medical School. *The Educational Review, USA*, *7*(6), 756–760. https://doi.org/10.26855/er.2023.06.021

Matondang, Z., Sitompul, H., & Wijaya, K. (2023). Implementation of Evaluation of Kirkpatrick Model in Statistics Course Based on Case Method at the Department of Building Engineering. *Jurnal Penelitian Pendidikan IPA*, *9*(5), 2677–2685. https://doi.org/10.29303/jppipa.v9i5.3371

Mosquera, J. M. L., Suarez, C. G. H., & Guerrero, V. A. B. (2023). Effect of Flipped Classroom and Automatic Source Code Evaluation in a CS1 Programming Course According to the Kirkpatrick Evaluation Model. *Education and Information Technologies*, *28*(10), 13235–13252. https://doi.org/10.1007/s10639-023-11678-9

Paull, M., Whitsed, C., & Girardi, A. (2020). Applying the Kirkpatrick model: Evaluating an "interaction for learning framework" curriculum intervention. *Issues in Educational Research*, *26*(3), 490–507. https://doi.org/10.3316/ielapa.509274814654719

Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. *Educational Assessment, Evaluation and Accountability*, *22*(3), 215–225. https://doi.org/10.1007/s11092-010-9098-7

Reio, T. G., Rocco, T. S., Smith, D. H., & Chang, E. (2017). A Critique of Kirkpatrick's Evaluation Model. *New Horizons in Adult Education and Human Resource Development*, *29*(2), 35–53. https://doi.org/10.1002/nha3.20178

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A Systematic Approach, 7th Edition*.

Smidt, A., Balandin, S., Sigafoos, J., & Reed, V. A. (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

SOBRE O ORGANIZADOR

Dr. José Luis Escamilla Reyes- Profesor del Tecnológico de Monterrey, Campus Ciudad de México desde 1998. Doctor en Física por la Universidad Autónoma Metropolitana, Unidad Iztapalapa. Cuenta con una experiencia docente de 32 años. Es coautor de Manuales de Física II y Física III, así como de dos ebooks, uno sobre Física General y otro sobre Óptica y Física Moderna. Está certificado en el Programa de Desarrollo de Habilidades Docentes del Tecnológico de Monterrey. Ha participado con varios trabajos en Congresos Nacionales e Internacionales relacionados con la Física de Semiconductores de los grupos IV y III-V. Sus áreas de interés son: fuentes alternativas de energía, Física del Estado Sólido, diseño y aplicaciones de los MEMS y modelación matemática de Sistemas Complejos. Ha publicado más de 15 trabajos arbitrados y memorias en congresos. Colaboró en el diseño y construcción de láseres pulsados de N2 en el Laboratorio de Óptica Cuántica de la Universidad Autónoma Metropolitana Iztapalapa (UAMI). En el Tecnológico de Monterrey Campus Ciudad de México, participó en el desarrollo de un prototipo de Celda de Combustible con membrana de intercambio protónico (PEMFC) de alta eficiencia. Obtuvo la Medalla al Mérito Académico por el mejor promedio de Maestría otorgada por la UAMI. Fue líder de la Cátedra de Investigación "Micro Sistemas Electromecánicos: Diseño y aplicaciones" del Tecnológico de Monterrey, Campus Ciudad de México y miembro del SNI.

ÍNDICE REMISSIVO