

Chai Chats: An Online Teacher-Training Program of Observation and Social Connectedness Evaluated via Contribution Analysis

Toni McLaughlan

Lancaster University, United Kingdom // t.mclaughlan@lancaster.ac.uk

ABSTRACT: This paper evaluates an international, online, content skills-based teacher education program sponsored by the U.S. State Department. The evaluation was designed using a RUFDATA framework (Saunders, 2000) to facilitate a complete, reflective assessment of the target program. Establishing causes-and-effects of the program's performance and data analysis involved adoption of a Contribution Analysis (Mayne, 2001; Mayne, 2008). Utilizing the six steps detailed in Mayne (2012), a credible contribution story emerged, highlighting strengths and weaknesses of transitioning teacher-training programs to virtual platforms. This evaluation has implications for teachers, teacher trainers, professionals planning similar programs particularly in developing regions, and individuals interested in how theory can be applied practically to impact continued teacher education processes. This paper contributes to knowledge as there are few formal evaluations of online international teacher education programs that facilitate observation of all aspects of a virtual course over an extended period of time and provide small-group engagement with course creators, especially with populations straddling the digital divide. It is also the first to conduct a theory-based evaluation of a U.S. English Language Specialist project despite the program's 1991 inception and current running rate of 150-200 projects annually worldwide (U.S. Department of State, 2021).

Keywords: Teacher training, Educational technology, Online, Developing regions, Contribution Analysis

1. Introduction

The integration of educational technology, especially resulting from Covid-19, has been growing for years. It is essential that educators are trained in effective teaching strategies involving technology and delivery of online instruction. One may question, then, whether using observation and practical applications of such tools may be an effective approach to teacher-training in these areas. This paper aims to evaluate an international, online, content skills-based teacher training program sponsored by India's Regional English Language Offices (RELO), part of the United States Bureau of Educational and Cultural Affairs (ECA), a division of the United States' Department of State.

This evaluation was first designed using the RUFDATA framework (Saunders, 2000) in order to facilitate the development of a complete and reflective assessment of the target program. To address remaining questions in analyzing the evidence gathered and establishing causes-and-effects of the program's performance, the evaluation adheres to a Contribution Analysis (Mayne, 2001; Mayne, 2008) approach and its six steps as detailed in Mayne (2012) in an effort to establish a strong, credible contribution story.

The paper begins with a situational analysis of the object of the evaluation, followed by a brief background of both RUFDATA and Contribution Analysis (CA). Subsequently, the methodology behind data collection in supporting the CA is outlined. The data analysis section then details each of CA's six steps, culminating in a revised contribution story. The paper concludes with discussion of the evaluation process, its limitations, and implications for future research, educators, and program developers.

1.1. Situational analysis of the object of the evaluation

1.1.1. Object of evaluation

The object of this evaluation is a professional development (PD) program for higher education (HE) instructors geared toward improving capabilities in the online teaching of academic English skills using educational technology throughout India. Though the program as a whole has multiple components, only the "Chai Chat" piece will be evaluated here. Participants observed a full 10-week virtual course that demonstrated the instruction of academic English reading or writing skills; each skill was taught for one hour weekly. This observation was done asynchronously; each lesson was taught and recorded through Zoom and shared with participants, who, while observing, completed weekly observation task forms (Wajnryb, 1992) that guided participants to focus on

several highlighted aspects of a given lesson. These task forms were then submitted and reviewed by the RELO-selected English Language Specialist for each Chai Chat group which consisted of roughly 15-30 participants. Chai Chat groups met synchronously on Zoom with their Specialist for an hour weekly to discuss content, educational technology, and overall teaching methodology of the week's online lesson. During these meetings, instructors were encouraged to share ideas on how to best adapt lessons to their own contexts, such as larger/smaller groups, disadvantaged populations, academically at-risk students, asynchronous online teaching, or in-person instruction. They were encouraged to share instances of their own adoption of lessons' approaches, ask questions regarding incorporation of new technologies, share resources, and suggest preferred platforms for doing so. By the end of the program, instructors engaged in observing all elements of a full course: lessons, homework, rubrics, grading, online gradebooks, LMSs, and feedback to students.

1.1.2. Situational analysis

Institutions of higher education (HEIs) throughout India were suddenly faced with challenges of transitioning to distance learning with the emergence of Covid-19 (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2020). Closures of student and faculty residences often accompanied closures of campuses (Careers360, 2020), posing challenges of accessibility to internet connectivity and technological devices for both students and faculty. India's Ministry of Rural Development (2018) reports that only 47% of Indian households can reliably access electricity for 12+ hours a day. Furthermore, a mere 23.8% of Indian homes have internet access; 10.7% have a computer at home (Ministry of Statistics & Programme Implementation, 2018). Therefore, thriving in distance education is unlikely for the majority of tertiary students and faculty, and a challenge even for those having *some* accessibility capabilities. Additional studies of the region highlight a need for instructor training in online educational tools and strategies in order for web-based instruction to be effectual (Lakshmi & Agarwal, 2017); similarly, students themselves may lack digital literacy skills for educational purposes (Naresh et al., 2016).

2. Informing frameworks of the evaluation

2.1. RUFDATA

In designing a solid foundation for program evaluation, the meta-evaluative RUFDATA tool (Saunders, 2000) was adopted in directing the various reflexive questioning processes. The framework requires its users in initial planning stages to define key concepts comprising the RUFDATA acronym as they relate to the evaluation. The RUFDATA framework (Saunders, 2000) was applied as follows:

- *Reasons and purposes:* To provide value on the effectiveness of an intervening program developed to fulfill an unexpected yet substantial need in online instructional strategies in the Indian HE system
- *Uses:* By U.S.- and internationally-based employees of the ECA section of the U.S Public Affairs and Public Diplomacy division
- *Foci:* The "Chai Chai" component of a larger distance learning initiative
- *Data and evidence:* From individuals working with key stakeholders, program participants, their students, and program students in the form of surveys, program evaluations, and interviews
- *Audience:* A wide range of professionals in education, curriculum design, and training; the project's stakeholders
- *Timing:* Several weeks after program completion
- *Agency:* RELO and its personnel under guidance of the evaluator, who also serves as the program's primary instructor and one of the five Specialists

2.2. Contribution analysis

Contribution Analysis (Mayne, 2001) is an evaluative approach that assists evaluators in addressing issues of cause-and-effect in assessing an intervening program's impacts. It is theory-based in that it involves the development of a *theory of change* (Weiss, 1995; Weiss, 1997) intended to model the expected changes of the program under evaluation, identifying various causal mechanisms that influence results of the program throughout (Patton, 2008). When an experimental evaluative design for inferring causality is not plausible, as is the case here, CA can be utilized in measuring a program's outcomes and impacts at various stages while providing explanations for why results do or do not occur.

3. Materials and methods

This program makes use of a mixed-methods design. Quantitative calculations present only descriptive data; no experimental element could be practically employed, and causation is instead addressed qualitatively through CA.

Throughout the program, participants were asked to complete observation task forms, adapted from Wajnryb (1992), who states that observation serves as a teacher training tool as educators gain skills of analyzing and interpreting while viewing others' techniques; this is then used for self-reflection. Observers completed the forms weekly, requiring them to record information from the lesson with a dual focus on a pedagogical element (i.e., classroom management, instructional language) and the content of the lesson's objective. This data informed Specialists on how to modify their Chai Chat meetings to be optimally relevant and guided the academic skills instructor in adapting the program's remaining lessons. For evaluation purposes, task form submissions were used primarily in reviewing attrition rates.

Three weeks following the program's end and into a new academic year, program evaluations were sent to participants via email. The program evaluations were developed and collected through Google Forms and included questions adapted from the Student Evaluation Quality Questionnaire (Marsh, 1982). The SEEQ is among the most utilized measures of course evaluation worldwide and is recognized for its consistent international reliability and validity, both attested over time (Coffey & Gibbs, 2001). Two weeks following review of program evaluation data, students of the academic skills courses were sent post-surveys inquiring about changes to their instructors' teaching habits using SurveyMonkey. This was done to corroborate participants' self-assessments on program evaluations. Similarly, Specialists received questionnaires via Google Forms regarding their participants' integration of program materials. Assumptions were further supported with data from student course evaluations. These evaluations were not initially intended for Chai Chat evaluation and were fully conducted under the direction of RELO, though review of its content reflects similar adaptation of the SEEQ (Marsh, 1982).

In addressing RELO stakeholders' program expectations, relevant data from Specialist questionnaires were reviewed, and semi-structured interviews were conducted with on-site RELO staff in India. Though representatives of English Language Programs with the U.S. Department of State are aware of and support the development of this evaluation and its possible publication, U.S.-based Embassy employees declined to provide information regarding in-house evaluation processes of individual Specialist projects. RELO stakeholders' goals mirror those listed on the ECA's website (Bureau of Educational and Cultural Affairs, n.d.).

4. Results

Using CA outlined in Mayne (2012), all six steps of the iterative analytical approach are detailed below. This includes a comprehensive Theory of Change (ToC), adapted from Mayne (2015).

4.1. Step 1: Set out the cause-effect issue to be addressed

In order to identify cause-effect issues related to the intervention, it is necessary to first clearly state the intended effects of the program. The primary goals of the intervention being evaluated are to:

- enhance instructors' effectiveness in teaching online and use of educational technologies,
- improve instructors' teaching efficacy by modeling a robust set of educational strategies in academic reading and writing in English,
- build public diplomacy through the promotion of English language programming, and
- foster mutual understanding between the U.S. and other nations through cultural exchange.

The attribution question here involves assessing whether impacts have been made in these areas, addressing whether any such developments are indeed the result of the intervening program, and, if they are, to what extent. The level of proof necessary is one of importance, as transitions to distance learning are imminent for most educators and likely to remain a reality to some extent for quite some time.

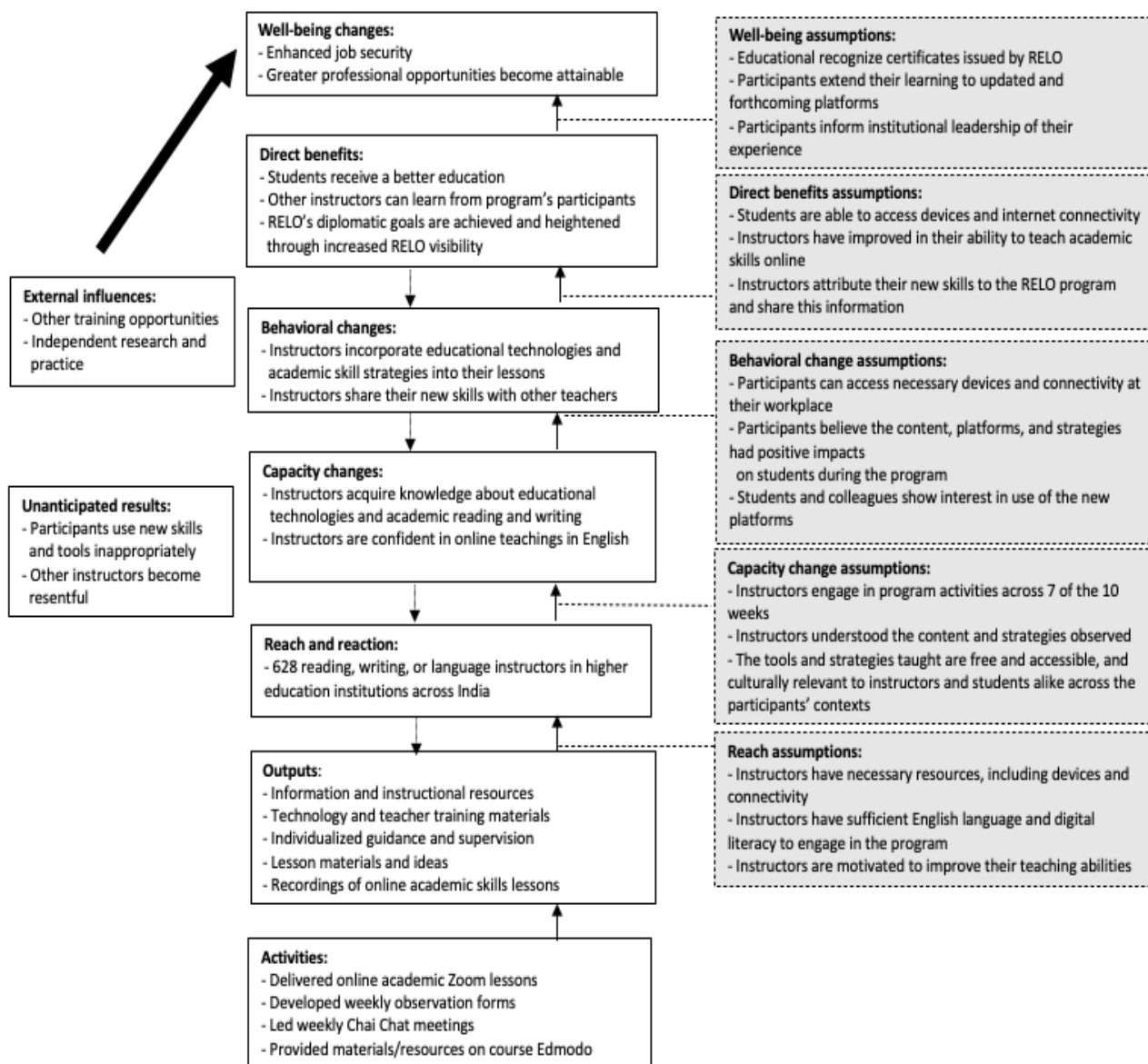
The intended contribution of the intervening program appears plausible. The problem is well-understood, though relevant baseline data exists only in the form of a pre-survey done by key stakeholders in assessing the direction

that the intervening program should take, showing a need in HE PD regarding teaching with technology, along with opportunities to improve English skills and non-technology related instructional methods.

4.2. Step 2: Develop the postulated theory of change and risks to it

The ToC reflects all four anticipated outcomes throughout the results chain, alongside the assumptions underlying each link (Figure 1). Key stakeholders were involved in discussing these assumptions, signifying shared understandings and program goals among key players and the evaluator.

Figure 1. The Theory of Change



This ToC takes into account an intermediate level of detail while clearly demonstrating the expected contribution of the program. Strengths and weaknesses of the chain and its corresponding assumptions are discussed throughout the paper.

4.3. Step 3: Gather the existing evidence on the theory of change

4.3.1. Evidence on outputs and activities

The outputs of the intervening program have tangible evidence in numerous formats. Participants gained access to academic skills content-related material in the forms of activities on several free, web-based platforms—all copiable, editable, and shareable—with in-class and web-based tutorials to guide their use. They received links to other open-education resources (OERs), lesson plans, and exercises for each topic discussed. All videos remain accessible and shareable on RELO India’s public Facebook website.

4.3.2. Evidence on assumptions

Existing evidence on the program’s observed results, as well as ToC assumptions and the potential for external factors that may influence outcomes, are gathered and presented.

4.3.2.1. Reach assumptions

The program was expected to reach approximately 900 instructors across nearly all states of India. ECA affiliate offices advertised the program and recruited from HEIs within their respective regions. Instructors were informed that technological devices, stable internet connectivity, and bandwidth or data capabilities that support hours-long videos over ten weeks were required. This, unfortunately, likely excluded potential participants, particularly those living in underserved areas. 628 registrants were instructors in HE; as this paper aims to evaluate the program’s value as a PD opportunity for tertiary instructors, only this data is included. Based on already-existing evidence gathered from the RELO office (Table 1), 128 of the original 628 (20%) registrants successfully completed the program, highlighting weakness in the *reach and reaction assumptions* link.

Table 1. Breakdown of completion status, registration, and survey respondents

Certificate of completion	Program registrants	Program evaluation respondents
Earned	128	96
Not earned	500	52
Total	628	148

In gauging the normalcy of such attrition, one may revert to the literature on similar programs. This is a free, voluntary program that results in no qualification, transferable credits, etc. Participants are not required to enroll nor complete the program, have not invested in it financially, and are intrinsically motivated. Though a completion certificate is issued, it is uncertain what, if any, advantage it carries in India’s highly-competitive academic job market.

Massive Open Online Courses (MOOCs) seem similar in structure to this program: participation is typically voluntary; participant work is required, assessed, and receives feedback but courses are typically intended for skill development over certificates. Thus, literature on MOOCs may serve as a logical basis for comparison. Several hosts offer over 4,500 MOOCs courses (Bouzayane & Saad, 2017), whose high dropout rates are fairly well-documented. According to Li et al. (2016), Coursera, a highly popular MOOC platform, sees successful completion rates averaging 7% to 9%. Later studies reviewing other platforms found less drastic albeit high attrition rates for their MOOCs as well: Cobos et al. (2017) monitored attrition rates of two courses requiring a cumulative score of 60% marks to earn a certificate, and two others simply requiring completion of 50% of coursework. Average certificate issuance rates were 13.25% and 22.25%, respectively. Given that certificate requirements of the program evaluated here falls between these two in rigor—obliging 70% of attendance and task form submission that were reviewed and recorded but not scored—the expected completion rate would fall between them accordingly, as is the case with the 20.4% of HE registrants successfully completing the intervening program.

Also addressed in this link is the digital divide that undeniably exists within India. Additional statistics on Indian infrastructure per surveys from London-based company QS (PTI, 2020) show that only 15% of Indian households have broadband, 53% of whom report poor connectivity, with 11% reporting electricity issues. 40% of internet users depend on mobile hotspots, 96% of whom cited connectivity issues. Still, participants acknowledged accessibility requirements upon registration. Thus, reasons *why* this attrition occurred must be explored.

4.3.2.2. Capacity change assumptions

Participants' engagement throughout the program is already demonstrated through certificate issuance as outlined above. Statistics on certificate completion and task form submissions can be found in Table 1 and Figure 3, respectively. Access to all course resources remained free and accessible throughout the program and at present.

In assessing the likelihood that participants learned, understood, and adopt lesson materials demonstrated in the skills-based classes observed weekly, we first turn to the research. In their book on impacts of PD in education, Condon et al. (2016) describe how PD indeed benefits its participants in the long-term, which is shown to have positive trickle-down effects on student outcomes. However, when delving more deeply into the integration of educational technology, McDaniel and Kenny (2013) found that, though instructors often believe such tools support student learning, their instructional strategies do not as often reflect this—typically a result of their own lack of knowledge on its use. Therefore, this opportunity, with its guidance and demonstration of applied technology, is expected to be impactful.

4.3.2.3. Behavioral change assumptions

Participants' motivations seem intrinsic; their interest appears established, verifying this element of the link. Whereas participants' ability to utilize the educational technology in their current situation requires home accessibility to technology, their ability to extend use to traditional classroom settings requires institutional connectivity. Presently, data on accessibility specifically in Indian HEIs is not available.

The opportunity is positioned to be impactful based on its adherence to PD best practices. Research suggests that one-time PD events do not sufficiently provide participants with opportunities to effectively integrate technology into their own lessons (Gunter & Gunter, 2015; Lacey et al., 2014). Research from Gunter and Reeves (2017) found that PD opportunities that extend training across weeks and demonstrate authentic use of technology in teaching relevant content increase participant engagement; providing participants with opportunities to practice such strategies heightens their likelihood of adopting techniques. Such findings are corroborated by Darling-Hammond et al. (2017), who extensively reviewed 35 publications centered on positive links between PD opportunities in teaching and their effects on instructional practices and student outcomes. The authors suggest seven tips in best practices of PD programs: PD should be content-focused, deliver active learning opportunities, encourage collaboration, incorporate modeling of lessons, provide expert guidance, offer reflective feedback, and occur over an extended period of time. The intervening program involved all seven recommendations.

The attitudes of Indian educators toward educational web-based materials are important in gauging their likelihood of adopting them. In extensive efforts on behalf of UNESCO, Hodgkinson-Williams and Arinto (2017) review global perspectives—including Indian attitudes toward them—on OERs, which were heavily featured in the intervening program. Researchers found that, though Indian educators typically had little awareness of OERs, they were very receptive toward integrating them following engagement with them.

4.3.2.4. Direct benefits assumptions

Given adherence to best practices, it is likely that students will receive a better education, assuming they can access it. Though issues of Indian infrastructure have been addressed, a recent study from the University of Chicago has gathered accessibility data specifically relating to university students in India. Mukhopadhyay (2020) describes that 27% of university students from urban backgrounds have at-home internet access; this number drops as low as 2% for students from rural homes. This study suggests that 47% of university students have access to an appropriate device at home, though this number includes mobile devices—many of which are shared.

There are no publications regarding RELO's goals, nor of participants' experiences from previous Specialist programs. Yet, because of its importance to key stakeholders, it will be explored using new data.

4.3.2.5. Well-being change assumptions

Stated by Montague et al. (2002) as cited in Mayne (2008), links involving *indirect influence* are more difficult to validate, as is the case here. An exploration of the connection between PD, especially RELO certificates, and

elevated job opportunities or security in India has not yet been undertaken. Nevertheless, Hodgkinson-Williams and Arinto (2017) found that educators in India do perceive a sense of heightened reputation after adopting and sharing OERs in their institutions.

4.4. Step 4: Assemble and assess the contribution story, and challenges to it

The contribution story is concisely detailed in Figure 2, as adapted from Mayne (2012), below:

Figure 2. Contribution analysis in evaluating existing evidence on program and stakeholder goals

<p>Strengths of the assumptions chain:</p> <ul style="list-style-type: none">• Developed a theory of changes with highly plausible assumptions• Examined pre-survey data in informing intervention construction to confirm alignment of stakeholder goals and program delivery• Reviewed program’s activities and outputs in confirming the program’s reliable, intended implementation and adherence to best practices in PD programming• Established a research-grounded baseline for attrition rates of similar programs• Detailed issues of infrastructure underlying India’s digital divide within higher education• Identified likely motivations for program participation• Illustrated logical expectations for PD impacts in the Indian context• Outlined research-based expectations for Indian attitudes toward applying PD-demonstrated practices <p>Weaknesses in need of further investigation:</p> <ul style="list-style-type: none">• Explore explanations underlying high attrition rates• Identify possible external influences• Outline perceived relevance and benefits of the program• Research the extent to which participants learned, understood, utilized program outputs• Gather data on achievement of RELO goals
--

The ToC, though iteratively revised during program planning, is finalized as presented in step 2. The contribution story overall has both strengths and weaknesses, as reflected in Figure 2.

4.5. Step 5: Seek out additional evidence

This CA step centers around integrating evidence gathered in the weeks following program completion and aims to strengthen the ToC’s validity. To assess whether the program contributed to changes in participants’ teaching and achieving RELO’s goals, further evidence was sought in the form of:

- participant program evaluations,
- student course evaluations,
- student post-surveys,
- Chai Chat Specialist questionnaires,
- weekly observation task forms, and
- interviews with key stakeholders.

4.5.1. Reach and reaction

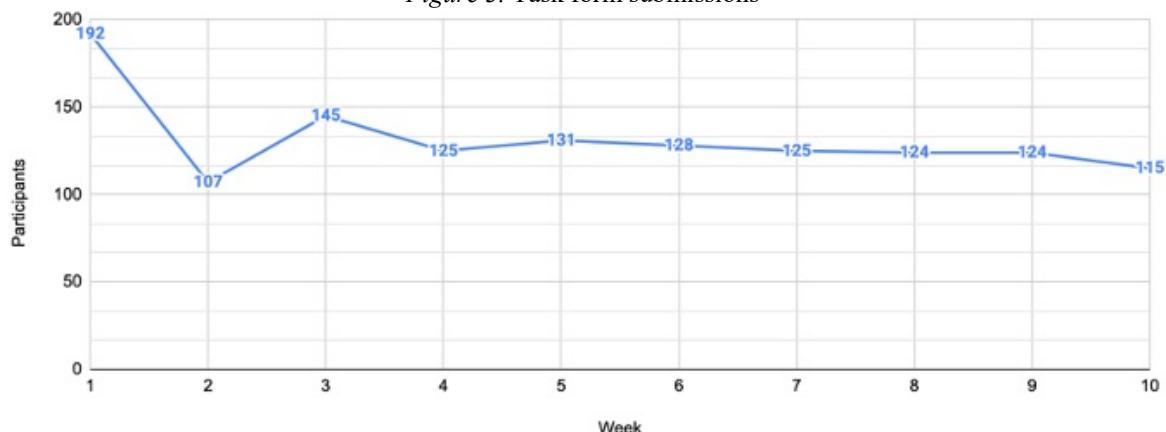
Evidence verifying this link involves deeper exploration of weekly observation task form submissions and program evaluation data.

4.5.1.1. Weekly observation task forms

As noted, the program faced high attrition rates. However, this is not uncommon for comparable initiatives, as outlined in step 3. Given these findings, the intervening program’s attrition rates are fairly expected. Crucially, investigation of weekly task form submissions suggests the majority of participants dropped out before program initiation, based on comparison of registrants (628) to submissions of week-one forms (192). Figure 3 shows the

number of observation task form submissions weekly, which average near the 128 issued certificates of completion ($\mu = 131.6$).

Figure 3. Task form submissions



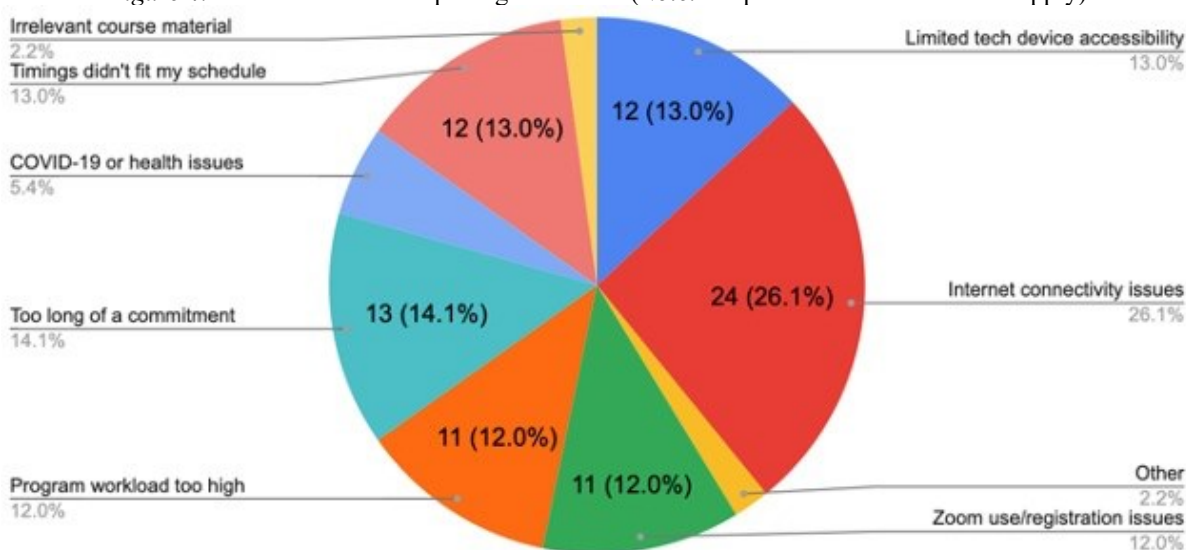
4.5.1.2. Program evaluation data of Non-CERT participants

Three weeks following completion of the Chai Chat program, participants were sent a program evaluation adapted from the SEEQ (Marsh, 1982). In total, 148 responses were received ($n = 148$) from educators in HE. Of these, 96 respondents had successfully completed the program (CERT group) and earned completion certificates. 52 respondents (Non-CERT group) did not receive a certificate (Table 1).

It is important to explore the reasons for attrition and to identify missteps as they correspond to assumptions in the results chain. Although the program evaluation was otherwise identical for all program registrants, the non-CERT group evaluation included an additional item inquiring about reasons for not completing the course.

Issues related to internet connectivity are cited (Figure 4) as the primary reason for attrition among the Non-CERT respondents ($n = 52$). Also commonly identified are issues of access to technological devices, and the digital literacy required to navigate registration to, access to, and materials within the course. Several responses address the timings of Chai Chat meetings (five per subject per week were offered) and the time commitment required of the program—valuable information for stakeholders in future planning initiatives. Only two individuals took issue with the content of the course itself.

Figure 4. Reasons for not completing the course (Note. Respondents selected all that apply)



Though this is strong indicative data, one cannot definitively state that the 52 Non-CERT respondents reflect the 500 registrants who did not earn a certificate nor the 436 who never began the program.

4.5.2. Capacity changes

The next link in the results chain addresses participants' growth in knowledge and ability, evaluated using program evaluation data of the CERT group. As non-CERT participants missed at least 40% of Chai Chat meetings, their data has been removed from calculations in the following sections; they likely gained insufficient exposure to course content.

4.5.2.1. Program evaluation data of CERT participants

The *capacity change* link appears to be well-verified by CERT participant data; continued participation may reflect their language and extended technology capabilities. According to responses on the program evaluation (Table 2), they have largely "learned and understood valuable... academic skills content, educational technology, and other materials" taught in the program. On a Likert-type scale of one (*Strongly Disagree*) to five (*Strongly Agree*), Table 2 shows mean scores for each area of learning and percentages of respondents who *Agree* or *Strongly Agree*. Average scores typically reach about 4.0 (Marsh, 1982); means presented here and throughout the paper are often above 4.5.

Table 2. Participants' program learning

Instructional area	Mean score	% of responses of <i>Agree</i> or <i>Strongly Agree</i>
Academic skills content	4.65	96%
Educational technology	4.75	95%
Other resources	4.63	95%

4.5.3. Behavioral changes

The next link in the ToC's results chain investigates participants' changes to instructional practices. The program's impact on participants' teaching is evidenced again from program evaluations, along with post-surveys from students, student course evaluations, and questionnaires from Specialists, who led the Chai Chat discussions.

4.5.3.1. Program evaluation data of CERT participants

Extracted from program evaluations, data from the 96 respondents of the CERT group show participants largely self-report having incorporated both academic skills content and educational technology tools demonstrated in the Chai Chat program. Table 3 shows percentages of respondents who claim to have added program teachings into their lessons, and the percentage of respondents who have otherwise adapted teaching methods as a result of the program.

Table 3. Changes made to teaching as a result of the intervening program

Area of change	% of affirmative respondents
Added new academic skills content	92%
Added new activities with educational technology	93%
Adapted other elements of previous teaching strategies	95%

4.5.3.2. Student post-surveys

Students in the Indian HE system are often taught by the same professors throughout their education. All students participating in the skills-based courses were HE students whose lessons suddenly transitioned online at the previous academic year's end; they would therefore be capable of noting changes to instructors' online teaching and were familiar with content exemplified in the skills-based courses, uniquely allowing them to identify these methods and changes to their instructors' strategies.

Given that Chai Chat participants were invited from the same universities as students participating in the academic skills courses, these students were expected to have both instructors who did and who did not participate in the PD program and may then be able to compare approaches employed between them. However, of the 62 students who completed the survey, 50% were not sure whether their professors had participated or not.

Of the 31 remaining, 13 students (21%) representing six different institutions could definitively state having instructors who participated in the intervening program while 18 students (29%) reported that none of their professors were involved.

Of the 13 students who had participating instructors, nine (69%) stated having noticed differences to instructors' methods while the remaining four students (31%) reported no changes in instruction. Based on open-ended survey data, the changes seem to be in both content and educational technology; several students also commented that participating instructors led more interesting, motivating lessons and that the online environments they created included more interaction and engagement.

4.5.3.3. Student course evaluations

81 of the 160 HE students registering in the synchronous academic skills classes responded to program evaluations. Students' opinions of the course and its technology are pertinent in assessing whether the Chai Chat program provided relevant content for its participating instructors.

Table 4 shows, on a Likert-type scale of 1 (*Very Poor*) to 5 (*Very Good*), students' mean scores in the areas of content and technology and the percentage of respondents who highly assessed the value of the course. It also shows students' perceived abilities in academic reading and writing in English before and after the course and perceived self-improvement.

Table 4. Student perceptions of abilities and self-improvement

Program area	Mean score	% of <i>Good</i> or <i>Very Good</i> responses	
Course Content	4.59	93.7%	
Use of Technology	4.7	92.3%	
Perceived ability before course	Perceived ability after course	Mean self-reported improvement	% of students reporting improvement
2.76	4.12	1.36	90.1%

4.5.3.4. Specialists' questionnaires

The program involved five Specialists, selected by the RELO office in and approved by the ECA, who led "Chai Chat" discussions with 15 to 30 individuals for one hour per week per academic skill.

Following the Chai Chat program's conclusion, Specialists were sent a brief questionnaire regarding their estimations of participants' integration of program information. Four of five Specialists responded ($n = 4$; 80%). All responding Specialists stated that participants had discussed using both the educational technology tools and the content presented in the skills courses. Specialists' approximations of how many participants incorporated program methods are detailed in Table 5.

Table 5. Specialists' estimations of participants who discussed program content, tools use

Specialist estimation, based on weekly Chai Chat discussions	Specialist 1	Specialist 2	Specialist 3	Specialist 4
	80%	60%	75%	70%
Average estimate of participants utilizing program material	71.25%			

4.5.4. Direct benefits

Here, it is proposed that students will receive a better education resulting from instructors' participation in the intervening program. Though difficult to verify, starting points for strengthening this link may include instructors' improvement in teaching academic skills, online, in English. A direct benefit for RELO and US government stakeholders involves the promotion of RELO programming, assuming participants attribute improvements to the program itself and are able and willing to share their new skills with colleagues.

Evidence that the program has positively impacted students' education has been gathered from CERT participants' program evaluations; to support goals of the RELO program, relevant data again from CERT

participants' program evaluations are reviewed along with Specialists' questionnaire responses and data from interviews with two RELO representatives.

4.5.4.1. Program evaluation data of CERT participants

The assumption of improved education is rooted in the notion that instructors are providing better lessons. Participants were asked to self-assess their confidence in the areas of teaching academic skills content, online, and in English, both before and after the program. The majority of participants perceive improvement in all three areas, with details of mean perceptions on a 5-point Likert-type scale and percentages of respondents perceiving improvement in Table 6.

Table 6. Perceived improvement in teaching abilities resulting from intervening program

Instructional skill area	Mean confidence before the program	Mean confidence after the program	Mean increase	% of respondents perceiving improvement
Teaching academic skills	3.47	4.58	1.11	75%
Teaching online	3.35	4.57	1.22	75%
Teaching in English	4.12	4.60	0.48	57%

In addressing whether participants are able and willing to share information regarding RELO programs and opportunities, participants would have to consider the program valuable. When asked about their attitudes toward the PD opportunity overall, the mean rating for the program was 4.81, with 97% of respondents stating it was either *Good* or *Very Good*.

4.5.4.2. Specialist questionnaires

As the goals of the RELO office are most familiar to and best understood by those in close proximity to stakeholders, evidence of the program's value was gathered from Specialists and two RELO representatives involved in the project.

Whereas the four responding Specialists believe the program successfully advanced public diplomacy through promotion of the English language, only three believe the goals of mutual understanding through cultural exchange were achieved; the fourth reported being "unsure."

All four Specialists believe the program successfully advanced public diplomacy. Reasons listed for this include participants' praise of RELO programs, resources, and opportunities during Chai Chat meetings; gratefulness for opportunities to practice English; shared best practices in English instruction; and observations of a master English teacher.

Specialists are, however, divided on whether the program successfully fostered mutual understanding between the U.S. and India through cultural exchange. Those responding affirmatively cite discussion of shared experiences and challenges among teachers in both cultures, such as connectivity issues in rural areas and sharing one device for multiple residents within a U.S. household, to the surprise of the Indian instructors. Similarly, good teaching strategies, including those intended to overcome such challenges, are shared between these two groups and appear to be cross-cultural. One of the Specialists pointed out that the mutual understanding seemed to be occurring as a consequence of the program, as intended, but that it often seemed to emerge as an effect of "hits and misses." For example, many instructors of students in rural areas noted that they at times have to deliver content through televised or radio programs, in which the lead teacher had no experience or training; adapting lessons to fit India's most developing and underserved regions was left largely up to discussion during the Chai Chat meetings between the Indian participants themselves, though future programming could be modified in light of this.

One Specialist reported uncertainty on the program's success in this area, stating that students seemed to be responding more to cultural differences than instructors, and that several Chai Chat instructors "seemed completely lost," though they enjoyed the differences in approaches when grasped.

4.5.4.3. Interviews with RELO representatives

Two representatives of an India-based RELO-affiliated office, with over ten years' experience and greatly familiar with all regional RELO offerings, agreed to semi-structured interviews about the program's success. The purpose of the interviews is to discuss the fulfillment of RELO's goals of achieving public diplomacy through promotion of English and the fostering of mutual understanding between the U.S. and other countries through cultural exchange. Though Specialists were asked about their views, they have primarily worked in peripheral, limited capacities with RELO. The local on-site staff interviewed, conversely, have worked extensively and are perhaps more familiar with these programs in the Indian context than anyone involved in the project or evaluation. For their expertise and unique ability to compare this program's success to previous RELO India projects, their interviews are considered instrumental in validating this element of the *direct benefits* link.

Interviewee A is an Indian native directly involved in RELO's HE programs and is familiar with nearly all of its initiatives. She stated that she believes the intervening program did fulfill both RELO goals in that (A) it was clearly centered around the promotion of the English language, and (B) mutual understanding through cultural exchange had more of an opportunity to flourish, as it occurred to perhaps a greater extent with this program than similar projects, being the first RELO-India PD or teacher training program that spanned both time and country, and facilitated individualized communications between participants and Specialists regularly. She explained that, because of the asynchronous format of lesson observations, a larger number of professionals were able to engage closely with cultural components. The observation task forms often required participants to reflect on cultural aspects of the lessons, highlighting cultural exchange at a time when she believed it may have been otherwise overlooked within the lessons' quick-moving content. The small-group, social nature of Chai Chat meetings also allowed participants to discuss all aspects of a full course and cultural differences that emerged throughout.

Interviewee B is also native to India. She typically works with educators of economically disadvantaged students, providing support across secondary into tertiary education. Interviewee B called the response to the program "unprecedented," acknowledging that praise and participants of the program had spread into nearby countries such as Uzbekistan, Saudi Arabia, and Bangladesh, and that video links and resources had been shared as far as Peru. She believed this to be a clear demonstration of the success of public diplomacy through the promotion of the English language and a wide-reaching representation of American culture as reflected in education. Regarding mutual understanding, Interviewee B explained that, in order to assess the achievement of this goal, one must first illustrate the average English-language classroom in the Indian context as a basis for comparison, which typically consists of teacher-centered approaches, often heightened in distance learning when instructors have insufficient training in online pedagogies. The expert Specialists contracted for this initiative elicited the student engagement and interaction that has gained traction in other cultures in the realm of distance learning; participants were able to recognize the value in and observe real instances of applications of virtual strategies, making adoption of such techniques seem practical and possible, as opposed to idealistic, as is sometimes a concern following briefer PD programs.

4.5.5. Well-being changes

The assumptions underlying potential well-being changes are difficult to explore, as the intervening program under evaluation was completed so recently. In assessing whether participants are able to adapt to new or updated platforms, a certain amount of time must pass to allow for change to occur. It may perhaps be beneficial to follow-up with participants of the CERT group to investigate the topics of job security and/or new opportunities that stemmed from completing the program. In adding to the uncertainty of the validity of these assumptions, Interviewee A stated that the value of the RELO-issued certificates varies between organizations. To the best of her knowledge, they are most commonly used in fulfilling PD requirements of local institutions.

4.5.6. Evidence on other influencing factors

Potentially influential external factors include participants' involvement in outside PDs or trainings, or participants' self-guided research of and practice with new methods. To distinguish the impact of these factors from those of the intervening program, evidence is gathered from student post-surveys and program evaluations of the CERT group.

4.5.6.1. Student post-surveys

69% of student survey respondents who reported that their instructors participated in the intervening program stated they had noticeably changed their online teaching habits. All of the students who completed the survey ($n = 62$) also have instructors who did not participate in the program. Interestingly, only three of the respondents (12.5%) noted changes in the methods of professors that did not participate in the intervening program, possibly lending support to the idea that few instructors were motivated to independently research and incorporate new approaches into their lessons, though this indicative data is admittedly weak.

4.5.6.2. Program evaluation data of CERT participants

The program evaluation explicitly inquired about participants' involvement in other PD or training opportunities. 55 of 96 responding CERT participants (57%) reported engaging in other PD activities. Though this accounts for more than half of the participants, the topic of these opportunities was not addressed. Furthermore, questions in the evaluation largely emphasized growth "as a result of" the program and incorporation of tools and content "from this specific program." Similarly, when asked whether they felt more knowledgeable on web-based educational resources and teaching compared to colleagues who did NOT participate in the program, 92% of respondents answered "yes."

4.6. Step 6: Revise and strengthen the contribution story

In revising a strengthened contribution story, all steps are combined and presented succinctly in an effort to highlight the strengths and remaining weaknesses of the ToC links.

4.6.1. The revised contribution story

New evidence gathered since step 4 is presented to underscore this evaluation's areas of success and those needing additional exploration in Figure 5 below. As most links involve a mix of both weak and strong components, CA steps are instead categorized, as adapted from Mayne (2012).

Figure 5. Contribution analysis in evaluating program impact and achievement of governmental goals

<p>Steps enacted via contribution analysis resulting in sufficient validation:</p> <ul style="list-style-type: none">• Determined dropout timing and why dropout likely occurred• Gauged participants' perceived learning, understanding, and improvements in and changes to teaching• Confirmed program's usefulness and relevance via skills course student evaluations• Gathered data from Specialists regarding achievement of program and RELO goals• Conducted interviews with RELO employees regarding fulfillment of RELO goals• Established an influence of external factors, including pressure from imminent transition to distance learning and outside PD opportunities <p>Steps enacted via contribution resulting in weak validation:</p> <ul style="list-style-type: none">• Surveyed academic skills course students regarding professors' modified techniques• Identified potential uses of completion certificates <p>Additional contribution analysis steps to consider in further validating the ToC:</p> <ul style="list-style-type: none">• Survey participants' current or future students to assess sustained integration of program strategies, platforms, and/or content• Inquire further on resulting professional opportunities and long-term benefits of completion certificates and the program overall• Investigate possible effects on RELO visibility in the region, perhaps through analysis of social media traffic, numbers of registrants of upcoming projects, etc.

The revised contribution story has indeed been made stronger and more credible through the inclusion of program- and participant-specific data. Nonetheless, as pointed out in Mayne (2012), it is nearly impossible to develop a "foolproof" narrative.

4.6.2. Discussion of remaining weaknesses

The credibility of each link has been re-assessed for its evidence and logic. The program's *activities, outputs, reach and reaction*, and *capacity change* links are deemed strong and credible.

Though the *behavioral change* link remains fairly strong, there is relative weakness in evidence from sources beyond participants' self-assessments reported in the program evaluation. It is possible, also, that multiple students could be identifying changes in the same individual instructor(s) in the student post-surveys, which yielded relatively little evidence corroborating instructors' self-reports of instructional change in the first place.

The *direct benefits* link includes largely different intended outcomes from varying stakeholders, with the validation of some being more readily quantifiable than others so soon after program completion. Though self-perceived improvement in CERT participants' use of educational technology and instruction of academic skills provides somewhat strong evidence of improved education for students, more objective evidence in the form of quasi-experimental approaches would provide stronger support, if possible. Furthermore, whether heightened visibility of RELO India and its projects is verified remains uncertain.

As acknowledged in section 5.5, evidence for the *well-being changes* link is not yet available beyond Interviewee A's estimation that RELO certificates are recognized for individual institutions' PD requirements. Future data on job security and professional opportunities will likely be speculative, at best, in the form of reported perceptions in follow-up surveys and largely confounded with external influences over time. Essentially, this link is the weakest in the chain, with little to no evidence.

5. Discussion

This paper has presented a comprehensive evaluation of a multi-regional online teacher-training/PD Specialist project, funded by the US Department of State. The evaluation's purpose was to assess the project's impact in successfully influencing participants' online teaching practices in the areas of academic content and use of educational technologies, and in achieving RELO's goals of building public diplomacy through English programming and fostering mutual understanding between the U.S. and foreign countries through cultural exchange. The RUFDATA framework (Saunders, 2000) guided initial design phases of the evaluation in efforts to ensure full consideration across all program facets and key stakeholders, a number of whom were involved in the evaluation's development. Contribution Analysis (Mayne, 2001; Mayne, 2008) informed the data collection and analysis and was selected for its theory-based approach and applicability with programs that cannot feasibly be evaluated under experimental conditions. CA was iteratively employed in order to develop a plausible and validated Theory of Change (Weiss, 1995; Weiss, 1997) and to establish a strong, credible contribution story that served to infer causality of the program, as opposed to external factors, in adding value for stakeholders, as illustrated through a revised, comprehensive, and verified contribution story. This paper contributes to knowledge in that there are few if any formal evaluations of fully online teacher training or PD programs that provide participants the opportunity to both observe all aspects of an online course while also engaging with the instructor and other course contributors in small groups over an extended period of time, especially with populations straddling the digital divide. It is also the first peer-reviewed, published study, to the best of the author's knowledge, to formally evaluate a RELO-sponsored Specialist project despite the program's inception in 1991 and its current running rate of 150-200 international projects per year (U. S. Department of State, 2021).

Implications stem from both the findings of the evaluation and the evaluation's enactment. Teachers, teacher trainers, and professionals involved in planning similar programs—especially in developing regions and/or in purely virtual contexts—may note the successes and potential of distance training through observation and social connectedness while keeping in mind the impacts of challenges to infrastructure, even when disclosing accessibility requirements at registration, as well as registrants' digital literacy. Familiarity with techniques that key actors utilize on the ground to navigate such obstacles is key in maximizing participation and program value. High attrition is expected; eliminating caps on registrants or creating waitlists may help in reaching a greater number of educators.

The evaluation encountered a number of limitations. The evaluator overestimated the academic skills course students' knowledge of their instructors' participation in the intervening program, which led to weakness in the *behavioral changes* link. Evaluators should aim to further substantiate participants' self-reports in integrating program materials; follow-up with participants' current students can most accurately attest to lesson content, and comparison to feedback from instructors' previous course populations may yield truer findings. Furthermore, the

ToC proved to be complex and layered due to differing expectations of various stakeholders, including instructors and RELO representatives. This complexity may be better unpacked by applying an actor-based approach to CA as outlined in Koleros and Mayne (2019), in which researchers integrate multiple actor-based ToCs to develop clear, distinct images of actor-specific changes that should occur. Moreover, in continued iterations of CA, evaluators may investigate whether participants shared program resources or information regarding RELO initiatives with colleagues; assess RELO's potentially heightened visibility in studies on its social media traffic, reach, and future program registrations; ask for specifics on external PD activities; and inquire about professional opportunities resulting from participants' program experience.

References

- Bouzayane, S., & Saad, I. (2017). Weekly predicting the at-risk MOOC learners using dominance based rough set approach. *European Conference on Massive Open Online Courses* (pp. 160–169). Springer.
- Bureau of Educational and Cultural Affairs. (n.d.). *English Language Specialist program: At-a-glance*. Exchange programs. <https://exchanges.state.gov/us/program/english-language-specialist-program>
- Careers360. (2020, March 21). *COVID-19: Students left in the cold as colleges shut down*. <https://news.careers360.com/covid-19-students-left-in-cold-colleges-shut-down>
- Cobos, R., Wilde, W., & Zaluska, E. (2017). FutureLearn data: What we currently have, what we are learning and how it is demonstrating learning in MOOCs. *The 7th International Learning Analytics and Knowledge Conference* (pp. 1-20). Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1145/3027385.3029433>
- Coffey, M., & Gibbs, G. (2001). The evaluation of the Student Evaluation of Educational Quality questionnaire (SEEQ) in UK higher education. *Assessment & Evaluation in Higher Education*, 26(1), 89-93.
- Condon, W., Iverson, E. R., Manduca, C. A., Rutz, C., & Willett, G. (2016). *Faculty development and student learning: Assessing the connections*. Indiana University Press.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://learningpolicyinstitute.org/product/effective-teacher-professional-development-report>
- Gunter, G. A., & Gunter, R. E. (2015). *Teachers discovering computers: Integrating technology in a changing world* (8th ed.). Cengage Learning.
- Gunter, G., & Reeves, J. L. (2017). Online professional development embedded with mobile learning: An Examination of teachers' attitudes, engagement and dispositions. *British Journal of Educational Technology*, 48(6), 1305-1317.
- Hodgkinson-Williams, C. & Arinto, P. B. (2017). *Adoption and impact of OER in the Global South*. African Minds, International Development Research Centre & Research on Open Educational Resources. <https://doi.org/10.5281/zenodo.1005330>
- Koleros, A., & Mayne, J. (2019). Using actor-based theories of change to conduct robust contribution analysis in complex settings. *The Canadian Journal of Program Evaluation*, 33(3), 292-315.
- Lacey, C. H., Gunter, G., & Reeves, J. (2014). Mobile technology integration: Shared experiences from three initiatives. *Distance Learning*, 11(1), 1-8.
- Lakshmi, G., & Agarwal, M. (2017). e-Readiness of state open universities: A Study of BRAOU and UOU. *Indian Journal of Open Learning*, 26(3), 177-202.
- Li, W., Gao, M., Li, H., Xiong, Q., Wen, J., & Wu, Z. (2016). Dropout prediction in MOOCs using behavior features and multi-view semi-supervised learning. *2016 International Joint Conference on Neural Networks* (pp. 3130–3137). IEEE.
- Marsh, H. W. (1982). SEEQ: A Reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, 52(1), 77-95.
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16(1), 1-24.
- Mayne, J. (2008). Contribution analysis: An Approach to exploring cause and effect. ILAC Brief No. 16: Rome: *The Institutional Learning and Change Initiative*. http://www.cgjarilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf
- Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3), 270-280.
- Mayne, J. (2015). Useful theory of change models. *The Canadian Journal of Program Evaluation*, 30(2), 119-142.

- McDaniel, R., & Kenny, R. (2013). Evaluating the relationship between cognitive style and pre-service teachers' preconceived notions about adopting console video games for use in future classrooms. *International Journal of Game-Based Learning*, 3(2), 55–76.
- Ministry of Rural Development. (2018). *Mission Antyodaya 2017-2018: Villages with availability of electricity for domestic use*. <https://missionantyodaya.nic.in/ma2018/preloginStateElectricityReport2018.html>
- Ministry of Statistics & Programme Implementation. (2018). *Key indicators of household social consumption on education in India*. http://mospi.nic.in/sites/default/files/publication_reports/KI_Education_75th_Final.pdf
- Mukhopadhyay, A. (2020). COVID-19: Online classes and the digital divide. *Ideas for India: For More Evidence-Based Policy*. TATA Centre for Development: University of Chicago. <https://www.ideasforindia.in/topics/poverty-inequality/covid-19-online-classes-and-the-digital-divide-during-the-times-of-corona1.html>
- Naresh, B., Reddy, D., & Pricilda, U. (2016). A Study on the relationship between demographic factor and e-learning readiness among students in higher education. *Global Management Review*, 10(4), 1-11.
- Patton, M. W. (2008). *Utilization-focused evaluation* (4th ed.). Sage.
- PTI. (2020, April 21). COVID-19: Indian Internet infrastructure not prepared for shift to online teaching-learning, says QS report. *Financial Times*. <https://www.financialexpress.com/education-2/covid-19-indian-internet-infrastructure-not-prepared-for-shift-to-online-teaching-learning-says-qs-report/1935287/>
- Saunders, M. (2000). Beginning an evaluation with RUFDATA: Theorizing a practical approach to evaluation planning. *Evaluation*, 6(1), 7-21.
- United Nations Educational, Scientific and Cultural Organization. (2020, June 18). *Education: From disruption to recovery*. <https://en.unesco.org/covid19/educationresponse>
- U.S. Department of State (2021). *What is the English language specialist program?* ELPrograms. <https://elprograms.org/specialist/>
- Wajnryb, R. (1992). *Classroom observation tasks: A Resource book for teachers and trainers*. Cambridge University Press.
- Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. Connell, A. C. Kubisch, L. B. Schorr & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65-92). The Aspen Institute.
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 76, 41–55.