



Construction Industry Training Assessment Framework

Hazem Jadallah^{1*}, Carol J. Friedland¹, Isabelina Nahmens², Charles Pecquet¹, Charles Berryman¹ and Yimin Zhu¹

¹Bert S. Turner Department of Construction Management, Louisiana State University, Baton Rouge, LA, United States,

²Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA, United States

OPEN ACCESS

Edited by:

Zhen Chen,
University of Strathclyde,
United Kingdom

Reviewed by:

Alex Opoku,
University College London,
United Kingdom
Igor Peško,
University of Novi Sad Faculty of
Technical Sciences, Serbia
John Schaufelberger,
University of Washington,
United States

*Correspondence:

Hazem Jadallah
hazem426@gmail.com

Specialty section:

This article was submitted to
Construction Management,
a section of the journal
Frontiers in Built Environment

Received: 09 March 2021

Accepted: 03 August 2021

Published: 13 August 2021

Citation:

Jadallah H, Friedland CJ, Nahmens I,
Pecquet C, Berryman C and Zhu Y
(2021) Construction Industry Training
Assessment Framework.
Front. Built Environ. 7:678366.
doi: 10.3389/fbuil.2021.678366

The construction industry suffers from a lack of structured assessment methods to consistently gauge the efficacy of workforce training programs. To address this issue, this study presents a framework for construction industry training assessment that identifies established practices rooted in evaluation science and developed from a review of archival construction industry training literature. Inclusion criteria for the evaluated studies are: archival training studies focused on the construction industry workforce and integration of educational theory in training creation or implementation. Literature meeting these criteria are summarized and a case review is presented detailing assessment practices and results. The assessment practices are then synthesized with the Kirkpatrick Model to analyze how closely industry assessment corresponds with established training evaluation standards. The study culminates in a training assessment framework created by integrating practices described in the identified studies, established survey writing practices, and the Kirkpatrick Model. This study found that two-thirds of reviewed literature used surveys, questionnaires, or interviews to assess training efficacy, two studies that used questionnaires to assess training efficacy provided question text, three studies measured learning by administering tests to training participants, one study measured changed behavior as a result of training, and one study measured organizational impact as a result of training.

Keywords: workforce training, training assessment, Kirkpatrick model, training framework, construction professionals

INTRODUCTION

Formal learning and training have been shown to increase an employee's critical thinking skills and informal learning potential in any given job function (Choi and Jacobs, 2011). Evaluating training through appropriate assessment is an important aspect of any educational endeavor (Salsali, 2005), especially for assessing training efficacy in real world studies (Salas and Cannon-Bowers, 2001). Examples of training assessment abound in literature across disciplines, for both professionals and non-professionals. For example, bus drivers who attended an eco-driving course achieved a statistically significant 16% improvement in fuel economy (Sullman et al., 2015); recording engineers with technical ear training achieved a statistically significant 10% improvement in technical listening (Sungyoung, 2015); and automatic external defibrillator training of non-medical professionals resulted in a statistically significant reduction in the time to initial defibrillation by 34 s, translating in a 6% increase in survival rate (Mitchell et al., 2008).

Many advancements have been made in construction education assessment at the university level (e.g., Mills et al., 2010; Clevenger and Ozbek, 2013; Ruge and McCormack, 2017). However, within the industry itself, the dearth of workforce training research (Russell et al., 2007; Killingsworth and Grosskopf, 2013) extends to the assessment of construction industry training, particularly assessments of how learning major construction tasks affects project outcomes (Jarkas, 2010). Love et al. (2009) found that poor training and low skill levels are commonly associated with rework, which is a chronic industry problem, representing 52% of construction project cost growth (Love, 2002). Given the potential for loss within the construction industry, in both economic and life safety terms (Zhou and Kou, 2010; Barber and El-Adaway, 2015), it is reasonable to expect that integration of construction industry training assessment practices across the industry would yield improved effectiveness amongst those trained.

To understand and improve current practices for industry training assessment, the following research questions are undertaken:

- What practices have been used to assess construction industry training?
- How closely do construction industry training assessments adhere to established training evaluation standards?
- What survey science practices are typically not integrated in construction industry training?
- What practices (i.e., optimal standards) are appropriate for implementation in construction industry training program assessment?

This paper presents a framework for construction industry training assessment that identifies established practices rooted in evaluation science and developed from a review of archival construction industry training literature. The Kirkpatrick techniques (Kirkpatrick, 1959) for training evaluation serve as the foundation for the framework and relevant survey science best practices are identified and integrated. Assessment methodologies contained within the studies that meet the inclusion criteria are summarized through comprehensive case review and categorized according to the Kirkpatrick Model (Kirkpatrick, 1959) levels. The identified assessment methods are then linked with Kirkpatrick Model guidelines to analyze how closely construction industry training studies have adhered to established training evaluation standards. By analyzing the identified studies and established survey science literature, optimal standards for assessing construction industry training programs are extracted and presented within a construction industry training assessment framework.

The contribution of this research is the creation of a framework with guidelines for assessing industry training that align with the Kirkpatrick Model and have been distilled from published industry training literature and survey science best practices. The case review results and synthesis provide a current snapshot of professional construction industry training assessment criteria, identifying how closely established evaluation standards are met, and more critically, what survey

science practices are integrated in assessments. This allows for the integration of established evaluation science into training assessment practices. The intended audience of this paper is construction education and training researchers, professionals, organizations, and groups. The practical implications of this framework are its direct implementation by those conducting training, basis in sound assessment science, and practices extracted from literature.

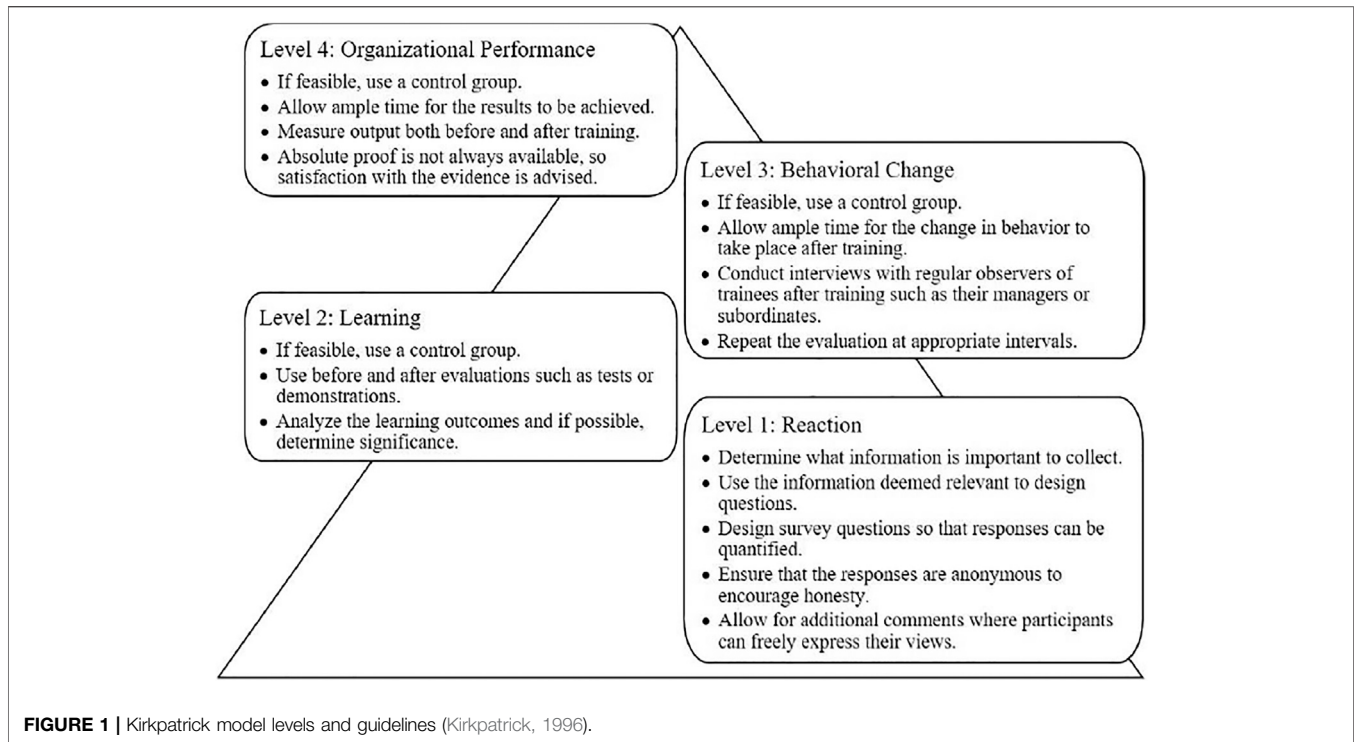
MATERIALS AND METHODS

Assessment Background

Overview of Evaluation Techniques

The reported efficacy of training has been shown to differ depending on the assessment methodology (Arthur et al., 2003), underlining the importance of the alignment of assessment levels and methods with outcome criteria. Kirkpatrick and Kirkpatrick (2016) define training efficacy as training that leads to improved key organizational results. Studies often use questionnaires after training for assessment; however, participant evaluations and learning metrics evaluate different aspects of success. Questionnaires administered directly following training tend to only measure immediate reaction to the training; therefore, to effectively evaluate training impacts beyond participant satisfaction, an assessment model is recommended. Kirkpatrick (1959) *Techniques for Evaluating Training Programs*, known as the Kirkpatrick Model, is likely the most well-known framework for training and development assessment (Phillips, 1991) and remains widely used today (Reio et al., 2017). It is comprised of four assessment levels: 1) Reaction, 2) Learning, 3) Behavior, and 4) Results.

Kirkpatrick asserts that training be evaluated using the four assessment levels described, and that these are sufficient for holistic training evaluation (Kirkpatrick, 1959). However, since its introduction, several other important evaluation models have been developed, many of which stem from the Kirkpatrick Model. For example, the input-process-output (IPO) model (Bushnell, 1990) begins by identifying pre-training components (e.g., training materials, instructors, facilities) that impact efficacy as the input stage. The process stage focuses on the design and delivery of training programs. Finally, the output stage essentially covers the same scope as the Kirkpatrick Model. Brinkerhoff (1987) six-stage evaluation model goes beyond assessment into training design and implementation. The first stage identifies the goals of training and the second stage assesses the design of a training program before implementation. The remaining four stages fall in line with Kirkpatrick's four levels. Kaufman and Keller (1994) present a five-level evaluation model where Level 1 is expanded to include enabling, or the availability of resources, as well as reaction; Levels 2 through 4 match the corresponding levels in the Kirkpatrick Model; Level 5 goes beyond the organization and presents a method of evaluating the training program on a societal level. Phillips (1998) presents a five-level model that adopts Kirkpatrick's first three levels and expands the fourth level by identifying ways that organizations can assess organizational impact. A fifth level is added that evaluates the true



return on investment (ROI) by comparing the cost of a training program with the financial gain of organizations implementing training.

While developing and designing effective programs are important, these criteria fall outside the scope of this study; which focuses on training assessment implementation and not evaluating the suitability of aspects of the training programs reviewed. Therefore, the Bushnell and Brinkerhoff models have no advantage above the Kirkpatrick Model for this analysis. Similarly, there is not enough information provided in the identified studies regarding social implications as a result of training to warrant use of Kaufman and Keller's or Phillips's five-level models as a basis. From an assessment aspect, the reviewed models essentially stem from and adhere to the four levels found in the Kirkpatrick Model. Because the focus of this research is the assessment of construction industry training programs, and not the design and development of training, the Kirkpatrick Model is well-suited for robust synthesis and extraction of optimal standards for training evaluation methodologies and is therefore used in this study.

The Kirkpatrick Model

Kirkpatrick (1996) asserts that the 1959 model is widely used because of its simplicity. Amongst the population of training professionals, there is little interest in a complex scholarly approach to training assessment. Definitions and simple guidelines are presented in the model to facilitate straightforward implementation (Figure 1). The following paragraphs describe each level in more detail.

Level 1: Reaction Within the first level, overall trainee satisfaction with the instruction they have received is

measured. While all training programs should be evaluated at least at this level (Kirkpatrick, 1996), learning retention is not measured here. Participant reactions are perceived to be easily measured through trainee feedback or survey question answers (Sapsford, 2006); therefore, surveys are a common means of assessment. From a robust reaction analysis, program designers assess training acceptance and elicit participant suggestions and comments to help shape future training sessions.

Level 2: Learning Within the second level, trainee knowledge gain, improved skills, or attitude adjustments resulting from the training program are measured. Because measuring learning is more difficult than measuring reactions (Level 1), before-and-after evaluations are recommended. These may include written tests or demonstrations measuring skill improvements. Analysis of learning assessment data and use of a control group are recommended to determine the statistical significance of training on learning outcomes, when possible.

Level 3: Behavior Within the third level, the extent to which training participants change their workplace behavior is measured. For behavior to change, trainees must recognize shortcomings and want to improve. Evaluation consists of participant observation at regular intervals following the training, allowing ample time for behavior change to occur. External longitudinal monitoring is more difficult than assessment practices in the previous two levels. A control group is recommended.

Level 4: Results Within the fourth level, the effect that training has on an overall organization or business is measured. Many organizations are most interested (if not only interested) in this level of evaluation (Kirkpatrick, 1996). In fact, "The New World Kirkpatrick Model" (Kirkpatrick and Kirkpatrick, 2016) asserts

that training programs should be designed in reverse order from Level 4 to Level 1 to keep the focus on what organizations value most. Common assessment metrics are improved quality, increased production, increased sales, or decreased cost following training. A control group is recommended.

Survey Science Best Practices

Multiple studies have focused on proper formulation of survey questions that can be used across industries. Lietz (2010) summarized the literature regarding questionnaire design, focusing on best practices such as question length, grammar, specificity and simplicity, social desirability, double-barreled questions, negatively worded questions, and adverbs of frequency. With regards to question length, Lietz (2010) recommends short questions to increase respondents' understanding. Complex grammar should be minimized and pronouns should be avoided. Simplicity and specificity should be practiced to decrease respondents' cognitive effort. Complex questions should be avoided and instead separated into multiple questions. Definitions should be provided within the question to give context. For example, a "chronic" health condition means seeing a doctor two or three times for the same condition (Fowler, 2004). The scale used to gauge responses with should also follow the concept of simplicity. Taherdoost (2019) found that while scales of 9 and 10 are thought to increase specificity, reliability, validity, and discriminating power were indicated to be more effective with scales of 7 or less. Social desirability may result in respondents' answering questions based on their perception of a position favored by society. To remedy this bias, Brace (2018) suggests asking questions indirectly, such as "What do you believe other people think?" where respondents may be more likely to admit unpopular views. "Doubled-barreled" questions contain two verbs and should be avoided. Negatively worded questions should similarly be avoided to clarify the meaning. This is particularly the case when the words "no" or "not" are used together with words that have a negative meaning such as "unhelpful." Finally, adverbs such as "usually" or "frequently" should be avoided and replaced with actual time intervals such as "weekly" or "monthly."

Methodology

The methodology consists of three steps:

1. Relevant literature is identified through inclusion criteria; case review is performed to extract and summarize key assessment aspects.
2. Identified construction assessment methodologies are evaluated against the corresponding Kirkpatrick Model level guidelines.
3. An assessment framework is constructed that integrates optimal assessment standards aligned with the Kirkpatrick Model.

Study Selection and Evaluation

A structured literature review is implemented to collect data describing construction industry training assessment for current

industry professionals. The objective is to understand how various construction industry training programs that have embedded established educational theory in their design or implementation assess training efficacy. Educational theory-embedded training was selected because it is indicative of a more robust training assessment. Peer reviewed archival literature is searched to determine the state of construction industry training studies that have been documented in scholarly works.

The main search keywords were "construction industry," "education theory," and "training." The main research engines were EBSCOhost library services and Google Scholar; and they were used to identify relevant studies. The following inclusion criteria were established to identify recent, relevant peer-reviewed construction industry training studies published after 2005 for investigation in this study:

1. The training focuses on the current construction industry workforce, including construction workers (W), project managers (M), and designers (D).
2. The training incorporated educational theory in its creation or implementation.

Using the keywords mentioned above, a literature search was conducted resulting in 475 research studies, which increased to 483 through identification of other sources referenced in the initial search results. After removing duplicates and applying the inclusion criteria and additional quality measures, 15 publications were identified for the review, indicating limited research conducted in this area. The selection process is illustrated in the Preferred Items for Systematic Review Recommendations (PRISMA) (Moher et al., 2009) flow chart in **Figure 2**.

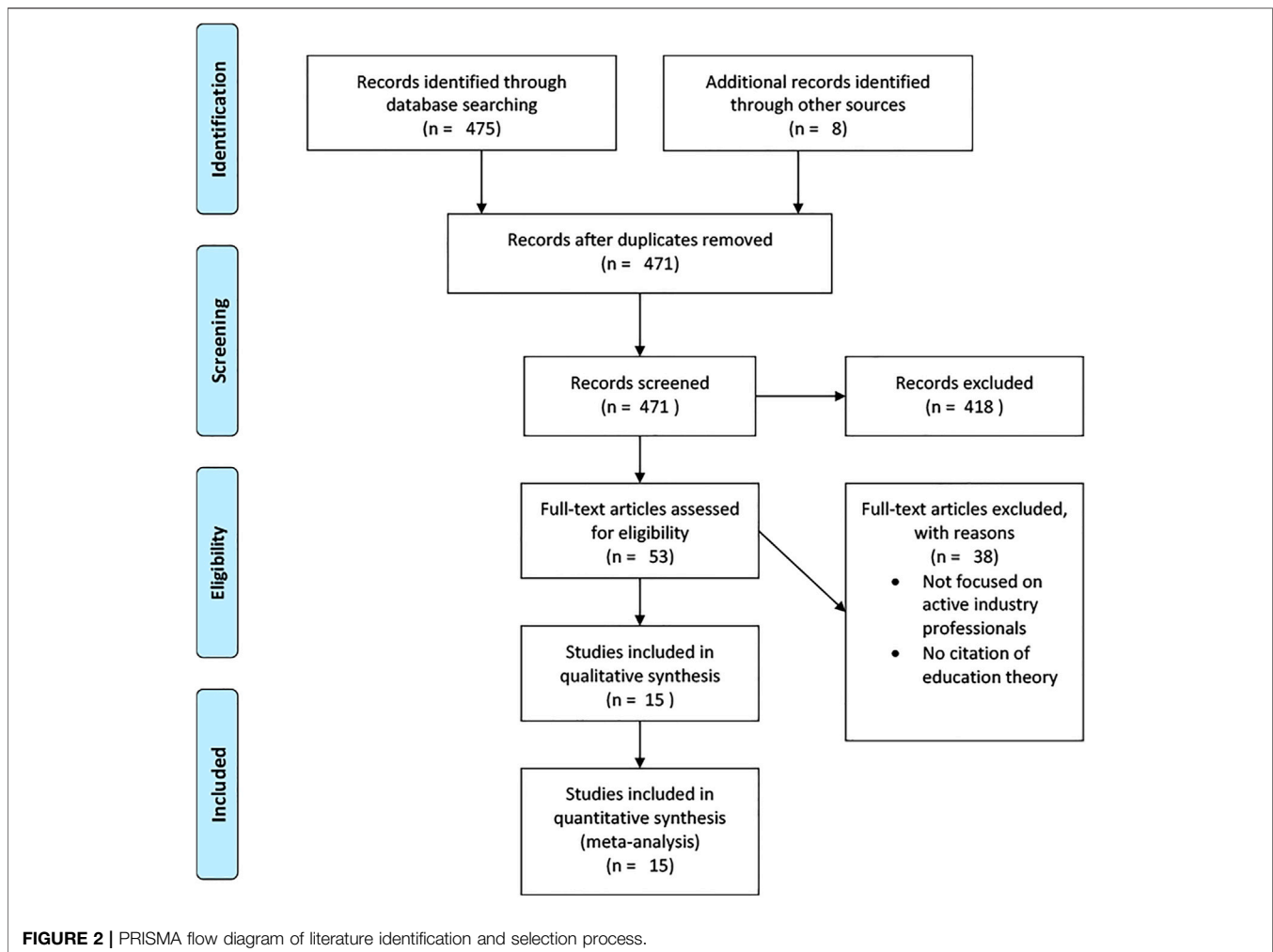
The following information was recorded from the relevant publications that met the inclusion criteria: location (i.e., country) where the study took place, educational theory employed, training subject, assessment level corresponding to the Kirkpatrick Model, and assessment methodology. Assessment tools were often referred to as questionnaires, surveys, or interviews. Each of these assessment types was recoded as "questionnaires." A case review summarizes the methods, assessment criteria, and results of the studies identified. The case review is created to provide context of the studies.

Kirkpatrick Model Synthesis

The assessment methodologies within the identified studies were linked to the corresponding guidelines established by the Kirkpatrick Model. The assessment methods within each training program study were evaluated, first to determine the corresponding Kirkpatrick Level, and second to identify adherence to the Kirkpatrick guidelines (Kirkpatrick, 1996) for each level.

Survey Science Synthesis

The identified studies that provided the text of the questionnaires administered to training participants were evaluated against the survey science best practices summarized by Lietz (2010). The



total occurrence of each practice is enumerated so that more common practices are identified.

Construction Industry Training Assessment Framework

The assessment review culminates in the presentation of a framework of optimal practices identified through the synthesis of assessment criteria used in the construction industry training studies and survey science best practices, aligned with the Kirkpatrick Model. The framework includes a summary of Kirkpatrick Model guidelines and practices resulting from the synthesis of identified construction literature and established survey science.

RESULTS, ANALYSIS, AND DISCUSSIONS

Study Selection and Evaluation

Fifteen studies describing education theory-integrated construction industry training met the inclusion criteria selected, listed in alphabetical order in **Table 1**. A short summary of assessment criteria used in each study is provided in the following case review and corresponding ties to the Kirkpatrick Model are established.

Study Number 1

Akanmu et al. (2020) implemented a virtual reality (VR) training focused on reducing construction worker ergonomic risks. The primary assessment method was participant feedback through a questionnaire with both rated questions (1 = strongly disagree, 5 = strongly agree) and open-ended questions, meeting Level 1 standards. Rating questions gauged whether the user interface for the postural training program interfered with the work surface (mean = 2.4), whether the virtual reality display affected performance (mean = 2.7), whether the display was distracting (mean = 1.3), and whether the avatar and color scheme enhanced their understanding of ergonomic safety (mean = 1.2). In open-ended questions, 9 out of 10 participants reported that the VR training helped adjust posture. Two out of ten participants complained that the wearable sensors obstructed movement. The study did not publish the assessment questions directly, and only provided results; therefore, they were not analyzed for survey science best practices outlined by Lietz (2010). It should be noted that mean scores of 1.3 and 1.2 do not appear to be positive as they favor the strongly disagree rating based on the key provided. Additionally, the exact open-ended question text is

TABLE 1 | Construction industry training studies reviewed.

Study number and reference	Country	Educational theory	Subject	Kirkpatrick level	Assessment methodology
1 Akanmu et al. (2020)	United States	Digital game-based learning	Ergonomic safety	1	Questionnaire
2 Begum et al. (2009)	Malaysia	Ajzen's theory	Waste management	1	Questionnaire
3 Bena et al. (2009)	Italy	Andragogy	Safety	4	Injury monitoring
4 Bhandari and Hallowell (2017)	United States	Andragogy	Safety and risk perception	1	Questionnaire
5 Bressiani and Roman (2017)	Brazil	Andragogy	Masonry brick laying	1	Questionnaire
6 Choudhry (2014)	China	Behaviorism	Safety	3	Questionnaire
7 Douglas-Lenders et al. (2017)	Australia	Self-efficacy	Leadership training for project managers	1	Questionnaire
8 Eggerth et al. (2018)	United States	Andragogy	Safety	1	Questionnaire
9 Evia (2011)	United States	The Kirkpatrick model	Safety	1	Questionnaire
10 Forst et al. (2013)	United States	Andragogy	Safety	2	Questionnaire
11 Goulding et al. (2012)	United Kingdom	Digital game-based learning	Offsite production	1	Feedback
12 Mehany et al. (2019)	United States	Long term retention	Confined space training	2	Testing
13 Lin et al. (2018)	United States	Andragogy	Safety	1,2	Questionnaire
14 Lingard et al. (2015)	Australia	Visual pedagogy	Construction health and safety	1	Video-based intervention
15 Wall and Ahmed (2008)	Ireland	Blended learning	Project management	1	Computer simulation implementation

TABLE 2 | Connections to the Kirkpatrick guideline from 15 construction industry training studies.

Kirkpatrick level	Attributes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total
1: Reaction ^a	Design survey questions so that responses can be quantified	x	x	—	x	x	—	x	x	x	—	—	—	x	—	—	8
	Ensure that the responses are anonymous to encourage honesty	x	x	—	x	x	—	x	x	x	—	—	—	x	—	—	8
	Allow for additional comments where participants can freely express their views	x	—	—	—	—	—	—	—	—	—	—	—	—	x	—	2
2: Learning	If feasible, use a control group	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
	Use before and after evaluations such as tests or demonstrations	—	—	—	—	—	—	—	—	—	x	—	x	x	—	—	3
	Analyze the learning outcomes and if possible, determine significance	—	—	—	—	—	—	—	—	—	x	—	x	x	—	—	3
3: Behavioral Change	If feasible, use a control group	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
	Allow ample time for the change in behavior to take place after training	—	—	—	—	—	x	—	—	—	—	—	—	—	—	—	1
	Conduct interviews with regular observers of trainees after training such as their managers or subordinates	—	—	—	—	—	x	—	—	—	—	—	—	—	—	—	1
4: Organizational Performance	Repeat the evaluation at appropriate intervals	—	—	—	—	—	x	—	—	—	—	—	—	—	—	—	1
	If feasible, use a control group	—	—	x	—	—	—	—	—	—	—	—	—	—	—	—	1
	Allow ample time for the results to be achieved	—	—	x	—	—	—	—	—	—	—	—	—	—	—	—	1
	Measure output both before and after training	—	—	x	—	—	—	—	—	—	—	—	—	—	—	—	1
	Absolute proof is not always available, so satisfaction with the evidence is advised	—	—	x	—	—	—	—	—	—	—	—	—	—	—	—	1

^aNote: The first two guidelines associated with Level 1 are not listed as they were not contained in the studies.

not provided, and the article states that they are asked to encourage improvement of training in the future. This does not follow established survey guidelines, as this question will not yield quantifiable results.

Study Number 2

Begum et al. (2009) administered a survey to local contractors in Malaysia to measure the attitudes and behaviors of contractors

toward waste management, categorizing this assessment as Level 1. The results found a positive regression coefficient ($\beta = 2.006$; $p = 0.002$) correlating education to contractor waste management attitude; making education one the most significant factors found in the study. The study did not provide the actual questions asked on the questionnaire, but instead stated that the following “attitudes” were assessed: general characteristics, such as contractor type and size; waste collection and disposal systems;

TABLE 3 | Survey science best practices connection to construction industry training studies.

Survey question best practices	1	2	4	5	7	8	9	11	13	14	15	Total (%)
Survey questions provided	x	x	x	✓	x	✓	x	x	x	x	x	18
Question length	-	-	-	✓	-	✓	-	-	-	-	-	100
Grammar	-	-	-	✓	-	✓	-	-	-	-	-	100
Specificity and simplicity	-	-	-	✓	-	✓	-	-	-	-	-	100
Social desirability	-	-	-	x	-	✓	-	-	-	-	-	50
Double-barreled questions	-	-	-	✓	-	✓	-	-	-	-	-	100
Negatively worded questions	-	-	-	✓	-	✓	-	-	-	-	-	100
Adverbs of frequency	-	-	-	✓	-	✓	-	-	-	-	-	100
Response scale is reasonable	✓	-	x	x	✓	-	x	-	-	-	-	40
Quantifiable results ^a	✓	✓	✓	✓	✓	✓	x	x	✓	x	x	64
Allowing for additional comments ^a	✓	-	-	x	x	-	x	-	-	-	-	25

^aNote: ✓ indicates the best practice was met; 'x' indicates the best practice was not met; '-' indicates adherence to the best practice could not be assessed.

waste sorting, reduction, reuse and recycling practices; employee awareness; education and training programs; attitudes and perceptions toward construction waste management and disposal; behaviors with regard to source reduction and the reuse and recycling of construction waste. With this information, it is difficult to determine how closely questionnaire guidelines were followed.

Study Number 3

Bena et al. (2009) assessed the training program delivered to construction workers working on a high-speed railway line in Italy. The assessment analyzed injury rates for workers before and after training and found that the incidence of occupational injuries fell by 16% for the basic training module, and by 25% after workers attended more specific modules. This is a Level 4 evaluation because the overall organizational outcomes were assessed.

Study Number 4

Bhandari and Hallowell (2017) proposed a multimedia training that integrated andragogy (i.e., adult learning) principles to demonstrate the cause and effect of hand injuries during construction situations, focusing on injuries caused by falling objects and pinch-points. A questionnaire asked participants to rate the intensity of different emotions using a 9-point Likert scale both before and after the training simulation was distributed. Overall, workers reported a statistically significant increase in negative emotions such as confusion ($p = 0.01$), fear ($p = 0.01$), and sadness ($p = 0.01$) after they had been trained. Statistically significant decreases in positive emotions such as happiness ($p = 0.01$), joy ($p = 0.01$), love ($p = 0.01$), and pride ($p = 0.01$) were also reported by trainees. Because gauging trainee response are the main assessment tool, this is classified as a Level 1 evaluation. In total, eighteen emotions were assessed, making the survey rather lengthy and possibly inducing cognitive fatigue or confusion. Additionally, a 9-point Likert scale adds a wide range of possible options to choose from, which is higher than the recommendation by Taherdoost (2019) of a 7-point scale. A shorter survey with fewer options might improve the results generated by this study.

Study Number 5

Bressiani and Roman (2017) used andragogy to develop a training program for masonry bricklayers. Questionnaires used to assess the participant feedback found that andragogical principles were met in more than 92% of responses. Because gauging trainee response are the main assessment tool, this is classified as a Level 1 evaluation. The study presented training participants with a 24-question survey found in the appendix of their study. The questions themselves are short, simple, and pertain to a singular topic, complying with survey best practices. However, the response options are given on a 0–10 scale. Similar to Bhandari and Hallowell's 9-point scale, this number of response choices can add confusion and complexity when respondents answer the questions.

Study Number 6

Choudhry (2014) implemented a safety training program based on behaviorism. Safety observers monitored the use of personal protective equipment (PPE) such as safety helmets, protective footwear, gloves, ear defenders, goggles or eye protection, and face masks over a 6-week period. Safety performance in the form of utilization of PPE increased from 86%, measured 3 weeks after training, to 92.9%, measured 9 weeks after training. This is classified as a Level 3 evaluation because behavior changes were observed and noted. Further, external observers were used and data were collected over time, adhering to Kirkpatrick Level 3 guidelines.

Study Number 7

Douglas-Lenders et al. (2017) found an increase in self-efficacy of construction project managers after a leadership training program was administered. This assessment was conducted through a questionnaire that presented questions on a 5-point Likert scale; which was used to gauge trainee self-perception as a result of training. Learning confidence, learning motivation, and supervisor support received average scores of 4.23, 3.86, and 3.84 respectively from training participants. Because surveys are the main assessment criteria this is classified as a Level 1 evaluation. The study did not publish the assessment questions directly, and

only provided results; therefore, they were not analyzed for survey science best practices.

Study Number 8

Eggerth et al. (2018) evaluated safety training “toolbox talks,” which are brief instructional sessions on a jobsite or in a contractor’s office. The study involves a treatment group that experienced training, as well as a control group answered a questionnaire. The trained group rated the importance of safety climate statistically significantly higher than the control group ($p = 0.026$). Because gauging trainee response are the main assessment tool, this is classified as a Level 1 evaluation. Sample questions are recorded in the study, however, the questionnaire in its entirety is not presented. However, based on the sample questions, it is likely that the questionnaire generally falls in line with survey standards.

Study Number 9

Evia (2011) evaluated computer-based safety training targeted toward Hispanic construction workers. Based on interviews with the participants, a positive reaction to the training with significant knowledge retention was achieved. This study also did not present the questionnaire in its entirety; however, it is mentioned that the evaluation measured reaction. Workers were able to give ratings such as “very interesting,” and “easy” with regards to a video watched during the training; however no numerical assessment was given. Because gauging trainee response are the main assessment tool, this is classified as a Level 1 evaluation. The study did not publish the assessment questions directly; therefore, they were not analyzed for survey science best practices.

Study Number 10

Forst et al. (2013) evaluated a safety training targeted toward Hispanic construction workers in seven cities across the United States. Questionnaires that were administered to the training participants indicate demonstrated improvements in safety knowledge. The results found a statistically significant knowledge gain for the questions regarding fall prevention and grounding from the pre-training and post-training questionnaires ($p = 0.0003$). This type of evaluation is classified as Level 2 because the learning outcomes of training were measured. The pre-training and post-training testing guidelines appear to have been met throughout this study.

Study Number 11

Goulding et al. (2012) present the findings of an offsite production virtual reality training prototype. Feedback of training was requested, and the feedback was summarized as being positive. Because gauging trainee response are the main assessment tool, this is classified as a Level 1 evaluation. No numerical assessment was provided and the study did not publish the assessment questions directly; therefore, they were not analyzed for survey science best practices.

Study Number 12

Mehany et al. (2019) evaluated a confined space training program administered to construction workers. A test was administered to

the training participants and the results found that the participants scored below average, even after attending the training on the subject. A score of 11/15 is taken to be the United States national average. The participants scored an average of 9.3/15. This average was further broken into a non-student sample (industry professionals) that scored an average mean of 8.3 and a student sample that scored 9.5. This is classified as a Level 2 evaluation because the learning outcomes of training were measured. Diversity in the population of examinees provided the authors with interesting analysis opportunities and the ability to speculate on the difference in scores between the two groups, which is desirable in learning evaluations.

Study Number 13

Lin et al. (2018) used a computer-based three-dimensional visualization technique, designed by adult education subject matter experts, to train Spanish-speaking construction workers on safety and fall fatality. Interviews were conducted to evaluate the training program. 64–90% of English-speaking workers achieved the intend results, 73–83% of Spanish-speaking workers achieved the intended results. 100% of Spanish-speaking workers reported that they would recommend the training materials to others while only 46% of English-speaking workers reported that they would recommend the training materials to others. Because both interviews and tests were conducted this is classified as a Level 1 and Level 2 evaluation. From a Level 1 perspective the study presents the results in an “evaluation of validation” format without referencing the exact questions asked. This makes it difficult to assess how closely question format guidelines were followed. From a Level 2 perspective a set of questions to assess knowledge gain is presented. Both English and Spanish speaking participants were tested. Six questions were included on the test to assess participant knowledge gain after the training. Similar to the previous study, the diversity in the populations provides analysis opportunities to assess learning outcomes as a result of training.

Study Number 14

Lingard et al. (2015) evaluated the use of participatory video-based training to identify safety concerns on a construction jobsite. As a result of this training, new health and safety rules were generated by participants. The training was based on viewing the recordings and success was measured by workers’ ability to establish new safety guidelines to enable compliance. Because feedback was taken into consideration this is classified as a Level 1 evaluation. This study culminated in the participants sharing their reactions to the training in a group setting. While the reactions were captured, the study did not publish the assessment questions directly; therefore, they were not analyzed for survey science best practices.

Study Number 15

Wall and Ahmed (2008) explore a training delivered to Irish construction project managers on construction management computerized tools. Participants reported the program increased their understanding of construction problems and

decisions. Because participant feedback was gathered this is classified as a Level 1 evaluation. However, the study did not capture participant responses in an explicit way, but rather it was presented that feedback was favorable and no numerical assessments were presented.

Case Review Summary

This case review found that ten studies (67%) used surveys, questionnaires, or interviews to assess the training programs, three studies (20%) measured learning by administering tests to training participants, one study measured changes in behavior resulting from training, and one study measured organizational impact a result of training. Attributes of the assessment methodologies that complied with Kirkpatrick standards or established survey science best practices were noted as positively complying with Level 1 assessment standards, which are summarized in the survey science synthesis. Studies that complied with Level 2–4 standards typically complied with the guidelines set forth by Kirkpatrick, however it is surprising that so few studies utilized these methodologies. This is especially the case with Level 4 evaluation standards. Organizations ultimately seek to understand how training might impact performance on an organizational level; yet of the 15 studies analyzed, one complied with this standard of evaluation. Gaps identified in the review of the studies inspired the guidelines outlined in the Construction Industry Assessment Framework presented in this paper.

Kirkpatrick Model Synthesis

Although the first two Level 1 guidelines were excluded from the analysis, amongst the remaining three Level 1 guidelines, one study (Akanmu et al., 2020) included all three assessment guidelines, while seven studies met two Level 1 guidelines, and one study met one Level 1 guideline. The three studies that met Level 2 guidelines were identical in that they excluded the use of a control group and adhered to all other guidelines. Similarly, the only study (Choudhry, 2014) that met Level 3 guidelines excluded the use of a control group and adhered to all other guidelines. One study (Bena et al., 2009) provided a Level 4 evaluation that met all associated guidelines. This information is shown in **Table 2**.

Survey Science Synthesis

Of the studies that used Level 1 criteria for their assessment methodology, two (18%) provided the text of the survey questions presented to training participants. The remaining studies did not publish the assessment questions directly. Bressiani and Roman (2017) presented the questionnaire in its entirety. All survey science recommendations summarized by Lietz (2010) were met except for guarding against social desirability, implementing a reasonable response scale, and allowing for additional comments. Eggerth et al. (2018) only presented sample questions from the questionnaire distributed to participants, however, all survey recommendations that could be analyzed were met. Analysis of the response scale reveals that of the five studies that provided their scales, two (40%) adhered to optimal scale standards of seven or less. 64% of studies provided results that could be quantified. 25% of studies that were analyzed for allowing additional comments were found to have done so.

The percentage was derived by dividing the number of times a practice was met by the number of times a practice was not met. When a practice could not be assessed for a study, this field was excluded from the calculation. This information is shown in **Table 3**.

Construction Industry Training Assessment Framework

Survey results may be skewed by the questions asked (Dolnicar, 2013), and poorly written questions often result in flawed data (Artino, 2017). When one considers that most construction industry training studies evaluate efficacy by attempting to collect the reaction of participants, it is important that the questions asked be made available for future study and analysis. For this reason, the framework provides extensive recommendations to improve Level 1 analyses. Additionally, because only 20% of studies that used questionnaires as their means of assessment provided the questionnaire text, the current adherence of Level 1 construction industry training assessment best practices remains widely unknown. Moving forward, it is of the utmost importance that this information be provided to support robust Level 1 assessment. Additionally, Taherdoost (2019) recommends a 7-point Likert type scale as to not overwhelm participants with a high number of response options. When composing open-ended questions, efforts should be made to frame the questions in a way that will yield results that are quantifiable. While analysis of open-ended questions is rare, the results can be very valuable (Roberts et al., 2014). Due to the lack of complete survey question text included in most studies, it is recommended that survey questions be contained within training studies so that the results can be fully analyzed.

The simplest method for analyzing learning development as a result of a training program is an evaluation to be administered before and after a training program (Kirkpatrick, 1996). Kirkpatrick recommends the use of a control group. However, in literature it was observed that a control group was rare. Cost, resources, and time could be contributing factors, however, for the sake of analysis these circumstances should be made clear. The study presented by Mehany et al. (2019) measured learning outcomes against an industry wide average, which provides a benchmark for the results of a given training program. If possible, this should be the norm, as it gives a standard by which a given training program is analyzed. Several studies analyzed the evaluation results for statistical significance. This should be done when possible to lend more credibility to the results.

To measure the extent to which training participants change their workplace behavior, observations are collected over time. Similar to the learning level, a rationale should be provided when a control group is not used. The study presented by Choudhry (2014) details the intervals at which observations are made. This should be standard practice and measurements at these intervals should be reports so that a progression can be seen. Additionally, it is known that people may change their behavior unexpectedly if they know that they are being observed (Harvey et al., 2009), and for this reason, observations should be made as inconspicuously as possible.

TABLE 4 | Framework for construction industry training assessment**Level 1: Reaction**

Design survey questions that will ensure the collection of relevant data from participants in a manner that can be quantified, allowing for anonymity and additional participant feedback

- To provide justification for survey results, present the process of identifying relevant information to be gathered by the surveys
- Generate questions that will encourage training participants to provide information that is relevant to the training designers
- Adhere to survey science best practices outlined in this paper
- Develop questions so that results may be quantified. Likert type scales should be no more than seven points to avoid confusion of participants
- While open-ended questions are encouraged, they should be framed in a way so that the responses are quantifiable
- Include survey question text in descriptions of the training (e.g., journal publications) to add to the body of knowledge

Level 2: Learning

Create evaluations for training participants that can be completed before and after a given training to measure learning progress. Analyze the results and determine the statistical significance of changes in knowledge

- Rationalize the lack of a control group if one is not utilized
- If possible, determine an industry average of test results to compare the results of trainees to the average of the overall industry
- Analyze the learning outcomes for statistical significance for each individual question so that specific learning outcomes can be identified, and improvement can be made

where no significance is found

Level 3: Behavioral Change

After the allotment of ample time for participants to change their behavior following training, conduct observations and interviews with regular observers to quantify the change in behavior, repeating the evaluation at appropriate intervals

- Rationalize the lack of a control group if one is not utilized
- Provide time intervals of when behavioral observations occur so change in behavior can be monitored over time
- If possible, monitor behavioral changes discretely so that participants are not only changing their behavior when they are being observed

Level 4: Organizational Performance

After allowing ample time for results to be achieved, measure the output before and after training

- Rationalize the lack of a control group if one is not utilized
- Generate a metric for organizational performance prior to training implementation so data can be more easily collected
- Be sure to note pre-training performance levels so changes in performance can be measured
- Identify other factors that may contribute to changes in performance to isolate the effect of training as a factor

When measuring organizational performance, the same care to rationalize the lack of a control group should be included in a training study; as is the recommendation for the Learning and Behavioral Change levels. While Kirkpatrick includes common metrics for measuring training effectiveness at this level such as decreased cost or increased revenue, these metrics are not always clearly defined. The metric by which an organization would like to measure effectiveness should be clearly identified in a training study. To accurately organizational change, pre-training levels must be noted. Kirkpatrick (1996) notes that factors other than training may also affect overall organizational performance. These factors should be identified and noted in a training study.

With this information in mind, the construction industry training framework (Table 4) is aligned using Kirkpatrick Model guidelines with the additional knowledge acquired by the synthesis of the identified studies and survey science best practices. Gaps found in the studies, such as the lack of information surrounding how survey questions were chosen, contribute to the framework by emphasizing this type of information that was notably missing across all studies analyzed.

CONCLUSION

This study provides a comprehensive literature review of educational theory-integrated construction industry training focusing on assessment methodologies used in construction industry training literature. Assessment practices identified through case review were compared against the Kirkpatrick Model, a well-known and widely used assessment model.

Assessment methodologies in the literature were synthesized with corresponding levels found in the Kirkpatrick Model to analyze how closely the industry adheres to established training evaluation standards. The studies that utilized questionnaires as their means of assessment and provided the text of the questions asked were evaluated against survey science best practices. This study culminates in the creation of a training assessment framework by extracting the practices used in the identified studies so that future assessment methodologies can be implemented, tested, and presented effectively, thus advancing construction industry training. The specific findings of this study are that two-thirds (67%) of identified studies used surveys, questionnaires, or interviews to assess training efficacy. Of the studies that met the inclusion criteria, 73% (11/15) were designed to assess reaction, 20% (3/15) assessed learning, and 7% (1/15) assessed each behavior and organizational impact. Kirkpatrick Levels 2 to 4 assessments implemented in construction literature typically met the Kirkpatrick guidelines; however, Level 1 guidelines were met by 18% (2/11) of the studies. Two of the ten studies (20%) that used questionnaires to assess training efficacy provided question text, and of these, one study followed survey science best practices completely. The following survey science best practices are typically not integrated: accounting for social desirability, implementing a reasonable response scale, and allowing for additional comments. Finally, archival construction industry training literature and survey science best practices were synthesized and aligned with Kirkpatrick (1959) *Techniques for Evaluating Training Programs* to create a framework for construction industry training assessment.

The issue of assessment methodologies is found within archival published literature and appears to be an industry-wide issue. Opportunity exists to implement training programs coupled with optimal assessment methodologies grounded in established educational assessment research. Further opportunities exist to present techniques for measuring organizational outcomes (Level 4), as only one of fifteen studies reviewed used this criterion to assess training. The findings of this research indicate that there is an opportunity to introduce more robust metrics prior to training implementation to assess training at the organizational level, rather than relying on Level 1 through 3 assessment results.

This paper is relevant to the current state of construction industry assessment by presenting a proposed construction industry assessment framework modified from the original Kirkpatrick Model to address gaps found in the model, identified best practices, and relevant practices found in the studies analyzed throughout this paper. While the assessment strategy proposed in this paper is based on best practices, such as survey science best practices to measure reaction, as well as identified practices that measure learning, behavior, and

organizational performance, future research is needed to apply the proposed framework to training programs so that its efficacy may be demonstrated.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

HJ provided the conceptual design of this research, conducted the literature review, collected data, developed the initial text, and revised the text through input of the co-authors. CF provided conceptualized the overall project methodology and presentation of results and contributed significantly to draft revisions. IN, CP, CB, and YZ revised early and late versions of the text and provided regular feedback and guidance.

REFERENCES

- Akanmu, A. A., Olayiwola, J., Ogunseju, O., and McFeeters, D. (2020). Cyber-Physical Postural Training System for Construction Workers. *Automation in Construction* 117, 103272. doi:10.1016/j.autcon.2020.103272
- Arthur, W., Tubre', T., Paul, D. S., and Edens, P. S. (2003). Teaching Effectiveness: The Relationship between Reaction and Learning Evaluation Criteria. *Educ. Psychol.* 23, 275–285. doi:10.1080/0144341032000060110
- Artino, A. R., Jr (2017). *Good Decisions Cannot Be Made from Bad Surveys*. Oxford, United Kingdom: Oxford University Press.
- Barber, H. M., Jr., and El-Adaway, I. H. (2015). Economic Performance Assessment for the Construction Industry in the Southeastern United States. *J. Manage. Eng.* 31, 05014014. doi:10.1061/(asce)me.1943-5479.0000272
- Begum, R. A., Siwar, C., Pereira, J. J., and Jaafar, A. H. (2009). Attitude and Behavioral Factors in Waste Management in the Construction Industry of Malaysia. *Resour. Conservation Recycling* 53 (6), 321–328. doi:10.1016/j.resconrec.2009.01.005
- Bena, A., Berchiolla, P., Coffano, M. E., Debernardi, M. L., and Icardi, L. G. (2009). Effectiveness of the Training Program for Workers at Construction Sites of the High-Speed Railway Line between Torino and Novara: Impact on Injury Rates. *Am. J. Ind. Med.* 52 (12), 965–972. doi:10.1002/ajim.20770
- Bhandari, S., and Hallowell, M. (2017). Emotional Engagement in Safety Training: Impact of Naturalistic Injury Simulations on the Emotional State of Construction Workers. *J. construction Eng. Manag.* 143 (12). doi:10.1061/(ASCE)CO.1943-7862.0001405
- Brace, I. (2018). *Questionnaire Design: How to Plan, Structure and Write Survey Material for Effective Market Research*. London, United Kingdom: Kogan Page Publishers.
- Bressiani, L., and Roman, H. R. (2017). A utilização da Andragogia em cursos de capacitação na construção civil. *Gest. Prod.* 24 (4), 745–762. doi:10.1590/0104-530x2245-17
- Brinkerhoff, R. O. (1987). *Achieving Results from Training: How to Evaluate Human Resource Development to Strengthen Programs and Increase Impact*. San Francisco: Pfeiffer.
- Bushnell, D. S. (1990). Input, Process, Output: A Model for Evaluating Training. *Train. Develop. J.* 44 (3), 41–44.
- Choi, W., and Jacobs, R. L. (2011). Influences of Formal Learning, Personal Learning Orientation, and Supportive Learning Environment on Informal Learning. *Hum. Resource Develop. Q.* 22 (3), 239–257. doi:10.1002/hrdq.20078
- Choudhry, R. M. (2014). Behavior-based Safety on Construction Sites: A Case Study. *Accid. Anal. Prev.* 70, 14–23. doi:10.1016/j.aap.2014.03.007
- Clevenger, C. M., and Ozbek, M. E. (2013). Service-Learning Assessment: Sustainability Competencies in Construction Education. *J. construction Eng. Manag.* 139 (12), 12–21. doi:10.1061/(ASCE)CO.1943-7862.0000769
- Dolnicar, S. (2013). Asking Good Survey Questions. *J. Trav. Res.* 52 (5), 551–574. doi:10.1177/0047287513479842
- Douglas-Lenders, R. C., Holland, P. J., and Allen, B. (2017). Building a Better Workforce. *Et* 59 (1), 2–14. doi:10.1108/et-10-2015-0095
- Eggerth, D. E., Keller, B. M., Cunningham, T. R., and Flynn, M. A. (2018). Evaluation of Toolbox Safety Training in Construction: The Impact of Narratives. *Am. J. Ind. Med.* 61 (12), 997–1004. doi:10.1002/ajim.22919
- Evia, C. (2011). Localizing and Designing Computer-Based Safety Training Solutions for Hispanic Construction Workers. *J. Constr. Eng. Manag.* 137 (6), 452–459. doi:10.1061/(ASCE)CO.1943-7862.0000313
- Forst, L., Ahonen, E., Zanoni, J., Holloway-Beth, A., Oschner, M., Kimmel, L., et al. (2013). More Than Training: Community-Based Participatory Research to Reduce Injuries Among Hispanic Construction Workers. *Am. J. Ind. Med.* 56, 827–837. doi:10.1002/ajim.22187
- Fowler, F. J. (2004). "The Case for More Split-Sample Experiments in Developing Survey Instruments," in *Methods for Testing and Evaluating Survey Questionnaires*. Editor J. M. R. Stanley Presser, C. Mick P., L. Judith T., M. Elizabeth, M. Jean, and S. Eleanor, 173–188.
- Goulding, J., Nadim, W., Petridis, P., and Alshawi, M. (2012). Construction Industry Offsite Production: A Virtual Reality Interactive Training Environment Prototype. *Adv. Eng. Inform.* 26 (1), 103–116. doi:10.1016/j.aei.2011.09.004
- Harvey, S. A., Olórtegui, M. P., Leontsini, E., and Winch, P. J. (2009). "They'll Change what They're Doing if They Know that You're Watching": Measuring Reactivity in Health Behavior Because of an Observer's Presence- A Case from the Peruvian Amazon. *Field Methods* 21 (1), 3–25. doi:10.1177/1525822x08323987
- Hashem M. Mehany, M. S., Killingsworth, J., and Shah, S. (2019). An Evaluation of Training Delivery Methods' Effects on Construction Safety Training and Knowledge Retention - A Foundational Study. *Int. J. Construction Educ. Res.* 17, 18–36. doi:10.1080/15578771.2019.1640319
- Jarkas, A. M. (2010). Critical Investigation into the Applicability of the Learning Curve Theory to Rebar Fixing Labor Productivity. *J. Constr. Eng. Manag.* 136 (12), 1279–1288. doi:10.1061/(ASCE)CO.1943-7862.0000236
- Kaufman, R., and Keller, J. M. (1994). Levels of Evaluation: Beyond Kirkpatrick. *Hum. Resource Develop. Q.* 5 (4), 371–380. doi:10.1002/hrdq.3920050408

- Killingsworth, J., and Grosskopf, K. R. (2013). syNERgy. *Adult Learn.* 24 (3), 95–103. doi:10.1177/1045159513489111
- Kim, S. (2015). An Assessment of Individualized Technical Ear Training for Audio Production. *The J. Acoust. Soc. America* 138 (1), EL110–EL113. doi:10.1121/1.4922622
- Kirkpatrick, D. (1996). Great Ideas Revisited: Revisiting Kirkpatrick's Four-Level Model, Training and Development. *J. Am. Soc. Training Directors* 50, 54–57.
- Kirkpatrick, D. (1959). Techniques for Evaluation Training Programs. *J. Am. Soc. Train. Directors* 13, 21–26.
- Kirkpatrick, J. D., and Kirkpatrick, W. K. (2016). *Kirkpatrick's Four Levels of Training Evaluation*. Alexandria, Virginia, US: Association for Talent Development.
- Lietz, P. (2010). Research into Questionnaire Design: A Summary of the Literature. *Int. J. Market Res.* 52 (2), 249–272. doi:10.2501/s147078530920120x
- Lin, K.-Y., Lee, W., Azari, R., and Migliaccio, G. C. (2018). Training of Low-Literacy and Low-English-Proficiency Hispanic Workers on Construction Fall Fatality. *J. Manage. Eng.* 34 (2), 05017009. doi:10.1061/(ASCE)ME.1943-5479.0000573
- Lingard, H., Pink, S., Harley, J., and Edirisinghe, R. (2015). Looking and Learning: Using Participatory Video to Improve Health and Safety in the Construction Industry. *Construction Manage. Econ.* 33, 740–751. doi:10.1080/01446193.2015.1102301
- Love, P. E. D., Edwards, D. J., Smith, J., and Walker, D. H. T. (2009). Divergence or Congruence? A Path Model of Rework for Building and Civil Engineering Projects. *J. Perform. Constr. Facil.* 23 (6), 480–488. doi:10.1061/(asce)cf.1943-5509.0000054
- Love, P. E. D. (2002). Influence of Project Type and Procurement Method on Rework Costs in Building Construction Projects. *J. Construction Eng. Manage.* 128. doi:10.1061/(ASCE)0733-9364(2002)128:1(18)
- Mills, A., Wingrove, D., and McLaughlin, P. (2010). "Exploring the Development and Assessment of Work-Readiness Using Reflective Practice in Construction Education," in *Proc 26th Annual ARCOM Conference*, Leeds, UK, 6–8 September 2010. Editor C. Egbu (Melbourne, Australia: Association of Researchers in Construction Management), 163–172.
- Mitchell, K. B., Gugerty, L., and Muth, E. (2008). Effects of Brief Training on Use of Automated External Defibrillators by People without Medical Expertise. *Hum. Factors* 50, 301–310. doi:10.1518/001872008x250746
- Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *BMJ* 339, b2535. doi:10.1136/bmj.b2535
- Phillips, J. J. (1991). *Handbook of Evaluation and Measurement Methods*. Houston: Gulf Publishing Company.
- Phillips, J. J. (1998). The Return-On-Investment (ROI) Process: Issues and Trends. *Educ. Technol.* 38 (4), 7–14.
- Reio, T. G., Rocco, T. S., Smith, D. H., and Chang, E. (2017). A Critique of Kirkpatrick's Evaluation Model. *New Horizons Adult Educ. Hum. Resource Develop.* 29 (2), 35–53. doi:10.1002/nha.3.20178
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., et al. (2014). Structural Topic Models for Open-Ended Survey Responses. *Am. J. Polit. Sci.* 58 (4), 1064–1082. doi:10.1111/ajps.12103
- Ruge, G., and McCormack, C. (2017). Building and Construction Students' Skills Development for Employability - Reframing Assessment for Learning in Discipline-specific Contexts. *Architectural Eng. Des. Manage.* 13 (5), 365–383. doi:10.1080/17452007.2017.1328351
- Russell, J. S., Hanna, A., Bank, L. C., and Shapira, A. (2007). Education in Construction Engineering and Management Built on Tradition: Blueprint for Tomorrow. *J. Constr. Eng. Manage.* 133 (9), 661–668. doi:10.1061/(ASCE)0733-9364(2007)133:9(661)
- Salas, E., and Cannon-Bowers, J. A. (2001). The Science of Training: A Decade of Progress. *Annu. Rev. Psychol.* 52, 471–499. doi:10.1146/annurev.psych.52.1.471
- Salsali, M. (2005). Evaluating Teaching Effectiveness in Nursing education: An Iranian Perspective. *BMC Med. Educ.* 5 (1), 1–9. doi:10.1186/1472-6920-5-29
- Sapsford, R. (2006). *Survey Research*. Thousand Oaks, California, US: Sage.
- Sullman, M. J. M., Dorn, L., and Niemi, P. (2015). Eco-driving Training of Professional Bus Drivers - Does it Work? *Transportation Res. C: Emerging Tech.* 58 (Part D), 749–759. doi:10.1016/j.trc.2015.04.010
- Taherdoost, H. (2019). What Is the Best Response Scale for Survey and Questionnaire Design; Review of Different Lengths of Rating Scale/Attitude Scale/Likert Scale. *Res. Methodol. ; Method, Des. Tools* 8, 1–10.
- Wall, J., and Ahmed, V. (2008). Use of a Simulation Game in Delivering Blended Lifelong Learning in the Construction Industry - Opportunities and Challenges. *Comput. Educ.* 50 (4), 1383–1393. doi:10.1016/j.compedu.2006.12.012
- Zhou, C., and Kou, X.-j. (2010). Method of Estimating Human Error Probabilities in Construction for Structural Reliability Analysis Based on Analytic Hierarchy Process and Failure Likelihood Index Method. *J. Shanghai Jiaotong Univ. (Sci.)* 15 (3), 291–296. doi:10.1007/s12204-010-1005-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Jadallah, Friedland, Nahmens, Pecquet, Berryman and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.