

Assessment of Training Effectiveness Adjusted for Learning (ATEAL)

Part II: Practical Application

Thomas Samuel¹, Razia Azen² & Naira Campbell-Kyureghyan^{1,3}

¹ Department of Industrial and Manufacturing Engineering, University of Wisconsin, Milwaukee, Milwaukee, WI, USA

² Department of Educational Psychology, University of Wisconsin, Milwaukee, Milwaukee, WI, USA

³ Department of Mechanical Engineering, Merrimack College, North Andover, MA, USA

Correspondence: Naira Campbell-Kyureghyan, School of Science and Engineering, Merrimack College, North Andover, MA, 01845 USA. E-mail: campbellnk@merrimack.edu

Received: July 8, 2020 Accepted: August 17, 2020 Online Published: September 11, 2020

doi:10.5539/jel.v9n5p134 URL: <https://doi.org/10.5539/jel.v9n5p134>

Abstract

Safety training programs are a popular method, in industry globally, to increase awareness of risks to employees and employers and plays a critical part in reducing safety incidents. The most frequently used method to assess the effectiveness of the training is to have the participants answer Multiple Choice Question (MCQ) and True/False (T/F) questions after the training. The metrics used to report the outcome of the assessments have drawbacks that make it difficult for the trainer and organization to easily identify the concepts that need more focus and those that do not. The goal of this research study is to compare how the methods used to measure training effectiveness of concepts in Level 2 post training assessment differ in how they assess training effectiveness using actual training results. Pre- and Post-training assessments were administered to the participants in 3 different utility industries and were analyzed for training effectiveness using the traditional metrics as well as using ATEAL method. The results were then compared and detailed recommendations of the best and least learned concepts by industry are presented based on these comparative analyses. The ATEAL method is further used to quantify the opportunities for improvement in the training programs based on the participant prior knowledge and any negative training impact observed. Results of the comparison of the various methods show that the proposed ATEAL method provides a quick, accurate and easy way to assesses the effectiveness of the training of concepts and the method identified that for 40% of the concepts trained a higher percentage of participants exhibited more prior knowledge than positive learning and for 6% of the concepts a higher percentage exhibited negative training. These results also provide a directional guide on the improvements that can be made to improve the training effectiveness of the programs. Additionally, it also shows that the ATEAL method can be used in any learning environment where there is a pre-/post-test evaluation of the change and is not limited in application to MCQ and T/F questions.

Keywords: training effectiveness, adult learning, control question, prior knowledge, knowledge gained, concepts trained

1. Introduction

Workplace training, globally, is an important way for organizations to increase the knowledge of their employees and it has been reported that organizations invest approximately \$55.3 billion to \$200 billion annually (Salas & Cannon-Bowers, 2001) on employee training. Brunello and Medio (2001) observed that different countries invest differing amounts in employee training based on tenure, and there is an overall approach globally to increase the knowledge of employees in an organization using formal training methods. With this level of fiscal and time investment being made in training it is important to ensure that the training is effective and will result in the expected changes in behavior among the participants.

Of the various topics that employees are trained on, safety training is particularly important due to the impact of poor safety practices (Campbell-Kyureghyan & Cooper, 2012). According to the Bureau of Labor Statistics, the number of fatal work injuries in the US for 2018 was 5,250, an increase of 2% (5,147) from 2017. Similar statistics have been reported by Ho and Dzung (2010) on occupational disasters in Taiwan and the International

Labor Organization who state that the construction industry has a disproportionately high rate of recorded accidents. This impact to human life and societies worldwide has necessitated a number of legislative acts and organizations being instituted to reduce occupational injuries and mandate workers to undergo safety education through training. This is a sound approach as training is a proven method to improve the safety conditions for workers worldwide with proven reduction in safety incidents on the worksite (Bahn & Barratt-Pugh, 2012; Ho & Dzeng, 2010; Burke et al., 2006; Becker & Morawetz, 2004; Demirkesen & Ardit, 2015; Campbell-Kyureghyan, Hernandez, & Ahmed, 2013). The importance of training is particularly more so in dynamic work environments such as construction which was noted by Campbell-Kyureghyan, Ahmed and Beschoner (2013) as traditional approaches to implement safety protocols with workstation redesigns are ineffective or not practical.

Blume, Ford, Baldwin and Huang (2010) and Tai (2006) noted that effective training can increase the knowledge, skills and abilities (KSA's) of the employees for organizational benefit. In the case of safety training this is particularly important as there is significant human and societal impact to the employee's application of their safety KSAs in the work environment. Alvarez, Salas and Garofano (2004) stated that training experts typically study training effectiveness through evaluation and, although training evaluation and training effectiveness are distinct concepts, they are related and models that integrate both concepts provide a better overall picture. The importance of effective safety training was also stated by Demirkesen and Ardit (2015) who observed that safety improvements may not be achieved unless special attention is paid to the effectiveness of learning during the training session.

The methods used to measure training effectiveness typically involve assessing the overall performance of the participants and no easy methodology exists to help organizations and trainers determine the learning gaps and to determine the best and least learned concepts while compensating for prior knowledge and guessing. Additionally, the current methods do not provide easy directional guidance on the countermeasures that need to be taken to improve the effectiveness of the training for each concept trained. The improvements that can be made to training on concepts related to safety is specifically impactful due to the human and societal benefits that changes in safety behavior have on participants and organizations.

In the companion paper (Part 1), we describe the Assessment of Training Effectiveness Adjusted for Learning (ATEAL) methodology that is able to assess the training effectiveness of each concept taught in a training session by adjusting for negative training impacts and prior knowledge of the participants. This research study presents the results of the different training effectiveness assessment methods of concepts for a pre-/post-test assessment model and determines how the models differ from each other on the concepts they report as best and least learned.

2. Method

2.1 Assessment Metrics

A complete description of the assessment metrics is contained in the companion paper and a brief summary is presented here. To align on nomenclature, the possible outcomes of answers in a pre- and post-test assessment are detailed below in Figure 1.

		Post-Test	
		Correct	Incorrect + IDK
Pre-Test	Correct	CC	CI
	Incorrect + IDK	IC	II

Figure 1. Terminology describing pattern of responses in a pre-/post-test assessment model

Each quadrant in Figure 1 contains the frequency or percentage of respondents that answered in a certain manner and can be interpreted as:

- CC: The question is answered correctly in both pre- and post-tests
- CI: The question is answered correctly in the pre-test and incorrectly or IDK in the post-test

- IC: The question is answered incorrectly or as IDK in the pre-test and correctly in the post-test
- II: The question is answered incorrectly or as IDK in both pre- and post-test assessments

It is important to note that for any question the combination of responses (correct and incorrect) can only lie in one of the four quadrants.

2.1.1 Total Percent Correct (TPC): the TPC measures the number of questions that the participants answered correctly in the post-training assessment or the number of participants who answered a certain question correctly and it is shown below in formula (1).

$$\text{Total Percent Correct (TPC)} = \frac{CC+IC}{CC+IC+CI+II} \quad (1)$$

2.1.2 Post – Pre-Training Percent Correct (PPPC): the PPPC measures the difference between the pre-/post-training scores, and can only be used when the same questions are administered before and after the training. It is computed as shown below in formula (2).

$$\text{Post – Pre-Training Percent Correct (PPPC)} = \frac{CC+IC}{CC+IC+CI+II} - \frac{CC+CI}{CC+IC+CI+II} = \frac{IC-CI}{CC+IC+CI+II} \quad (2)$$

2.1.3 Prior Knowledge (PK): the PK measures the proportion of all participants who answered a question correctly in the post-training assessment who also answered correctly in the pre-training assessment, as is shown in formula (3).

$$\text{Prior Knowledge (PK)} = \frac{CC}{CC+IC} \quad (3)$$

2.1.4 Positive Training Impact (PTI): the PTI, shown in formula (4), measures the proportion of all the participants who needed to learn the concept (responded incorrectly or IDK in the pre-test assessment) who actually did learn the concept as indicated by their response changing to correct in the post-test.

$$\text{Positive Training Impact (PTI)} = \frac{IC}{IC+II} \quad (4)$$

2.1.5 Negative Training Impact (NTI): the NTI, shown in formula (5), measures the proportion of participants who presumably knew the concept prior to training (answered correctly in the pre-training assessment) who answered incorrectly or IDK in the post-test assessment.

$$\text{Negative Training Impact (NTI)} = \frac{CI}{CC+CI} \quad (5)$$

2.1.6 Learning Adjustment Coefficient (LAC): the LAC measures the necessity of the training by comparing the positive impacts of the training (PTI) to the prior knowledge (PK) of the participants, and it is calculated as shown in formula (6).

$$\text{LAC} = \frac{1 + \left(\frac{IC}{IC+II} - \frac{CC}{CC+IC} \right)}{2} \quad (6)$$

2.1.7 Net Training Impact Coefficient (NTIC): the NTIC measures the net impact of the training session by comparing the positive impacts of the training (PTI) to the negative impact of training (NTI) of the respondents, and it is calculated as shown in formula (7).

$$\text{NTIC} = \text{PTI} - \text{NTI} = \frac{IC}{IC+II} - \frac{CI}{CC+CI} \quad (7)$$

2.1.8 Training Effectiveness Matrix (TEM): The LAC and the NTIC can be summarized in a Training Effectiveness Matrix (TEM) that allows for visual identification of the training effectiveness for a concept/question, as shown in Figure 2. The quadrants of the matrix are described below.



Figure 2. Training Effectiveness Matrix with the quadrant layout

Quad 1: Contains questions/concepts for which the participants experienced more positive training impact than either prior knowledge or negative learning impact. The color gradient ranges from yellow to green which indicates increasing levels of positive training impact for the participants.

Quad 2: Contains questions/concepts for which the participants had more prior knowledge than positive training impact but did not experience more negative training than positive training.

Quad 3: Contains the questions/concepts for which the participants had higher negative training impact and it outweighs any positive training impact.

2.2 Industry Application

Workplace safety and ergonomic training was developed and deployed for multiple sectors of the utility industry by a team of researchers at the University of Wisconsin-Milwaukee under a DOL Susan Harwood Training Grant. Table 1 illustrates the number of participants, their roles, number of questions based on types and the usage of Control Question (CQ) and “I Don’t Know” (IDK) option in the three energy sectors. The results from these training sessions will be used to evaluate the performance of the assessment metrics by comparing and contrasting how each of metrics illustrates participant performance for the concepts taught.

Table 1. List of the number of training participants, assessment questions, and usage of CQ and IDK option in each industry

Utility Sector	Participant Role	# of Participants	# of MCQ Assessments	# of T/F Assessments	MCQ Assessments	
					CQ	IDK
Natural Gas	Employee – Tier 1	414	7	8	X	
	Employee – Tier 2	375	7	8	X	
	Manager – Tier 1	86	7	8	X	X
Electric Transmission	Employee – Tier 1	54	9	5		X
	Manager – Tier 1	7	9	5		X
	Employee – Tier 2	359	10	5	X	X
Power Generation	Employee – Tier 2	157	10	5	X	X
	Manager – Tier 1	14	13	9	X	X
Total = 1,466	Managers =	107				
	Employees =	1,359				

Further details of the training methods, the content, and knowledge testing, are detailed in a prior paper written by the same authors (Samuel, Azen, & Campbell-Kyureghyan, 2019). The training sessions were all face-to-face and instructor-led with the number of training participants ranging from 6–40 per class. The pre-test and

post-training assessments contained Multiple Choice Questions (MCQs) and True or False (T/F) items to determine the knowledge of the content for each participant. The pre-training assessment was completed just prior to the training session and collected on completion. The training session typically lasted from 1-3 hours and the same assessment was administered as the post-training assessment. The order of the options and the questions were not rearranged between the pre-test and post-test assessments. The number of MCQ and T/F questions for each of the utility sectors, based on the role of the participants, is summarized in Table 1. In the MCQ assessment, one question, in both the pre- and post-training assessment, was a question contextually similar to the content being trained but was not specifically covered in the training class. This is referred to as the Control Question (CQ) and is used to explain if the participants had prior knowledge of the concept or were guessing in the assessment. Further details of the CQ are provided in Samuel et al., (2019) and Caston, Cooper, and Campbell-Kyureghyan (2009). Additionally, for the pre- and post-training assessments for the Electric Transmission and Power Generation utility sectors an additional “I Don’t Know” (IDK) option was added, as indicated in Table 1.

Training content and concepts were based on research that specifically targeted the areas of safety and ergonomics in non-repetitive work environments (Ahmed & Campbell-Kyureghyan, 2014). To define the ergonomic risks onsite visits were conducted, and data gathered from interviews with managers and employees and direct observations using videotaping methods. Due to the differences in the types of utilities and the work performed concepts were changed to best cater to each industry and combined with information from nationwide industry and fatality statistics for utility industries (Campbell-Kyureghyan & Cooper, 2012). Table 2 details the concepts trained and the number of questions in the assessments by concept for the various training groups in each utility sector. Both employees and mid-level management were trained as it has been reported that management’s commitment to safety results in lowering injury rates and improving the company safety culture (Demirkesen, 2015).

Table 2. Concepts trained and number of assessment questions for each utility industry sector

	Natural Gas		Electric Transmission		Power Generation	
	Employee – Tier 1 & 2	Manager – Tier 1	Employee – Tier 1	Employee – Tier 2	Manager – Tier 1	Employee – Tier 2
Confined Space	0	0	0	0	1	1
Control Question	1	1	0	1	1	1
Electric Safety	0	0	2	2	0	0
Employee Rights & Responsibilities	1	1	1	1	0	1
Environment	2	1	1	1	1	1
General	2	2	1	1	1	2
Hearing Loss	1	1	1	1	2	1
Overexertion	3	1	3	3	5	5
PPE	0	0	2	2	0	0
Program Implementation	0	2	0	0	3	0
Root Cause Analysis	0	1	0	0	1	0
Slips, Trips & Falls	3	3	2	2	2	2
Struck by/caught between	0	0	0	0	1	1
Vehicle Safety	1	1	1	1	0	0
Vibration	1	1	0	0	0	0
Workplace Assessment	0	1	0	0	4	0

The assessment metrics were calculated for each of the training groups and are compared and contrasted to identify the metrics that best help determine the performance of the participants and the direction of training improvements required.

3. Results

The pre- and post-training assessment results for the participants from the various utilities are calculated using the TPC, PPPC, and ATEAL measures to help identify the concepts which were best learned, the concepts for which the participants had the most prior knowledge, and the concepts for which the participants experienced higher negative impact. Additionally, the responses of the participants on the Control Question and its representation by the various metrics is examined. Ideally, in all cases, we would expect the CQs to be at (0.5, 0) in the TEM when using the ATEAL method, and zero when using the PPPC or the TPC as this would indicate

zero learning. However, if there was some prior knowledge on the CQ, we would expect the TPC to be greater than zero and the CQ to lie in either Quad 1 or 2 when using the ATEAL method. The following sections present the results by each utility as the concepts trained varied by the industry.

3.1 Natural Gas Utility

Table 3 illustrates the training performance metrics calculated for the Tier 1 Employee training group in the Natural Gas Utility sector. A total of 405 participants answered each question/concept in this training group. If one used TPC to measure training effectiveness, the conclusion would be that Vehicle Safety, Employee Rights & Responsibilities, and Slips, Trips & Falls are the best learned concepts by this group. However, the PPPC indicates that Hearing Loss is the best learned concept by a large margin.

Table 3. Natural Gas Utility – Tier 1 Employees (n = 405) assessment result metrics

	Concept	TPC	PPPC	LAC	NTIC
1	Control Question	11%	4%	0.41	-0.47
2	Employee Rights & Responsibilities	99%	9%	0.50	0.90
3	Environment	88%	9%	0.44	0.64
4	General	82%	27%	0.56	0.58
5	Hearing Loss	84%	67%	0.83	0.72
6	Overexertion	82%	31%	0.62	0.71
7	Slips, Trips & Falls	99%	22%	0.61	0.99
8	Vehicle Safety	100%	4%	0.52	1.00
9	Vibration	87%	-2%	0.34	0.50

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

In applying the ATEAL method and plotting these 9 concepts on the Training Effectiveness Matrix, as shown in Figure 3, it is clear that Hearing Loss is the best learned concept, and that the prior knowledge level was low. Employee Rights & Responsibilities had about equal number of participants who had prior knowledge as participants who learned the concept. The participants all experienced positive learning for the concepts of General, Hearing Loss, Overexertion, Slips Trips & Falls, and Vehicle Safety.

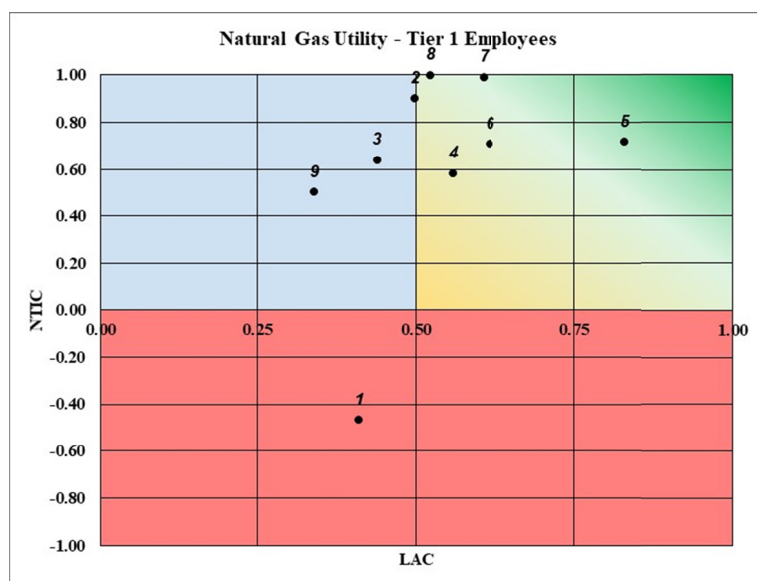


Figure 3. Training Effectiveness Matrix for the Natural Gas Utility – Tier 1 employees

The participants exhibited considerable prior knowledge for the concepts of Environment and Vibration (numbered 3 & 9 in Figure 3). A total of three questions were associated with these two concepts and, since there was such a high amount of prior knowledge among the participants, it would potentially have been a better use of participant time to reduce the number of questions and amount of training on these concepts and instead focus

on the other concepts that needed to be learned. Finally, the ATEAL method does an excellent job in identifying the CQ (numbered 1 in Figure 3) among the concepts taught. As indicated previously, the CQ is a concept that was not taught in the training, but was thematically similar to the rest of the content tested, and was used to estimate the amount of guessing by the participants. The results show that there was more negative training impact than positive training on the CQ and that the participants were having difficulty answering the question. This is the only question for which the NTIC is less than zero. By having the CQ and using it along with the other assessment results, we can clearly see that the ATEAL method helps provide considerably higher resolution in understanding the effectiveness of the training of each concept compared to the PPPC metric.

Table 4 illustrates the training performance metrics calculated for the Tier 2 Employee training group in the Natural Gas Utility sector. A total of 347 participants answered each question/concept in this training group. Similar to the Tier 1 Employee group, the TPC metric does not indicate that Hearing Loss is the best learned concept as it includes the prior knowledge in the final assessment results reported. However, the PPPC identifies Hearing Loss as the best learned concept.

Table 4. Natural Gas Utility – Tier 2 Employees (n = 347) assessment result metrics

	Concept	TPC	PPPC	LAC	NTIC
1	Control Question	42%	27%	0.55	0.10
2	Employee Rights & Responsibilities	96%	8%	0.50	0.85
3	Environment	87%	6%	0.42	0.60
4	General	77%	23%	0.52	0.49
5	Hearing Loss	81%	56%	0.78	0.63
6	Overexertion	78%	21%	0.53	0.58
7	Slips, Trips & Falls	97%	16%	0.53	0.88
8	Vehicle Safety	100%	7%	0.51	0.96
9	Vibration	83%	-7%	0.30	0.41

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

In applying the ATEAL method, the Training Effectiveness Matrix for these 9 concepts, shown in Figure 4, clearly identifies Hearing Loss (numbered 5 in Figure 4) as the best-learned concept. The rest of the concepts have very similar results to those observed with the Tier 1 Employee training group, with the participants having higher prior knowledge for the Environment and Vibration concepts (numbered 3 & 9 in Figure 4). The CQ (numbered 1 in Figure 4) for Tier 2 trainees lands in Quad 1, whereas for the Tier 1 training group it was in Quad 3. This indicates that there was more positive learning on the CQ than both prior knowledge and negative training. However, its magnitude is very low (close to 0.5, 0) indicating that the net learning was almost zero. This could be explained by the fact that more participants in the Tier 2 Employee group guessed correctly on the CQ compared to the Tier 1 Employee group.

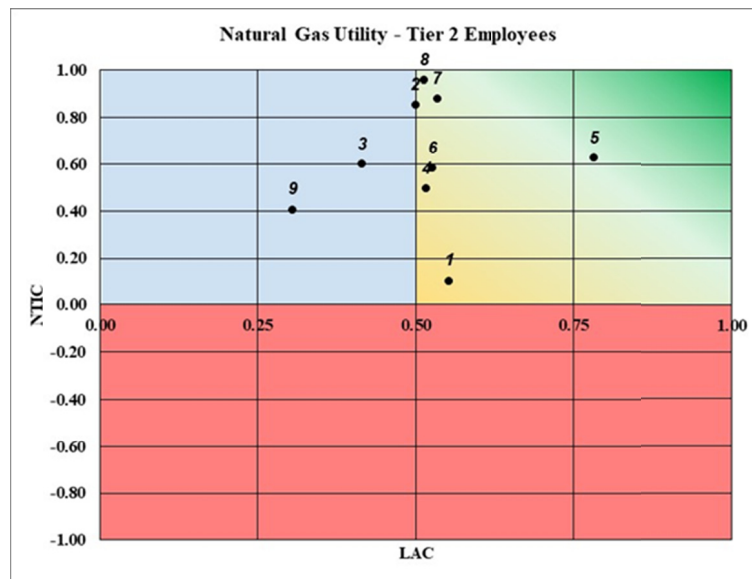


Figure 4. Training Effectiveness Matrix for the Gas Utility – Tier 2 Employees

Table 5 illustrates the training performance metrics calculated for the Manager training group in the Natural Gas Utility sector. A total of 78 participants answered each question/concept in this training group. Similar to the Tier 1 and Tier 2 Employee groups, the TPC metric does not indicate that Hearing Loss is the best learned concept by the training participants as it includes the prior knowledge in the final assessment results reported; however, the PPPC identifies Hearing Loss as the best learned concept.

Table 5. Natural Gas Utility – Manager (n = 78) assessment result metrics

	Concept	TPC	PPPC	LAC	NTIC
1	Control Question	14%	-15%	0.22	-0.62
2	Employee Rights & Responsibilities	96%	9%	0.51	0.87
3	Environment	91%	10%	0.48	0.74
4	General	88%	29%	0.62	0.77
5	Hearing Loss	87%	62%	0.80	0.76
6	Overexertion	78%	50%	0.75	0.56
7	Program Implementation	97%	5%	0.48	0.87
8	Root Cause Analysis	71%	26%	0.54	0.43
9	Slips, Trips & Falls	98%	18%	0.58	0.95
10	Vehicle Safety	100%	9%	0.54	1.00
11	Vibration	92%	1%	0.31	0.53
12	Workplace Assessment	56%	21%	0.58	0.15

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

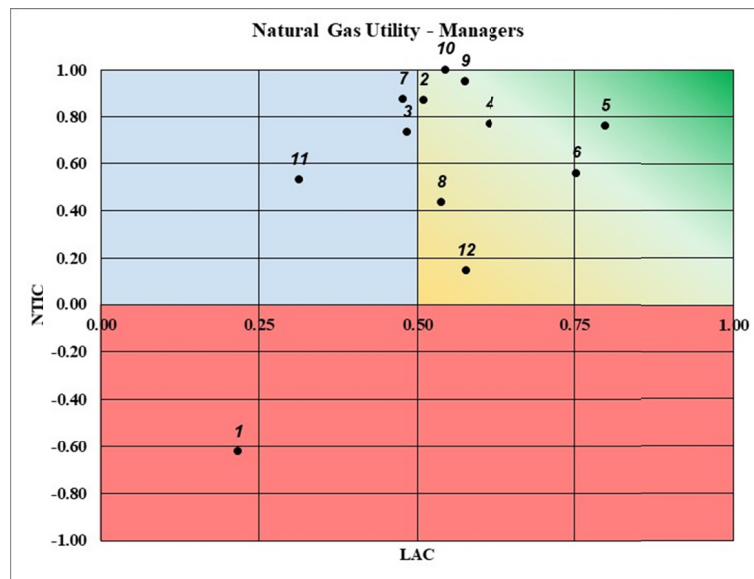


Figure 5. Training Effectiveness Matrix for the Gas Utility – Managers

In applying the ATEAL method, Figure 5 presents the Training Effectiveness Matrix for the 12 concepts for the Manager training group, and clearly shows that the CQ (number 1 in Figure 5) is unlike the other questions. The participants had higher prior knowledge for the concepts of Environment, Program Implementation and Vibration (numbered 3, 7 & 11 respectively in Figure 5), for which there were a total of 4 questions. The Manager training group is observed to exhibit positive learning for the other 8 concepts covered in the training program.

3.2 Electric Transmission Utility

Table 6 illustrates the training performance metrics calculated for the Tier 1 Employee and Manager training group in the Electric Transmission Utility sector, as both groups were administered the same MCQ & T/F assessments. A total of 60 participants answered each question/concept in this training group. If TPC was used to measure training effectiveness, we would have concluded that Vehicle Safety, Slips, Trips & Falls and PPE are the best learned concepts by this group. However, PPPC indicates that General is the best learned concept by a large margin.

Table 6. Electric Transmission Utility – Tier 1 Employees (n = 60) assessment result metrics.

	Concept	TPC	PPPC	LAC	NTIC
1	Electric Safety	93%	21%	0.53	0.78
2	Employee Rights & Responsibilities	95%	13%	0.49	0.80
3	Environment	88%	12%	0.50	0.70
4	General	63%	43%	0.73	0.23
5	Hearing Loss	87%	27%	0.52	0.68
6	Overexertion	79%	13%	0.48	0.56
7	PPE	96%	29%	0.40	0.43
8	Slips, Trips & Falls	96%	15%	0.60	0.95
9	Vehicle Safety	97%	8%	0.56	0.96

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

In applying the ATEAL method and plotting the Training Effectiveness Matrix as shown in Figure 6, it is clear that this group had more learning than prior knowledge for the General (numbered 4 in Figure 6) concept; however, they did experience more negative training for this concept than the concepts of Slips Trips & Falls and Vehicle Safety (numbered 8 & 9 in Figure 6). The participants had considerably higher prior knowledge for the concepts of Employee Rights & Responsibilities, Environment, Over Exertion and PPE (numbered 2, 6 & 7 respectively in Figure 6). Thus, despite having PPPC scores of 13%, 12%, 13% and 29% respectively, they still

lie in Quad 2, thus indicating that there was a lower need to train on these concepts. There was no CQ for this training group.

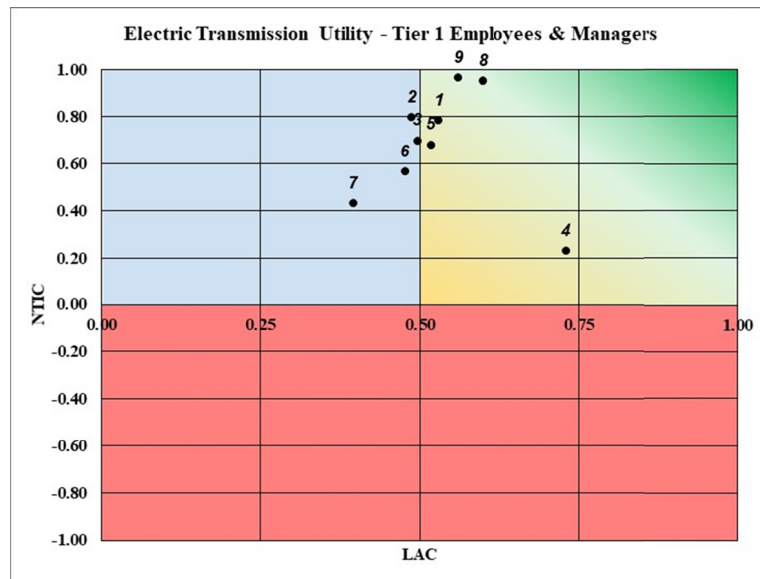


Figure 6. Training Effectiveness Matrix for the Electric Transmission Utility – Tier 1 Employees & Managers

Table 7 illustrates the training performance metrics calculated for the Tier 2 Employee training group in the Electric Transmission Utility. If we used the TPC to measure training effectiveness, we would have concluded that Vehicle Safety, Hearing Loss and Slips, Trips & Falls are the best learned concept by this group. However, PPPC would have us conclude that the General and Control Question concepts were the best learned by a large margin.

Table 7. Electric Transmission Utility – Tier 2 Employees assessment result metrics.

	Concept*	TPC	PPPC	LAC	NTIC
1	Control Question (n=293)	45%	33%	0.60	0.19
2	Electric Safety (n=292)	88%	22%	0.54	0.69
3	Employee Rights & Responsibilities (n = 217)	88%	0%	0.32	0.48
4	Environment (n=217)	87%	2%	0.40	0.58
5	General (n=217)	68%	38%	0.59	0.49
6	Hearing Loss (n=287)	93%	21%	0.56	0.81
7	Overexertion (n=293)	87%	7%	0.42	0.62
8	PPE (n=217)	77%	5%	0.31	0.36
9	Slips, Trips & Falls (n=293)	93%	11%	0.43	0.70
10	Vehicle Safety (n=293)	94%	-1%	0.24	0.60

Note. *Where 'n' is the number of participants answering questions on that specific concept; TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

The PPPC indicates that there is negative learning for the concept of Vehicle Safety. This, however, is different from the information we observe when using the ATEAL method. In reviewing the Training Effectiveness Matrix for this training group, shown in Figure 7, we observe that none of the concepts exhibited negative learning. The participants have higher prior knowledge on the concepts of Employee Rights & Responsibilities, Environment, Overexertion, PPE, Slips Trips & Falls and Vehicle Safety (numbered 3, 4, 7, 8, 9 & 10 respectively in Figure 7). The trainees exhibit positive learning for the concepts of Electric Safety, Hearing Loss and General (numbered 2, 6 & 5 respectively in Figure 7). Of these concepts, the General and Hearing Loss concepts (numbered 5 & 6 in Figure 7) are the best learned concepts. The CQ also lies in Quad 1 for this training group, potentially due to correct guessing by the participants.

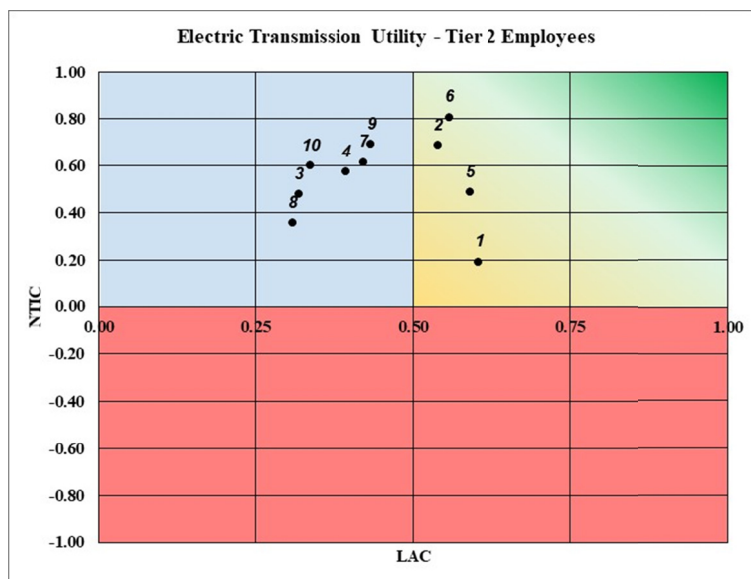


Figure 7. Training Effectiveness Matrix for the Electric Transmission Utility – Tier 2 Employees

3.3 Power Generation Utility

Table 8 illustrates the training performance metrics calculated for the Managers training group in the Power Generation Utility. A total of 12 participants answered each question/concept in this training group. If the TPC was used to measure training effectiveness, we would have concluded that Confined Space and Environment are the best learned concept by this group. However, PPPC would have us conclude that Confined Space, Hearing Loss and Struck by/caught between are the best learned concepts. Additionally, the PPPC illustrates that the Managers had a positive learning experience for the CQ.

Table 8. Power Generation Utility – Managers (n = 12) assessment result metrics

	Concept	TPC	PPPC	LAC	NTIC
1	Confined Space	100%	33%	0.67	1.00
2	Control Question	58%	58%	0.79	0.58
3	Environment	100%	8%	0.54	1.00
4	General	92%	17%	0.42	0.67
5	Hearing Loss	92%	33%	0.63	0.88
6	Overexertion	94%	13%	0.60	0.92
7	Program Implementation	97%	17%	0.55	0.92
8	Root Cause Analysis	25%	17%	0.64	-0.73
9	Slips, Trips & Falls	92%	13%	0.51	0.78
10	Struck by/caught between	42%	33%	0.58	0.36
11	Workplace Assessment	27%	27%	0.64	0.27

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

Using the ATEAL method, the Training Effectiveness Matrix for the Power Generation Utility Managers is shown in Figure 8. It shows that more of participants have prior knowledge for the General concept (numbered 4 in Figure 8) than those that learned the concept. The matrix shows that the participants experienced significant negative learning for the Root Cause Analysis concept (numbered 8 in Figure 8), a detail that could not be discerned by looking at the TPC or the PPPC metrics. The matrix also shows that, other than the two concepts detailed above, the participants experienced positive training for all the other concepts including the CQ.

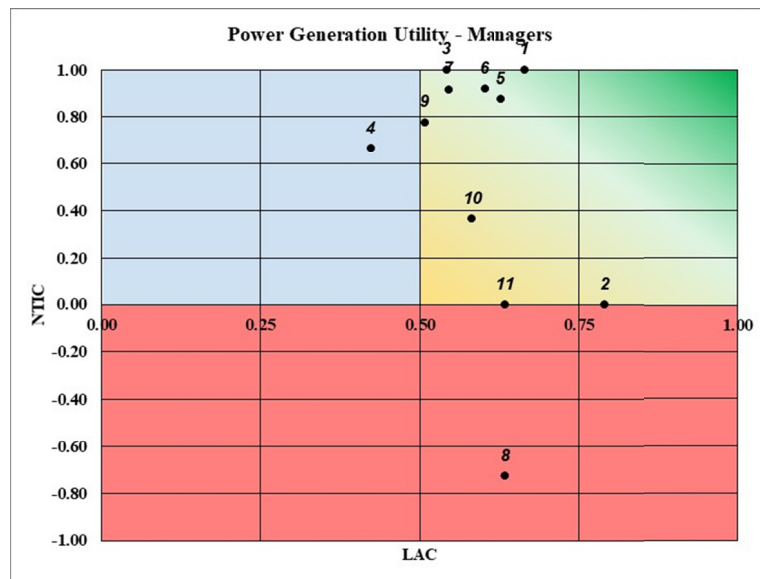


Figure 8. Training Effectiveness Matrix for the Power Generation Utility – Managers

Table 9 illustrates the training performance metrics calculated for the Employees training group in the Power Generation Utility. A total of 176 participants answered each question/concept in this training group. If we use the TPC to measure training effectiveness, we would conclude that Environment and Slips Trips & Falls are the best learned concept by this group. However, PPPC would have us conclude that Confined Space and Struck by/caught Between are the best learned concepts.

Table 9. Power Generation Utility – Employees (n = 176) assessment result metrics.

	Concept	TPC	PPPC	LAC	NTIC
1	Confined Space	73%	22%	0.44	0.46
2	Control Question	30%	10%	0.36	-0.09
3	Employee Rights & Responsibilities	90%	9%	0.38	0.59
4	Environment	94%	7%	0.43	0.72
5	General	85%	13%	0.40	0.55
6	Hearing Loss	54%	15%	0.35	0.21
7	Overexertion	91%	14%	0.47	0.70
8	Slips, Trips & Falls	93%	12%	0.47	0.75
9	Struck by/caught between	76%	21%	0.45	0.48

Note. TPC – Total Percent Correct; LAC – Learning Adjustment Coefficient; PPPC – Post – Pre-Training Percent Correct; NTIC – Net Training Impact Coefficient.

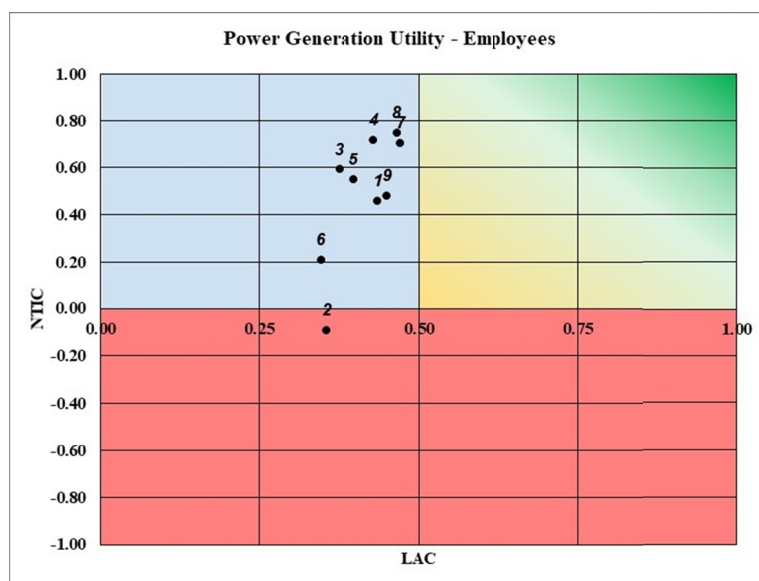


Figure 9. Training Effectiveness Matrix for the Power Generation Utility – Employees

Using the ATEAL method and the Training Effectiveness Matrix shown in Figure 9 for the Power Generation Utility – Employees training group, we observe that, despite the results reported by the TPC and the PPPC, the larger number of trainees had prior knowledge on all of the training concepts. The method also clearly isolates the CQ (numbered 2 in Figure 9) by illustrating that it lies in Quad 3 and that the participants were having a difficult time answering it. This implies that considerable improvement is needed on the concepts trained for the training to be useful to the participants and the organization.

4. Discussion

The results from the different training groups using the various assessment metrics demonstrated that the ATEAL method is easy to interpret and is helpful in determining the steps that need to be taken to improve the training. For consistency and flow, the discussion is presented by utility as was the case in the Results section.

4.1 Natural Gas Utility

For the Natural Gas Utility Tier 1 Employees, the ATEAL method identifies 'Hearing Loss' as the best learned concept followed by the concepts of Slips Trips & Falls, Overexertion, General and Vehicle Safety. The participants exhibited higher prior knowledge than positive training for the concepts of Environment and Vibration. This implies that the trainer could either reduce the time spent on training these concepts to the participants, or train content within the concepts that would be more value added to the participants in terms of gaining new knowledge. It is very interesting to note that the ATEAL matrix clearly identifies the CQ as a question the participants had trouble answering and places it in Quad 3. Hence, we can easily conclude that there was more negative learning than positive learning for the CQ. Since this question was not taught in the training session, it is appropriate and correct that the matrix separates it from the rest of the questions taught. Similar trends are observed in the Natural Gas Utility Tier 2 Employees, where Hearing Loss is the best learned concept and the participants had very high prior knowledge for the concepts of Environment and Vibration. In the case of the CQ, however, the matrix shows that there was almost zero learning as it is very close to (0.5, 0). This is the expected outcome in the ATEAL method as the CQ concept was not taught during the training and we would expect no positive or negative learning. Finally, the Natural Gas Utility Managers exhibit the same trends as the Tier 1 Employees. That is, Hearing Loss is the best learned concept and the CQ is located in Quad 3, clearly separated from the rest of the questions.

In the companion paper, we observe that the TPC metric is overly optimistic in its depiction of participant performance in the case of the scenario and simulation analysis. These trends are again seen when analyzing actual training data and the implications are more profound. In the Natural Gas Utility, TPC shows Vehicle Safety to be the best learned concept for the Tier 1 Employees, Tier 2 Employees and Managers. This, however, is due to the participants having very high prior knowledge (over 95%) of this concept; that is, they were able to answer it correctly in both the pre- and post-test assessments. Looking strictly at the TPC metric, the trainer would have interpreted that 'Vehicle Safety' was the best learned concept among the 875 participants in the

Natural Gas Utility and that its content and method of delivery was highly effective due to its positive impact with such a large number of participants. However, because of the high level of prior knowledge, there should have only been a cursory overview of this concept and an argument can be made that it did not need to be tested in the post-test assessment.

The results of the PPPC in the scenario and simulation analysis in the companion paper show that it is better at compensating for prior knowledge than TPC. This benefit is further observed when looking at the results of actual training and assessments conducted on the participants from the various utility companies. For the Tier 1 Employees in the Natural Gas Utility, the PPPC identifies the concept of 'Hearing Loss' to be the best learned concept and 'Vibration' to be the worst learned concept. In looking at the actual performance of the participants for 'Vibration' (CC = 80%; IC = 7%; CI = 9%) we observe that its negative PPPC value is due to the high prior knowledge among the participants and the small number of participants who experienced negative learning. The PPPC metric also does not isolate the CQ and, although it reports a low performance of the participants for the CQ, it is in line with the results for 'Vehicle Safety' which had a low score due to very high prior knowledge. The same trends for the concepts are observed for Natural Gas Utility Tier 2 Employees. For the Managers, the concept of 'Hearing Loss' is identified as the best learned concept and the CQ receives a negative score.

4.2 Electric Transmission Utility

Using the ATEAL method to analyze the data for the Electric Transmission Utility, for Tier 1 Employees there was positive learning on six of the nine concepts taught, with the 'General' concept being the best learned because the participants had the least prior knowledge and comparatively learned the most on this concept. For the Electric Transmission Utility Tier 2 Employees, the concept of 'Hearing Loss' was the best learned concept and the participants exhibited positive learning on four of the ten concepts tested. The CQ, as seen before, exhibited low learning and, although it is in Quad 1, it is the closest of all the concepts taught to (0.5, 0). When using the TPC to analyze the data in the Electric Transmission Utility, the concept of 'Vehicle Safety' again seems to be the best learned concept by the Tier 1 & 2 Employees due to the high level of prior knowledge (over 85%) among the participants. In using the PPPC to analyze the data of the Tier 1 Employees, the 'General' concept is identified as the best learned concept and 'Vehicle Safety' as the least learned concept. This is the exact opposite of the results from the TPC metric, and is a more accurate representation of participant knowledge levels as the participants had the highest amount of prior knowledge for 'Vehicle Safety'. Similarly, for the Tier 2 Employees in the Electric Transmission Utility, the 'General' concept is identified as the best learned concept. Due to high prior knowledge and a small number of participants experiencing negative learning, the metric identifies 'Vehicle Safety' and 'Employee Rights & Responsibilities' as the worst learned concept.

4.3 Power Generation Utility

Using the ATEAL method, we observe that there was positive learning on eight of the eleven concepts on which Managers were tested. It is extremely interesting to observe that the CQ was the best learned concept, as over 50% of the participants went from incorrect and IDK responses to the CQ in the pre-test assessment to correct responses in the post-test assessment. This could imply that the concept was inadvertently trained in the class by the trainer or that the participants were able to correctly guess the post-test answer. We observe that the concept of Root Cause Analysis had considerable negative training impact and very low prior knowledge. This is a critical issue as this concept is key for the Managers to diagnose safety issues correctly and implement countermeasures to improve the safety of the employees. In further researching the results, we observe that 58% of the participants exhibited zero learning; thus, it is important for the trainers to revisit this concept with this group to ensure that they understand and learn the concepts. It is not possible to quickly arrive at this conclusion when solely looking at TPC and PPPC metrics. Hence, this shows that using the ATEAL method is better and quicker at helping discern participant learning and helps trainers determine countermeasures in an expeditious manner. For the Power Generation Utility Employees, we observe that the CQ lies in Quad 3 and we observe that, for all the other concepts taught, the participants exhibited considerably higher prior knowledge than learning. This is concerning as it shows that a majority of the participants did not learn anything new and the effective use of their time comes into question.

Using the TPC the concept of 'Environment' is shown to be the best learned concept for both the Employees and the Managers due to high prior knowledge (over 84%). The Managers of the Power Generation Utility are also shown to have high learning for the concept of 'Confined Space' as reported by this metric. For this concept there was considerably less prior knowledge (66%) and 33% of the Managers learned the concept. In using the PPPC to analyze the results for the Managers in the Power Generation Utility, we observe that the CQ is reported as the best learned concept. Although this is counterintuitive, the results are due to the 0% prior knowledge and

50% of the participants who answered correctly in the post-test assessment. The other concepts ranked lower mainly due to the fact that participants had higher prior knowledge. Finally, for the Employees of the Power General Utility, the concept of 'Confined Space' is reported to be the best learned concept, although 48% of the participants had prior knowledge of this concept, and 'Environment' is the least learned concept due to 84% of the participants having prior knowledge of this concept.

A common observation through the results and discussion across all of the utilities is the level of prior knowledge that the participants possess for the various concepts trained. Using the ATEAL method we can clearly determine when there are more participants exhibiting prior knowledge than learning. This is impossible to determine when using the TPC metric as it does not compensate for prior knowledge and reports it as learning. Using the PPPC, it takes more time to discern if the low (or) negative values are due to high prior knowledge or negative learning. The metric does not separate the elements, so it requires additional detailed review of the raw score that takes time and effort and may not always be conducted.

The limitation of ATEAL is that the method requires the presence of matched pre- and post-training assessment results, as the analysis is based on baseline knowledge and learning and cannot be used when there are only post-training assessment results. The application of the method may also require some basic training for trainers and organizations. This training, however, is minimal, as the calculations are simple and graphics are easily implemented by using widely available software packages such as MS Excel.

One of the generalizable benefits of the ATEAL method is that it can be used for any type of assessment situation where there is a pre- and post-test assessment. For example, suppose assembly workers were being trained to improve assembly practices, and the assessment was made by an assessor observing the assembler for performance in the categories of quality, speed, efficiency, following standard work, etc. If the assessment is made on the assembler prior to training, and a score obtained for the various categories, the training conducted and the assembler can then be reassessed on their performance post training and the ATEAL method can be used to measure the training effectiveness in this scenario. Thus, the method is more widely applicable than in just the case of MCQ assessments. This may have remarkable implications for the organizations and the participants as the training time can be reduced and the effectiveness improved simultaneously. Additionally, reduction in training time may have fiscal impacts that result in a higher return on investment (ROI) for the training with a higher focus on concepts for which the participants genuinely have knowledge gaps.

5. Conclusion/Future Direction

Metrics to quantify the amount of learning that training participants exhibit for a particular training course, or concepts within the course, are critical to understanding the effectiveness of the training, specifically in the context of workplace safety-related concepts. Using the ATEAL method to measure training effectiveness for training conducted with 1,466 participants from a variety of utility industries, and comparing the results to traditional measurement metrics, we observe that the ATEAL method proves very effective in quickly identifying the learning gaps that the participants experienced and in giving direction on the countermeasures that should be taken for each concept trained.

Some recommendations that can be derived from this study are:

- Using only the TPC in the post-test assessment to evaluate training effectiveness (or) how much the participants learned is shown to be a highly inaccurate method and does not give clear guidance on areas of improvement.
- The PPPC is shown to be a better metric than the TPC to evaluate training effectiveness; however, it lacks the ability to quickly provide information on the changes needed in the training content or its delivery to improve training effectiveness.
- The ATEAL method uses metrics that are of greater accuracy, are easy to calculate, and provide intuitive output that allows for easy visualization of the training effectiveness results. It provides a great way to illustrate the training effectiveness of each concept taught to the participants and can be used to quickly determine the countermeasures that need to be taken by the trainer with regards to content delivery or development as part of the training program. This then provides information on how to improve training effectiveness in future training sessions on the topic. Organizations can also benefit considerably from this method as it helps them understand the concepts that the participants can be held accountable for as well as the specific concepts that need further reinforcement to ensure the employees have safe work practices in their work environment.
- Using the ATEAL method, the trainers and the organizations are able to quickly identify the concepts for

which the participants had considerable prior knowledge. This enables them to focus on concepts for which the participants truly have knowledge gaps and ensure the best return of investment on the training provided and the time used for the training.

Acknowledgments

This study was partially funded by the US DOL Susan Harwood Grants: SH-20840-SH0, #SH-22220-SH1, #SH-23568-SH2, #SH-24880-SH3. The authors also express their gratitude to Karen Cooper, Sruthi Boda and Madiha Ahmed for assisting with the test development and administration. We would additionally like to thank all the companies and employees who participated in this study.

References

- Ahmed, M., & Campbell-Kyureghyan, N. (2014, June). *Reliability of learning assessment*. Proceedings of the 26th Annual International Occupational Ergonomics & Safety Conference, El Paso, TX. June 5–6, 2014
- Alvarez, K., Salas, E., & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review*, 3(4), 385–416. <https://doi.org/10.1177/1534484304270820>
- Bahn, S., & Barratt-Pugh, L. (2012). Emerging issues of health and safety training delivery in Australia: Quality and transferability. *Procedia-Social and Behavioral Sciences*, 62, 213–222. <https://doi.org/10.1016/j.sbspro.2012.09.035>
- Bar-Hillel, M., Budescu, D., & Attali, Y. (2005). Scoring and keying multiple choice tests: A case study in irrationality. *Mind & Society*, 4(1), 3–12. <https://doi.org/10.1007/s11299-005-0001-z>
- Becker, P., & Morawetz, J. (2004). Impacts of health and safety education: Comparison of worker activities before and after training. *American Journal of Industrial Medicine*, 46(1), 63–70. <https://doi.org/10.1002/ajim.20034>
- Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, 36(4), 1065–1105. <https://doi.org/10.1177/0149206309352880>
- Brunello, G., & Medio, A. (2001). An explanation of international differences in education and workplace training. *European Economic Review*, 45(2), 307–322. [https://doi.org/10.1016/S0014-2921\(99\)00065-3](https://doi.org/10.1016/S0014-2921(99)00065-3)
- Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American Journal of Public Health*, 96(2), 315–324. <https://doi.org/10.2105/AJPH.2004.059840>
- Campbell-Kyureghyan, N., Ahmed, M., & Beschoner, K. (2013, March). *Measuring training impact 1–5*. Paper presented at the US DOL Trainer Exchange Meeting, Washington DC, March 12–13, 2013.
- Campbell-Kyureghyan, N., & Cooper, K. (2012). Impact of customized training on learning across demographic groups. *JPIIE*, 9(1), 25–31.
- Campbell-Kyureghyan, N., Hernandez, A. P., & Ahmed, M. (2013). *Effectiveness of first and second tier safety and ergonomics training in power utilities*. Proceedings of the 25th Annual Occupational Ergonomics and Safety Conference, Atlanta, GA, USA.
- Caston, S., Cooper, K., & Campbell-Kyureghyan, N. (2009). *Assessment of ergonomic training in small business foundries*. In Proc. 21st Annual International Occupational Ergonomics and Safety Conference (Vol. 9).
- Demirkesen, S., & Arditi, D. (2015). Construction safety personnel's perceptions of safety training practices. *International Journal of Project Management*, 33(5), 1160–1169. <https://doi.org/10.1016/j.ijproman.2015.01.007>
- Ho, C. L., & Dzeng, R. J. (2010). Construction safety training via e-Learning: Learning effectiveness and user satisfaction. *Computers & Education*, 55(2), 858–867. <https://doi.org/10.1016/j.compedu.2010.03.017>
- International Labor Organization. (n.d.). *World Statistic*. Retrieved from https://www.ilo.org/moscow/areas-of-work/occupational-safety-and-health/WCMS_249278/lang--en/index.htm
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52(1), 471–499. <https://doi.org/10.1146/annurev.psych.52.1.471>
- Samuel, T., Azen, R., & Campbell-Kyureghyan, N. (2019). Evaluation of learning outcomes through multiple choice pre-and post-training assessments. *Journal of Education and Learning*, 8(3), 122. <https://doi.org/10.5539/jel.v8n3p122>

Tai, W. T. (2006). Effects of training framing, general self-efficacy and training motivation on trainees' training effectiveness. *Personnel Review*, 35(1), 51–65. <https://doi.org/10.1108/00483480610636786>
<https://www.bls.gov/news.release/pdf/cfoi.pdf>

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).