

Developing General Training Effectiveness Scale for the Malaysian Workplace Learning

Siti Fardaniah Abdul Aziz

PhD, Human Resource Development, Senior Lecturer, School of Psychology and Human Development,
Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia
Email: daniah@ukm.edu.my

Doi:10.5901/mjss.2015.v6n4s1p47

Abstract

Information on evaluating training effectiveness via a general instrument is limited despite its obvious need in human resource development (HRD) especially in the Malaysian context. To this end, the proposed General Training Effectiveness Scale (GTES) offers a useful measurement for HRD practitioners to evaluate the self-report of workplace learning at minimal cost, time, and energy. The GTES was developed using theories in HRD, and tested in four studies including jury validation, pilot study, survey study¹, and extension study with the Malaysian sample. The first version of GTES comprises of 30 items; however, only 10 items left in the final version. The proposed GTES constitutes a general instrument to measure training effectiveness at various levels including learning performance, individual performance, and organizational performance using a self-report that can be used by researchers, practitioners, and public.

Keywords: training evaluation, psychometric property, employee training, human resource development, General Training Effectiveness Scale, Malaysia

1. Introduction

Organizing employee training has grown into an international trend especially for those organizations listed in stock exchange board and multinational companies (Bersin, 2008; Griffin, 2012). Powell (2009) reported that U.S.A. spends \$134.39 billion on such programs, while Griffin (2010) reported that U.K. spends £38.6 billion annually. In developing country, such as Malaysia, training is also a major concern, in which employers must contribute as much as one per cent annually from their employees' salary for training (Maimunah and Aahad, 2013). Given this willingness for employee training, measuring its effectiveness becomes a necessary component in improving productivity as well as to develop the human capital (Schonewille, 2001; Ramos et al., 2004). To this end a handful of models and approaches to evaluate its effectiveness have been suggested (e.g., Kirkpatrick, 1996; Brinkerhoff, 2006; Pineda- Herrero et al., 2011).

Among the approaches to evaluate training effectiveness, the four levels of Kirkpatrick's (1959/1996) model has been continually used as the favourite model for researchers and practitioners because it is simple and easy to understand (Noe, 2010; Giangreco et al., 2010). According to Kirkpatrick (1959/1996), training effectiveness can be evaluated using the four levels: reaction, learning, behavioural changes, and results. However, Kirkpatrick argues that the higher the level of training effectiveness the more difficult, complex, and expensive the measure. This explains why it is very difficult to evaluate all levels of training effectiveness.

Some researchers argue that to conclude the effectiveness of training using selected levels of training evaluation based on the Kirkpatrick's model offers an incomplete and partial measure (e.g., Brinkerhoff, 2006; Noe, 2010; Ford, 2014). Ironically, Giangreco et al. (2009) found that only large organizations tend to evaluate training effectiveness at all levels, most of small to medium organizations have never evaluate beyond the second level although the training evaluation seems to be incomplete. In fact, Chong (2005) found that from 106 manufacturing companies in Malaysia, 35% evaluated the reaction, 25% evaluated learning changes, 16.5% evaluated the behavioural changes, 11% evaluated results, and 12.5% did not organize any training evaluation. Hence, there is a need to develop and test a self-report instrument to conclude general training effectiveness.

In view of the above, the purpose of this article is to report the development of a general instrument to measure a self-report on training effectiveness named the General Training Effectiveness Scale (GTES). Rationally, the GTES can

¹ Development of the proposed GTES was part of a PhD research by the author during her candidature at University Putra Malaysia with the title, "Mediation effect of training motivation on the relationships between trainee, training, and organizational characteristic support and training effectiveness".

be used as a valid and reliable tool to determine general training effectiveness in order to make early decisions whether to terminate, continue, or improve formal employee training. The findings are useful for HRD practitioners to prepare records for training audit in addition to supporting the decision making process of top management regarding future formal training for workplace learning especially in the Malaysian context.

2. Literature Review

Cannon-Bowers *et al.* (1995) and Holton (2005) believe that training effectiveness can be determined by assessing the learning, individual, and organizational performance but not the trainees' reaction as proposed by Kirkpatrick (1959/1996). Hence, in the current article, the construction of GTES uses the HRD theory/model proposed by Cannon-Bowers *et al.* (1995). Meanwhile, Bersin (2008) redefines these measurements into learning performance, individual performance, and organizational performance as a better term to evaluate training effectiveness. This is because he argues that those terms, such as "individual performance" is a more appropriate term than "behavioural changes" since it distinguishes the effect of training transfer on job performance instead of workplace behaviour.

Cannon-Bowers *et al.* (1995) proposed a complex model of training evaluation by explaining the process of developing training from training needs analysis to evaluating the training effectiveness. They argue that trainees' reaction could have effect on training effectiveness but should not be considered as a level in assessing training effectiveness. Therefore, training can be evaluated using learning, individual, and organizational performance. This is supported by Alliger's *et al.* (1997) study that finds trainees' reaction to have a low or insignificant correlation with other measurement including learning performance. In fact, Holton (2005) argue that *reaction* does not relate to the achievement of the training's objectives, hence, it should not be considered as training evaluation.

Further, some scholars prefer to evaluate training effectiveness using selective levels (e.g., Kraiger *et al.*, 1993; Kontoghiorghes, 2004) and some proposed multiple constructs of training evaluation to conclude training effectiveness (e.g., Cervero, 1988; Quinones, 1997). This is because prior researchers argue that the purpose of training will not always involve the complete levels though some argued the opposite. For example, some training is organized to provide employees with knowledge that can only be used in the future; hence, it is not expected to be transferred to the job setting. This explains why Baldwin *et al.* (1991) focus only on learning performance.

Nonetheless, Garcia (2005) combines the evaluation of *reaction*, *knowledge acquired* (learning performance), *knowledge applications* (individual performance), and *costs and benefits* (organizational performance) simultaneously to determine the status of training effectiveness; meanwhile, Horgan and Muhlau (2006) combine *skills acquired* (learning performance) and *job performance* (individual performance) to measure training effectiveness. This implies that training effectiveness can also be determined by combining training evaluation at all levels; this approach is consistent with that suggested by Cervero (1988) and Quinones (1997).

Interestingly, prior researchers found that the learning, individual performance, and organizational performance that usually used to determine training effectiveness were significantly correlated (Tracey *et al.*, 2001; Pilati and Borges-Andrade, 2008; Chiaburu and Tekleab, 2005; Scaduto *et al.*, 2008). However, the strength of relationship between these constructs varied from low to strong correlation. For example, Bell and Ford (2007) found $r = .20$ for the relationship between knowledge-acquired self-efficacy (learning) and individual performance, but Tziner *et al.* (2007) found $r = .705$ for the relationship between declarative knowledge (learning) and individual performance. This goes the same with the relationship between other constructs. Hence, it can be seen that training effectiveness can be evaluated by assessing all the learning, individual, and organizational performance.

2.1 Learning Performance

Learning performance is the improvement or changes in declarative knowledge, procedural knowledge, and meta-cognition (Kraiger *et al.*, 1993; Colquitt *et al.*, 2000). Bersin (2008) stresses that learning performance is used to measure the achievement of training objectives, especially to improve the capability and qualification of an individual to perform in his or her job. Kraiger *et al.* (1993) and Kirkpatrick (1959/1996) argue that learning performance is very important because it can determine whether training can educate trainees and provide them with knowledge, skills, and attitude. Hence, almost all models of training effectiveness have embraced learning performance as a measurement in training evaluation including those suggested by Kirkpatrick (1959/1996), Holton (2005), and Griffin (2010). Learning performance is usually measured using academic tests (e.g., Bell and Ford, 2007; Tziner *et al.*, 2007); however, some researchers used trainees' perception in self-report (e.g., Stanford, 2000; Chen and Chih, 2012). This is because some researchers, such as Stehle *et al.* (2012), have proven that there is no significant difference between learning performance measured

by learning test and learning perception if both are measuring the same criterion of learning to achieve the same learning objective.

2.2 Individual Performance

Individual performance is a measurement to determine whether what is learned from training is transferred to the workplace (Chiaburu *et al.*, 2010) and can be evaluated using individual improvements in job competencies, efficiencies, and effectiveness (Kirkpatrick, 1959/1996). Baldwin and Ford (1988) argue that the measurement is very important because it is the main reason trainees are sent for training; meanwhile, Bersin (2008) stresses the importance of exhibiting the training improvements in job performance. Hence, contemporary models of training effectiveness have identified it as the main measurement in training evaluation including those suggested by Kontoghiorghes (2004), Bersin (2008), and Kirkpatrick and Kirkpatrick (2010). Individual performance is usually measured using either trainees' perception in self-report (e.g., Facticeau *et al.*, 1995; Nikandrou *et al.*, 2009) or their supervisor's perception in self-report (e.g., Axtell *et al.*, 1997; Tziner *et al.*, 2007). However, it is rarely measured using trainees' job performance record because of its political affect; in which, Garavan *et al.* (1997) argue that general record of job performance for promotion is usually rely on supervisor's personal interest instead of actual performance. Interestingly, most HRD researchers including Axtell *et al.* (1997), Nijman *et al.* (2006), and Tziner *et al.* (2007) find that there is no significant difference between individual performance as reported by trainees and their supervisors in training records.

2.3 Organizational Performance

Organizational performance is the improvement or changes in productivity, teamwork, customer satisfaction, and achievement of an organization's goals and reputation due to training outcomes (Kirkpatrick, 1959/1996). Griffin (2012) argues that organizational performance is used to determine the impact and profit from training on organizational effectiveness. Hence, some training evaluation can be used to determine organizational performance including the ROI (return on investment) technique that popularized by Phillips (1997), and the tangible and intangible result technique by Kirkpatrick and Kirkpatrick (2010). Further, Brinkerhoff (2006) suggests that organizational performance can be evaluated using the effect of training on training function, managers, systems, and senior leaders. Ironically, Chong (2005) found that only 11% from 106 manufacturing companies in Malaysia evaluated *results* or organizational performance; meanwhile, Giangreco *et al.* (2009) found that most organizations never evaluated it because of the associated high cost, complexity, and inability to manage. Therefore, a general instrument is needed to offer an early picture of organizational performance due to training.

2.4 Operational Definition

A general instrument to measure a self-report of training effectiveness at all levels is developed and tested to determine the early picture of training effectiveness; this is named as the General Training Effectiveness Scale (GTES). Based on the literature review, each level of training effectiveness proposed by Cannon-Bowers *et al.* (1995) is operationalized in Table 1.

Table 1. Operational definition

Dimension	Description
Learning Performance	Learning performance evaluation is the improvement or changes in declarative knowledge, procedural knowledge, and meta-cognition. Declarative knowledge includes knowledge about theories, facts, and method to use the training outcomes. Procedural knowledge includes the ability to use or apply training outcomes. Meta-cognition includes believes, certainty, and confidence in mastering knowledge and skills emphasized in training.
Individual Performance	Individual performance evaluation is the improvement or changes in competencies, efficiencies, and effectiveness in workplace due to training outcomes. Competencies include the ability to solve work-related problems, produce better quality works, and fewer mistakes. Efficiencies include faster working; effectiveness includes the increase in total number of works.
Organizational Performance	Organizational performance evaluation is the improvement or changes in teamwork, customer satisfaction, and organization's goals achievement and reputation improvement due to training outcomes.

3. Methodology and Findings

To construct GTES, the current research follows Hinkin's (1998), Neuman's (2010), and Hair's et al. (2010) suggestions. Hinkin (1998) suggests organizing item generation, questionnaire administration, initial item reduction, confirmatory factor analysis (CFA), convergent/discriminant validity, and replication to develop questionnaire. Meanwhile, Neuman (2010) suggests testing face, content, construct, and criterion validity (e.g., concurrent validity) to validate an instrument; Hair et al. (2010) suggest testing reliability using Cronbach's Alpha test and construct reliability (CR) before performing construct validity. Hence, these suggestions were followed by organizing a quantitative approach through four studies including jury validation, pilot study, survey study, and extension study. In these studies, the GTES' reliability, face and content validity, CFA, construct validity, construct reliability, and concurrent validity were assessed.

The first version of GTES comprises of 30 items; then, four studies were organized to refine the GTES, leaving only 10 items remaining in the final study. The first version of GTES comprises of three different sections to measure learning performance, individual performance, and organizational performance. The GTES was developed by conceptualizing and operationalizing each construct as suggested by Cannon-Bowers's et al. (1995) model; then, items were self-constructed and tested in Malay language; however, for the purpose of reporting research outcomes, the English Version is used. Data were analysed using SPSS and AMOS.

3.1 Study 1 - Jury Validation

In jury validation, three experts were appointed as juries to determine face and content validity. These experts were lecturers in University Putra Malaysia (UPM) from various specializations within HRD, such as in research method, HRD policies, and training evaluation. Using a quantitative approach, these experts play the role as juries to validate the instrument by giving their scores from one to 10, as well as comments to improve each item. Score one represents "strongly disagreed" and 10 represents "strongly agreed". Using average scores, only items that have score greater than 7 were included in the GTES, this is consistent with suggestion of accepting reliability (0.7) by Hair *et al.* (2010). Based on study 1, the average score for nine selected items in learning performance was 8.96, 10 selected items in individual performance was 9.07, and seven selected items in organizational performance was 8.90. Then, each item was modified according to these experts. In this study, only 26 items remained in the GTES.

3.2 Study 2 - Pilot Study

A pilot study was organized to determine the reliability for GTES. It involved 33 academic staff in a faculty in a public university in Malaysia in the early part of 2011. These respondents previously attended various training programs organized by the Academic Development Centre (ADC) at the university. The GTES was distributed using formal e-mails to 246 academic staff at the faculty, of which only 33 were willing to become involved in this study and returned the e-mail. To become involved in the pilot test, these respondents have to remember one training program organized by the ADC that they attended, and proceed to answer the questions honestly. They were given a week to return the e-mail. Findings indicated that the Cronbach's Alpha for learning performance was 0.812, individual performance was 0.797, and organizational performance was 0.826. In this study, 26 items remained in the GTES.

3.3 Study 3 - Survey Study

A survey was organized for academic staff in a public university in Malaysia (the same university subject to the pilot study) to test for the hypotheses as well as to determine the CFA, construct reliability, and construct validity. These academic staff previously attended one of 17 training programs organized by the ADC in 2011. ADC is a department in the university that provides training programs for academic staff within the university. According to Hair *et al.* (2010), items should be reduced in CFA to get acceptable GOF (goodness of fit) in order to make sure that these items could represent a smaller number of constructs; this suggestion is followed before performing construct reliability and validity. Jonck (2015) also applied CFA to develop a valid instrument.

From a population of 338 participants, only 84% were willing to participate in the study; however, the sample size was reduced from 284 to 281 due to outliers. The 281 sample size had 0.99 value of power (1- β error probability) when tested using G-Power software. This indicates an adequate sample size (Fritz and MacKinnon, 2007). Participants were informed at the beginning of the training that participation in this research is voluntary and confidential. Data were collected at least two to three months after the completion of each training program. To maintain confidentiality and

anonymity, respondents were not required to give their names. Using SEM-AMOS, CFA was organized to determine appropriate items for each construct of training effectiveness used in the GTES. Then, the construct reliability and construct validity were determined, followed by determining the relationships between these constructs of training effectiveness.

A CFA was organized for each construct of training effectiveness in GTES. The findings indicated that each construct of GTES had acceptable GOF (see Table 2). Hair *et al.* (2010) stress that the GFI should be greater than 0.90 and the value for average variance extracted (AVE) should be greater than 0.5. In CFA, only 15 items remained for the GTES; these include evaluating training at learning, individual and organizational performance's level.

Table 2. Goodness-of-fit indices for each dimension of GTES based on CFA

GOF index	Number of selected items	CMIN (χ^2)	df	GFI	CFI	TLI	RMSEA	AVE
Learning Performance	6	17.131 ($p < .047$)	9	.980	.991	.985	.057 (PCLOSE < .345)	0.57
Individual Performance	5	9.856 ($p < .079$)	5	.986	.994	.988	.059 (PCLOSE < .330)	0.63
Organizational Performance	4	4.582 ($p < .101$)	2	.992	.995	.986	.068 (PCLOSE < .267)	0.63

Further, construct validity and construct reliability were assessed in the measurement model (see Table 3); the formula was taken from Hair *et al.* (2010). The measurement model includes selected items in CFA; however, the number of items for each construct of training effectiveness was reduced to get an acceptable GOF. The measurement model reduced the 15 items into 10 items to get significant p value. The measurement model had acceptable GOF with $\chi^2(32) = 40.837$ with $p < .136$, $\chi^2/df = 1.276$, GFI = .917, CFI = .995, TLI = .993, and RMSEA = .031 with PCLOSE < .0868. Each construct of training effectiveness had an AVE greater than 0.5 indicating convergent validation, was significantly correlated with each other indicating nomological validation, and the value for squared correlation between each construct was lesser than the value of AVE for each construct indicating discriminant validation. Additionally, the construct reliability for each construct was also greater than 0.7 indicating sufficient level of construct reliability. In sum, there were 10 items left in the measurement model, in which the measurement model had an acceptable level of GOF, construct validity, and construct reliability. The 10 items can be seen in Table 4.

Further, findings indicated a low relationship between learning and individual performance, and that between learning and organizational performance; although there is a high correlation between individual performance and organizational performance. This is consistent with previous research findings that find varied relationship from low to high. This implies that different sample will result in different results; this is demonstrated by findings in Bell and Ford (2007) and Tziner *et al.* (2007). In fact, variation in organizational culture (Bunch, 2007), variation in trainee characteristic including trainee's culture in region and occupational types (Tharenou, 2001), and variations in training context including technical-based skills, physical-based skills and cognitive-based skills (Cheng and Ho, 1998; Bell and Ford, 2007) can affect these relationships.

On the other hand, Klimoski (1991) argue that the constructs should be correlated to present a model; meanwhile, Tabachnick and Fidell (2012) argue that the relationship between constructs should be more than 0.3 to demonstrate acceptable relationship in a measurement model. Hence, combination of learning, individual, and organizational performance in GTES can be seen to evaluate general training effectiveness.

Table 4. Items selected for the final version of General Training Effectiveness Scale (GTES)

Levels of evaluation	Num. Items (English Version)	Items (Malay Version)
Learning Performance	1	I can lists down all the important things emphasized in this training. <i>Saya boleh menyenaraikan semua perkara penting yang diajar dalam latihan ini.</i>
	2	I know how to solve certain job problems using the skills taught in this training. <i>Saya tahu bagaimana untuk menyelesaikan masalah kerja tertentu dengan menggunakan kemahiran yang diajar dalam latihan ini.</i>
	3	I know how to work more efficient using the knowledge learned in this training. <i>Saya tahu bagaimana untuk bekerja dengan lebih cekap dengan menggunakan ilmu yang dipelajari dalam latihan ini.</i>
	4	I have the capability to perform the skills taught in this training. <i>Saya berkebolehan menggunakan kemahiran yang diajar dalam latihan ini.</i>

Individual Performance	5	My personal competencies have improved after attending this training.	Kompetensi diri saya telah meningkat selepas menghadiri latihan ini.
	6	I am being more professional in certain tasks after attending this training.	Saya telah menjadi lebih profesional dalam tugas tertentu selepas menghadiri latihan ini.
	7	My job performance has improved as a result of applying the skills emphasized in this training.	Prestasi kerja saya telah meningkat hasil daripada menggunakan kemahiran yang ditekankan dalam latihan ini.
Organizational Performance	8	The productivity of my department has improved due to the skills that I learned and used in this training either directly or indirectly.	Produktiviti jabatan saya telah meningkat sejak saya menggunakan kemahiran yang diajar dalam latihan ini sama ada secara langsung atau tidak.
	9	What I have learned in this training has improved my job performance and subsequently my organizational performance.	Saya percaya apa yang saya pelajari dalam latihan ini telah meningkatkan prestasi kerja saya dan seterusnya prestasi organisasi saya.
	10	I have contributed to improving my organization's reputation due to the outcome of this training either directly or indirectly.	Saya percaya saya telah menyumbang kepada peningkatan reputasi organisasi saya kesan daripada latihan ini sama ada secara langsung atau tidak.

Table 3. Average variance extracted (AVE), constructs reliability (CR), correlation, and squared correlation among GTES constructs

AVE	CR	Number of items		LP	IP	OP
.708	.906	4	LP	-	.31	.31
.711	.881	3	IP	.0961	-	.80
.669	.858	3	OP	.0961	.6400	-

Notes: Values above the diagonal are correlations as produced by SEM. Values below the diagonal are squared correlation. All covariance are significant at .0001 level of significant, LP = learning performance, IP = individual performance, OP = organizational performance

Furthermore, Figure 1 shows the structural model for GTES with acceptable GOF (the same value as measurement model). The structural model shows that the most important construct of training effectiveness was organizational performance ($\beta < .89$), followed by individual performance ($\beta < .89$), and learning performance ($\beta < .35$). Table 4 shows the final version of GTES as produced in the survey study.

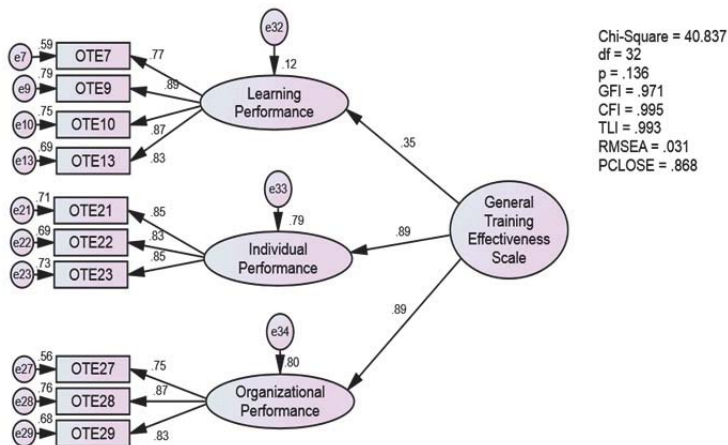


Figure 1. Structural Model for General Training Effectiveness Scale (GTES)

Notes: All regression weights and variances are significant at least at .05 level of significant

3.4 Study 4 - Extension Study

An extension study was organized to determine the concurrent validity of GTES among 31 non-academic staff in a faculty at the public university in Malaysia (the same university subject to the pilot and survey study) who has attended a motivation training program on 27 to 29 December 2013 at the Commonwealth Forest Park & Resort Rawang, Selangor, Malaysia. The objectives of the training were to improve teamwork, positive attitudes, and work motivation among trainees. Hence, some established instruments were used to evaluate the training effectiveness. The instrument developed by Stevens and Campion (1994) was used to measure learning performance by determining the level of teamwork's knowledge, skill, and ability before and after the completion of training using trainees' self-report. Meanwhile, the instrument developed by Hilliard (2000) was used to assess individual performance before and three months after the completion of training by the trainees' supervisor. Besides, GTES was distributed three months after the completion of training.

The findings indicated a significant difference of learning performance [$t(30) = 16.3, p < 0.0001$] as measured before ($M = 4.0645, SD = .4439$) and after the completion of training ($M = 2.9055, SD = .46342$) using the instrument developed by Stevens and Campion (1994). Further, there was a significant difference of individual performance [$t(30) = 20.7, p < 0.0001$] as measured before ($M = 4.6994, SD = .1989$) and three months after the completion of training ($M = 3.8581, SD = .3070$) using the instrument developed by Hilliard (2000). Interestingly, Table 5 shows a significant correlation between post learning performance and post individual performance ($r = .835, p < .001$); and that between post individual performance (as measured by Hilliard, 2000) and training effectiveness as measured by GTES ($r = .769, p < .001$). These have shown that the improvement of individual performance due to training that measured using instrument developed by Hilliard (2000) had a significant, positive, and strong correlation with training effectiveness as measured by GTES. The higher scores for individual performance will increase the higher scores for training effectiveness; this implies the concurrent validity for GTES. However, training effectiveness (as measured by GTES) had a moderate relationship with post learning performance that measured using instrument developed by Stevens and Campion (1994). This is because instrument by Stevens and Campion (1994) is developed to measure the level of teamwork's knowledge, skill, and ability instead of general learning performance.

Table 5. Correlations between GTES and other instrument

Types of instrument		General Training Effectiveness Scale (GTES)	Individual Performance	Learning Performance of teamwork's knowledge, skill, and ability
General Training Effectiveness Scale (GTES)	Pearson Correlation	-	.769**	.548**
	Sig. (2-tailed)		.000	.001
	N		31	31
Individual Performance	Pearson Correlation	.769**	-	.835**
	Sig. (2-tailed)	.000		.000
	N	31		31
Learning Performance of teamwork's knowledge, skill, and ability	Pearson Correlation	.548**	.835**	-
	Sig. (2-tailed)	.001	.000	
	N	31	31	

4. Limitations and Future Research

There are some limitations to the proposed GTES. First, the internal and external validities were not verified because the research did not apply experimental research design; this should be considered in future research. Second, GTES was tested among academic and non-academic staff in a large public university in Malaysia using the Malay version; it might have different effects for the English version using a different sample, such as organizations in different regions, co-business, culture, job position, and etc. Hence, it is recommended for future researchers to test GTES and compare the different language versions using other samples. Third, data were collected using self-report; in which, some people believe that data should be compared with other sources, such as cost and profit report, productivity report, customer complain report, and other kinds of data. Hence, these should be considered in future research. Fourth, GTES is an instrument to evaluate training effectiveness in general; HRD researchers and practitioners should make intensive training evaluation if they need detailed information to improve their training programs.

5. Conclusion and Implications

The objective of this article was to discuss research findings on developing and testing the validity and reliability of a self-report instrument named the General Training Effectiveness Scale (GTES) for workplace learning. The GTES items were developed using training effectiveness model by Cannon-Bowers *et al.* (1995). Four studies were organized to determine the psychometric property for GTES using a Malaysian sample including Study 1 (jury validation), Study 2 (pilot test), Study 3 (survey study), and Study 4 (extension study). The first version of GTES comprised of 30 items; however, it was reduced to 10 items in the final version of GTES. The final version of GTES had passed some tests including face validity, content validity, CFA, construct validity, cronbach alpha reliability, construct reliability, sufficient GOF for structural model of GTES, and concurrent validity. The GTES can be used to determine the effectiveness of employees training using a self-report that usually used in survey to determine the overall picture of training effectiveness with minimum expenses, time or energy. However, since GTES includes the evaluation of individual and organizational performance, GTES should be used at least two to three months after the completion of training.

General Training Effectiveness Scale (GTES) is useful for the society, including for researchers, practitioner, and public to improve workplace learning by distinguishing effective and ineffective employee training. Using GTES, a self-report of training participant can be produced to determine the level of training effectiveness. The low level of GTES's score will picture ineffective training, and the other way around. In fact, HRD practitioners especially in SMEs (small to medium enterprise) can use GTES as an easier way to picture training effectiveness without applying complicated training evaluation that usually involves high cost, time, and expertise; in which, the cost of training evaluation can be higher than the cost of training itself (Chong, 2005; Giangreco *et al.*, 2010). This can be seen as a more economic manner to determine training effectiveness since practitioners in SMEs usually get training from training vendors instead of developing the training program. Additionally, GTES can also benefit the public by offering a self-report instrument to determine the effect of training that they have attended. For researchers, GTES can be used in a survey to determine the relationship or comparison between training effectiveness and other variables, such as organizational climate, training characteristic, and employee's criterion to increase the effectiveness of training. In fact, it can also benefits the undergraduate/postgraduate students that needs to study about training since there is limited instrument to measure general training effectiveness.

References

- Alliger, G. M., Tannenbaum, S. L., Bennett, W. Jr., Traver, H., & Shetland, A. (1997), "A meta-analysis of the relations among training criteria", *Personnel Psychology*, Vol. 50, pp. 341-358.
- Axtell, C. M., Maitlis, S., and Yearta, S. K. (1997), "Predicting immediate and longer-term transfer of training", *Personnel Review*, Vol. 26 No. 3, pp. 201-203.
- Baldwin, T. T., and Ford, J. K. (1988), "Transfer of training: A review and directions for future research", *Personnel Psychology*, Vol. 41, pp. 63-105.
- Baldwin, T. T., Magjuka, R. J., and Loher, B. (1991), "The perils of participation: Effects of the choice of training on trainee motivation and learning", *Personnel Psychology*, Vol. 44, pp. 51-65.
- Bell, B. S., and Ford, J. K. (2007), "Reactions to skill assessment: The forgotten factor in explaining motivation to learn", *Human Resource Development Quarterly*, Vol. 18 No. 1, pp. 33-62.
- Bersin, J. (2008), *The training measurement book*, San Francisco: Pfeiffer/John Wiley & Sons.
- Brinkerhoff, R. O. (2006), "Increasing impact of training investments: An evaluation strategy for building organizational learning capability", *Industrial and Commercial Training*, Vol. 38 No. 6, pp. 303-307.
- Bunch, K. J. (2007), "Training failure as a consequence of organizational culture", *Human Resource Development Review*, Vol. 6 No. 2, pp. 142-163.
- Cannon-Bowers, J. A., Salas, E., Tannenbaum, S. I., and Mathieu, J. E. (1995), "Toward theoretically based principles of training effectiveness: A model and initial empirical investigation", *Military Psychology*, Vol. 7 No. 3, pp. 141-164.
- Cervero, R. M. (1988), *Effective continuing education for professionals*, San Francisco: Jossey-Bass Publishers.
- Chen, P. S., and Chih, J. T. (2012), "The relations between learner motivation and satisfaction with aspects of management training", *International Journal of Management*, Vol. 29 No. 2, pp. 545-561.
- Cheng, E. W. L., & Ho, D. C. K. (1998), "The effects of some attitudinal and organizational factors on transfer outcomes", *Journal of Managerial Psychology*, Vol. 13, No. 5/6, pp. 309-317.
- Chiaburu, D. S., and Tekleab, A.G. (2005), "Individual and contextual influences on multiple dimensions of training effectiveness", *Journal of European Industrial Training*, Vol. 29 No. 8, pp., 604-626.
- Chiaburu, D. S., Dam, K. V., and Hutchins, H. M. (2010), "Social support in the workplace and training transfer: A longitudinal analysis", *International Journal of Selection and Assessment*, Vol. 18, pp. 187-200.
- Colquitt, J. A., LePine, J. A., and Noe, R. A. (2000), "Toward an integrative theory of training motivation: A meta-analytic path analysis of

- 20 years of research", *Journal of Applied Psychology*, Vol. 85 No. 5, pp. 678-707.
- Fcateau, J. D., Dobbins, G. H., Russell, J. E., Ladd, R. T., and Kudisch, J. D. (1995), "The influence of general perceptions of the training environment on pre-training motivation and perceived training transfer", *Journal of Management*, Vol. 21 No. 1, pp. 1-25.
- Ford, J. K. (Ed.). (2014), *Improving training effectiveness in work organizations*, New York: Psychology Press.
- Fritz, M. S., and MacKinnon, D. P. (2007), "Required sample size to detect the mediated effect", *Psychological Science*, Vol. 18 No. 3, pp. 233-239.
- Garavan T. N., Morley, M., and Flynn, M. (1997), "360 degree feedback: Its role in employee development", *Journal of Management Development*, Vol. 16 No. 2, pp. 134-147.
- García, M. (2005), "Training and business performance: The Spanish case", *International Journal of Human Resource Management*, Vol. 16 No. 9, pp. 1691-1710.
- Giangreco, A., Carugati, A., Denmark, A., and Sebastiano, A. (2010), "Are we doing the right thing? Food for thought on training evaluation and its context", *Personnel Review*, Vol. 39 No. 2, pp. 162-177.
- Giangreco, A., Sebastiano, A. and Peccei, R. (2009), "Trainees' reactions to training: An analysis of the factors affecting overall satisfaction with training", *The International Journal of Human Resources Management*, Vol. 20 No. 1, pp. 96-111.
- Griffin, R. P. (2010), "Means and ends: Effective training evaluation", *Industrial and Commercial Training*, Vol. 42 No. 4, pp. 220-225.
- Griffin, R. P. (2012), "A practitioner friendly and scientifically robust training evaluation approach", *Journal of Workplace Learning*, Vol. 24 No. 6, pp. 393-402
- Hair, J. F. J., Black, W. C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2010), *Multivariate data analysis* (7 ed.), Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Hilliard, P. A. (2000), *Comparison of the predictive validity of a written test, an integrity test, a conscientiousness questionnaire, a structured behavioural interview and a personality inventory in the assessment of job applicants' background investigations, and subsequent task and contextual job performance*, Doctoral dissertation, University Of Southern California, USA (UMI No. 3018089).
- Hinkin, T. R. (1998), "A brief tutorial on the development of measures for use in survey questionnaires", *Organizational Research Methods*, Vol. 1 No. 1, pp. 104-121.
- Holton, E. F., III. (2005), "Holton's Evaluation Model: New evidence and construct elaborations", *Advances in Developing Human Resources*, Vol. 7 No. 37, pp. 37-54.
- Horgan, J., and Muhlau, P. (2006), "Human resource systems and employee performance in Ireland and the Netherlands: A test of the complementarity hypothesis", *International Journal of Human Resource Management*, Vol. 17 No. 3, pp. 414-439.
- Jonck, P. (2015). Validating an employer graduate-employability skills questionnaire in the faculty of management sciences. *Mediterranean Journal of Social Sciences*, 6(2), 230-237.
- Kirkpatrick, D. (1959), "Techniques for evaluating training programs", In Kirkpatrick, D. (1996), "Great ideas revisited: Revisiting Kirkpatrick's four-level model", *Training and Development*, Vol. 50, pp. 54-57.
- Kirkpatrick, D. (1996), "Great ideas revisited: Revisiting Kirkpatrick's four-level model", *Training and Development*, Vol. 50, pp. 54-57.
- Kirkpatrick, D. L., and Kirkpatrick, J. D. (2010), *Evaluating training programs: The four levels* (3 ed.), USA: Berrett-Koehler Publishers.
- Klimoski, R. (1991), "Theory presentation in Human Resource Management", *Human Resource Management Review*, Vol. 1 No. 4, pp. 253-271.
- Kontoghiorghes, C. (2004), "Reconceptualizing the learning transfer conceptual framework: Empirical validation of a new systemic model", *International Journal of Training and Development*, Vol. 8 No. 3, pp. 210-221.
- Kraiger, K., Ford, J. K., and Salas, E. (1993), "Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation", *Journal of Applied Psychology*, Vol. 78, pp. 311-328.
- Maimunah Ismail, and AAhad M. Osman Gani (Ed.), (2012), *Human Resource Development in Malaysia*, Petaling Jaya, Selangor, Malaysia: Pearson.
- Neuman, W. L. (2010), *Social research methods: Qualitative and quantitative approaches* (7 ed.), Massachusetts: Allyn and Bacon.
- Nijman, D. J. M., Nijhof, W. J., Wognum, A. A. M., and Veldkamp, B. P. (2006), "Exploring differential effects of supervisor support on transfer of training", *Journal of European Industrial Training*, Vol. 30 No. 7, pp. 529-549.
- Nikandrou, I., Brinia, V., and Bereri, E. (2009), "Trainee perceptions of training transfer: An empirical analysis", *Journal of European Industrial Training*, Vol. 33 No. 3, pp. 255-270.
- Noe, R. A. (2010), *Employee training and development* (5 ed.), Boston: McGraw Hill.
- Phillips, J. J. (1997), *Return on investment in training and performance improvement programs* (1 ed.), USA: Butterworth-Heinemann.
- Pilati, R., and Borges-Andrade, J.E. (2008), "Affective predictors of the effectiveness of training moderated by the cognitive complexity of expected competencies", *International Journal of Training and Development*, Vol. 12 No. 4, pp. 226-237.
- Pineda-Herrero, P., Belvis, E., Moreno, V., Duran-Bellonch, M. M., and Ucar, X. (2011), "Evaluation of training effectiveness in the Spanish health sector", *Journal of Workplace Learning*, Vol. 23 No. 5, pp. 315-330.
- Powell, J. L. (2009), *Transfer initiation and maintenance of training: Employees' perception of the relative influences of transfer intentions, general self-efficacy (GSE) and supervisor support*, Unpublished doctoral dissertation, Chesterfield University, Falls Church, Virginia, USA.
- Quinones, M. A. (1997), "Contextual influences in training", In M. A. Quiñones and A. Ehrenstein (Eds.), *Training for a Rapidly Changing Workforce*, pp. 177-99, Washington, DC: American Psychological Association.
- Ramos, V., Rey-Maqueira, J., and Tugores, M. (2004), "The role of training in changing an economy specialising in tourism",

- International Journal of Manpower*, Vol. 25 No. 1, pp. 55-72.
- Scaduto, A., Lindsay, D., and Chiaburu, D.S. (2008), "Leader influences on training effectiveness: Motivation and outcome expectation processes", *International Journal of Training and Development*, Vol. 12 No. 3, pp. 158-170.
- Schonewille, M. (2001), "Does training generally work? Explaining labour productivity effects from schooling and training", *International Journal of Manpower*, Vol. 22 No.1/2, pp. 158-172.
- Stanford, M. A. (2000), *Individual and situational influences on pre-training motivation: Its impact on training motivation*, Doctoral dissertation, North Carolina State University, USA (UMI No. 9994056).
- Stehle, S., Spinath, B., and Kadmon, M. (2012), "Measuring teaching effectiveness: Correspondence between students' evaluations of teaching and different measures of student learning", *Res High Educ*, Vol. 53, pp. 888-904.
- Stevens, M. J., and Campion, M. A. (1994), "The knowledge, skill, and ability requirements for teamwork: Implications for human resource management", *Journal of Management*, Vol. 20 No. 2, pp. 503-530.
- Tabachnick, B. G., & Fidell, L. S. (2012), *Using multivariate statistics* (6 ed.), Needham Hights, USA: Allyn and Bacon.
- Tharenou, P. (2001), "The relationship of training motivation to participation in training and development", *Journal of Occupational and Organizational Psychology*, Vol. 74, pp. 599-62.
- Tracey, J. B., Hinkin, T. R., Tannenbaum, S. I., and Mathieu, J. E. (2001), "The influence of the individual characteristics and the work environment on varying levels of training outcomes", *Human Resource Development Quarterly*, Vol. 12 No. 1, pp. 210-214.
- Tziner, A., Fisher, M., Senior, T., and Weisberg, J. (2007), "Effects of trainee characteristics on training effectiveness", *International Journal of Selection and Assessment*, Vol. 15 No. 2, pp. 167-174.