Professional Development:
The International Journal of
Continuing Social Work Education

The Human Services Training Effectiveness Postcard (HSTEP): A Tool for Research and Evaluation of Human Services Training

Journal:	Professional Development: The International Journal of Continuing Social Work Education
Article Title:	*The Human Services Training Effectiveness Postcard (HSTEP): A Tool for Research and Evaluation of Human Services Training*
Author(s):	*Dale Curry, and Theodore Chandler*
Volume and Issue Number:	*Vol. 2 No. 2*
Manuscript ID:	*22043*
Page Number:	*43*
Year:	*1999*

Professional Development: The International Journal of Continuing Social Work Education is published three times a year (Spring, Summer, and Winter) by the Center for Social Work Research at 1 University Station, D3500 Austin, TX 78712. Journal subscriptions are $110. Our website at www.profdevjournal.org contains additional information regarding submission of publications and subscriptions.

ISSN: 1097-4911

URL: www.profdevjournal.org **Email: www.profdevjournal.org/contact**

The Human Services Training Effectiveness Postcard (HSTEP): A Tool for Research and Evaluation of Human Services Training

Dale Curry, PhD, and Theodore Chandler, PhD

In recent years, there has been dramatic growth and transformation in the area of training for human service professionals. For example, VanderVen (1990) described several changes in the area of child- and youth-care worker training: (1) a greater growth in agency sponsored training compared to academic education programs; (2) an increase in the scope of clientele for care-giving personnel; (3) an increase in distance models and alternative delivery systems to increase availability; (4) an increase in the variety of sponsors of training; and (5) a growing infusion and exchange in international training and education activities.

While the emphasis on training in human services has expanded, the development of evaluation and tools for this training have been negligible. Most training evaluation measures still rely on participant response/reaction assessments upon the immediate conclusion of training (Bramely, 1991; Clark & Voogel, 1985; Garavaglia, 1993; Krein & Weldon, 1994). Although helpful, immediate reaction evaluations are limited partly because they probably are influenced by a number of factors such as weak versus powerful endings to sessions, early dismissal from training, and trainers who provide candy and gifts for the participants during the training.

Reaction evaluations, completed at the conclusion of training, could be more helpful if accompanied by other indicators of training success or failure (Curry, Caplan & Knuppel, 1994; Jones, Stevenson, Leung & Cheung, 1995). Here is described an evaluative tool which can be used to better assess the effectiveness of human services training as well as to provide information on the application process. The use of the Human Services Training Effectiveness Postcard (HSTEP) as an outcome indicator in a research study on transfer of learning at the Northeast Ohio Regional Training Center is emphasized.

DESCRIPTION OF THE HUMAN SERVICES TRAINING EFFECTIVENESS POSTCARD

The Northeast Ohio Regional Training Center (NEORTC), a component of the Ohio Child Welfare Training Program, has utilized a brief mail survey questionnaire in the form of a postcard to assess training. The HSTEP questionnaire uses a scale which ranges from strongly agree to strongly disagree. It was designed to incorporate the four major areas/levels of training evaluation in Kirkpatrick's (1975) model into a single self-reporting assessment index. Kirkpatrick's widely accepted 4-level evaluation model includes: (1) reaction/satisfaction, (2) learning, (3) behavior/transfer of learning, and (4) results of transfer of learning (Bell & Kerr, 1987; Birnbrauer, 1987; Krein & Weldon, 1994; Pine & Tingley, 1993). Table 1 lists items on the HSTEP and their association with the four levels.

Table 1

Relationship Of Evaluation Postcard Items to the Four Levels of Kirkpatrick's Model

Level	Item Number and Description
1. Reaction/satisfaction	#1-Overall, I am very satisfied with the workshop.
2. Learning	#2-During the workshop, I learned a substantial amount of information.
3. Behavior/transfer of learning	#3-I have used the knowledge and skills I learned from the workshop on the job.
4. Results of transfer of learning	#4-As a result of using the knowledge/skills from the workshop, I have observed client progress.
	#5-As a result of the workshop, I am a more effective worker.

Dale Curry is Training Coordinator, Northeast Ohio Regional Training Center, Summit County Children Services, Akron, OH. Theodore Chandler is Professor of Educational Psychology, Kent State University, Kent, OH 44204.

The HSTEP is intended also to provide acceptable return rates from training participants who may have high caseloads and little time for assisting in the evaluation of training programs. Utilization of the HSTEP in a NEORTC study of 598 child protection social workers from 14 counties in northeast Ohio indicated that the postcard approach can elicit adequate returns from participants (Curry, 1997).

In addition to providing quantitative outcome data, the instrument was designed to elicit qualitative information on the process of transfer of learning. The participants are requested to list factors which helped or hindered their application of learning. The participant responses can be used to identify reasons for transfer success and/or failure as well as suggest possible ways to intervene in the learning and transfer process, before, during, and after training.

Although this study emphasizes the use of the HSTEP with child protective services (CPS) workers, the HSTEP was intended to be used for assessment of training for a wide variety of human service worker populations and training content area

METHOD

Participants

Study participants were direct line CPS social workers who attended training at the Northeast Ohio Regional Training Center over a three-month period of time. Ninety-six percent of the training participants (598), representing 14 counties in northeast Ohio, participated in at least the first phase of the study. Four hundred and ninety-seven (83.5%) were female and 98 (16.5%) were male. Five hundred and three (84.8%) were white/European-American and eighty (13.5%) were black/African-American. Public child welfare experience ranged from less than one year to 44 years with a mean age of 6.5 years. Ages ranged from 22 to 65 years, with a mean of 34.46. Caseload size ranged from zero to 99 cases, with a mean of 19.

PROCEDURE

Phase One: Data Collection

Participants of all NEORTC CPS social worker training during a three-month period were asked to complete a questionnaire (Transfer Potential Questionnaire) at the end of each workshop. The Transfer Potential Questionnaire (TPQ) was designed to assess factors which, according to research, influence transfer of learning. In addition to demographic information, the TPQ consisted of 68 transfer-related items which were combined into eleven transfer factors based upon the results of principal components analysis. (Table 2.) Table 3 provides a sampling of items from the TPQ. The reader should refer to Curry (1997) for a detailed description of the TPQ and discussion of its psychometric characteristics.

Frequencies, means, and standard deviation scores were obtained on each of the TPQ items. The 68 items were averaged to obtain a "transfer potential" score for each participant. High and low "transfer potential" scores were determined based on the percentile score. High "transfer potential" scores were those on the 50th percentile or higher. Low scores were below the 50th percentile.

Phase Two

Three months after each training workshop, the HSTEP and an instructional letter was mailed to each participant. The letter provided instructions for completion of the postcard and its return within two weeks. To facilitate questionnaire return, a reminder notice was mailed to all participants one week after sending the initial postcard questionnaire.

The rationale for administering the HSTEP three months subsequent to training was two-fold: (1) to provide sufficient time to apply training, and (2) to minimize self-reporting limitations due to forgetting over time. Rackham (1979) emphasized the importance of what he described as the "results dip" which can occur when trying out newly learned skills. He said that when a worker tries out

Table 2
Eleven Transfer Factors and Their Correlation with Composite "Transfer" Variable and Composite of All Five Items on the HSTEP

TPQ Factors	Correlation with "transfer"	Correlation with all five HSTEP items
1. Trainer Adult Learning and Transfer Strategies	.54	.60
2. Training Relevance and Applicability to the Job	.59	.61
3. Supervisory Support for Training and Application	.25	.23
4. Top Management Support for Training and Application	.26	.24
5. Application Plan	.49	.48
6. Participant Perceived Learning	.52	.60
7. Participant Motivation to Attend Prior to Training	.16	.14
8. Participant Prior Experience with Training and Application (Value of Previous Training)	.25	.23
9. Co-worker Support for Training and Application	.29	.27
10. Training/Organization Congruence	.31	.52
11. Pre-training Preparation	.38	.36

Note: All factors except factor 7 significant ($p < .001$) ; factor 7 significant ($p < .01$).

Table 3
Sample Items on the Transfer Potential Questionnaire

Item #	Item Description
1	As a result of training, I learned a substantial amount of information.
8	My supervisor is familiar with the content of this training.
18	The trainer helped me to see how the training can be applied on the job.
23	I have already made a plan with a co-worker to utilize this training.
35	I am very confident that I will use this training on the job.
46	The trainer/training helped me to become aware of underlying principles (rules) which can be used with different cases and situations.
47	Prior to attending the workshop, I heard that this training was "worthwhile"/ valuable.
63	This content is consistent with my agency's mission, philosophy, and goals.
65	This training will help me continue learning in this topic area.

new skills, there is often an initial, temporary decrease in the effectiveness of performance. According to Rackham, if a newly learned skill is not supported in some manner, behavior frequently returns to the pre-training level.

Three months after training has been mentioned as a key time to show that behavior change has occurred (Bickerstaffe, 1993; Garavaglia, 1993). Three months has generally been considered sufficient time for a participant to put training into practice and maintain new behavior.

Based upon information from a more extensive study, newly hired workers on probationary employment status were excluded from this study. Curry (1977) found no statistically significant relationship between the TPQ and the HSTEP for newly hired workers on probationary status. It was hypothesized that these newly hired workers were unable to assess accurately many of the transfer variables at the time of training, due to inexperience or concern about employment status. Results and implications from this study cannot be generalized to newly hired workers attending training during their probationary period.

In order to provide information about the reliability and validity of HSTEP, internal reliability analysis and principal components analysis were conducted on the five items which compose the HSTEP. In addition, a comparison between those who returned the HSTEP and those who did not was conducted on the 11 TPQ transfer factors and the overall "transfer potential" score.

The transfer items (3, 4 & 5) were combined to form a composite variable named "transfer." Frequencies, means, and standard deviation scores were obtained on all of the HSTEP items and the "transfer" composite variable. High and low "transfer" scores were also obtained, depending upon their percentile score. A high score was defined as one on the 50th percentile or above. A low score was defined as falling below the 50th percentile. A two by two matrix was created based upon high and low "transfer" and "transfer potential" scores.

In addition to the "transfer" variable, a composite variable consisting of all five HSTEP items was created. In order to understand the relationship between the HSTEP and the TPQ, correlational analysis was conducted with "transfer" and the composite HSTEP variable with the overall "transfer potential" score and each of the 11 transfer factors.

Since two of the 68 transfer items on the TPQ were identical to two of the five items on the HSTEP, a paired t-test analysis on the two items was utilized to explore potential differences. This was intended to provide some information concerning whether or not the time, cost, and effort involved in administering an evaluation tool three months after training is warranted. In order to examine the possible effect of application experience upon the ratings, one-way analysis of variance was also conducted, comparing the high and low "transfer" groups. In addition, correlational analysis was performed on the difference scores and "transfer."

The relationship between several demographic variables and "transfer" was also explored. One-way Analysis of Variance was conducted on participant gender, race, and size of agency to examine their effect upon "transfer." Similarly, a correlational analysis was conducted with caseload size, age, and years of experience with "transfer."

In addition to the quantitative analysis, a content analysis of these open-ended statements was conducted. Participants were requested to list on the card factors which helped or hindered their application of learning on the job. The comments were summarized and categorized according to the 11 factors assessed by the Transfer Potential Questionnaire (Table 2). An "other" category was also created. Similar responses in the other category were given an appropriate category name (e.g., time and caseload demands). Two Training Coordinators from the Northeast Ohio Regional Training Center independently assigned the responses into the categories. An inter-rater reliability percentage agreement rate was obtained.

RESULTS

Reliability

Results of the NEORTC study indicated that the HSTEP has excellent reliability as a quantitative outcome indicator and good reliability as a method for exploring the transfer process. Internal reliability analysis of the five items resulted in a .91 reliability coefficient (Cronbach's alpha). Inter-rater reliability of the content analysis of the open-ended responses (factors which helped or hindered transfer) resulted in an 83% agreement rate.

When comparing those who returned the HSTEP (74% of the participants) with those who did not on the previously completed Transfer Potential Questionnaire, Oneway Analysis of Variance revealed that only one of eleven transfer of learning factors (participant motivation to attend prior to training) was significantly influenced by the 26% non-return rate. Apparently, those participants who were less motivated to attend training were also less motivated to return the HSTEP. However, the overall average of the TPQ's items was not significantly affected. In general, the non-returns appeared to have only a minimal effect on the study's results (Curry, 1997).

Validity

The most crucial issue in instrument construction is validity. A training evaluation instrument should measure what it is intended to measure as well as provide useful information to training evaluation professionals. The Standards for Educational and Psychological Testing (AERA, APA, NCME, 1985) emphasize three major types of validity: (1) content-related, (2) criterion-related, and (3) construct-related. The findings from the study provided partial support for validity in each of these three areas.

Content-Related Validity

Content-related validity evidence refers to the extent that HSTEP items are representative of what they are trying to measure. The HSTEP is intended to provide a general measure of human service training effectiveness as well as identify factors which affect application of learning from training to the job. It is also intended to be utilized for a wide variety of human service training content areas.

As stated previously, the items on the HSTEP were developed to provide a general assessment of the four levels in Kirkpatrick's (1975) widely accepted training evaluation model (reaction/satisfaction, learning, transfer of learning, and the results of transfer). Items one through three address levels one, two, and three. Items four and five pertain to level four (results). Although additional items addressing each level may have increased the tool's reliability and validity as a training effectiveness indicator, the convenience of the brief postcard approach, and associated adequate return rate, would probably diminish. In a general manner, the items on the HSTEP address the most cogent elements of training effectiveness. (The results of how the participants responded to each item are provided in a later section on training effectiveness data.)

In addition to providing a quantitative index of training effectiveness, the HSTEP provided training participants the opportunity to feed back information on factors which influenced their application efforts. The participants were asked to provide information on factors which helped and/or hindered their application. (Information on how the participants responded is provided in a subsequent section on qualitative analysis.)

Criterion-Related Validity

Another major approach to determining validity is criterion validity. This approach compares an instrument's scores with another outside measure. Although comparisons of the HSTEP with other concurrent indicators of transfer on the job (e.g., supervisor ratings, completion of participant action plan objectives) has not yet occurred, the overall average "transfer potential" score on the Transfer Potential Questionnaire, which was completed at the conclusion of training, significantly correlated with a composite "transfer" score (items 3, 4, and

5) on the HSTEP, which was completed three months after training ($r=.62$, $p < .001$). The use of a composite score of all five HSTEP items resulted in a correlation of .64 ($p < .001$). The HSTEP "transfer" score as well as the composite mean of all five HSTEP items also significantly correlated with each of the eleven transfer factors assessed by the Transfer Potential Questionnaire. (Table 2.)

In addition to the quantitative HSTEP ratings, a content analysis of the open-ended statements of factors which influenced application resulted in the identification of factors consistent with the quantitative results from the TPQ. This will be discussed in a later section on qualitative analysis.

Construct-Related Validity

Principal Components Analysis of the HSTEP indicated the presence of one factor which accounted for 73% of the variance. Item five had the highest communality ("As a result of the workshop, I am a more effective worker"). The HSTEP appears to assess a worker's perception of improved competence on the job as a result of training. While the individual items can provide information regarding each of Kirkpatrick's four levels, a composite score of the five items is probably a more valid outcome indicator of training effectiveness.

Convergent and discriminant validity are considered to be aspects of construct validity. Conceptually, the HSTEP should correlate well with factors which theoretically are highly similar. However, it should have low or negative correlations with variables to which it is theoretically dissimilar (Groth-Marnat, 1990).

As stated previously, the overall mean score from the TPQ ("transfer potential") had a .62 correlation with the HSTEP variable "transfer" and a .64 correlation with the composite score of all five HSTEP items. Since transfer of learning is affected by many factors, the overall "transfer potential" score should be the best predictor of "transfer" on the job. As expected, no single item or factor on the TPQ had a higher correlation with HSTEP than the overall TPQ "transfer potential" score.

Other variables which correlated highly included "training relevance and applicability" (factor 2), "perceived learning" (factor 6), and factor 1, "trainer adult learning and transfer strategies" (Table 2). Since these factors are probably more under the influence of the training design, one would expect these factors to be more highly associated with training effectiveness than environmental variables such as supervisor, co-worker, and administrative support. Factors such as "pre-training motivation" and "previous experience with training" would also be expected to be somewhat lower rated since they are farther removed in time from when learning and transfer occur, compared to training design factors.

Variables which were not hypothesized to have a relationship with training effectiveness included demographic variables such as gender, race, and age. The research on transfer of learning has not associated these factors with transfer. As expected, these variables had no statistically significant relationship.

The following section on incremental validity is focused on the relationship between ratings completed at the immediate conclusion of training and ratings completed at a later time, after one has an opportunity to try out new learning. As expected, satisfaction ratings were generally lower after the training. It is hypothesized that the initial "euphoria" which is sometimes associated with training dissipates, and ratings are influenced by other factors such as application success or failure. Since the HSTEP appears to measure a worker's perception of improved competence on the job, an assessment of satisfaction three months after training should be positively influenced by application success and negatively influenced by application failure. Since the HSTEP appears to measure a worker's perception of improved competence on the job, an assessment of satisfaction and perceived learning three months after training could be influenced positively by application success and negatively affected by application failure. (The next section on incremental validity provides some support for this conjecture.)

Incremental Validity

In addition to the three major areas of validity, it is important to make a case for the value that is added by the increased effort/cost/time involved in administration of an evaluation instrument (Groth-Marnat, 1990). Since most training evaluations currently being conducted utilize participant reactions, does the HSTEP provide additional information of value? To what extent does the self-rating three months after training differ from a reaction evaluation conducted at the conclusion of training. Does the value added by administering a mail questionnaire postcard survey outweigh the increased effort and cost?

As previously discussed, two identical items on the 68 item Transfer Potential Questionnaire and the Evaluation Postcard (items 1 and 68 on the TPQ and 1 and 2 on the HSTEP) were explored. Items 1 on the TPQ and 2 on the HSTEP ("I learned a substantial amount of information") significantly correlated with each other ($r=.59$, $p<.001$). Paired t-test analysis found no significant overall change in rating. The mean rating at the immediate conclusion of training was 4.10 (S.D.=.90) compared to a mean rating three months later of 4.05 (S.D.=.88), ($p=.21$). However, when comparing high vs. low "transfer" participants, low "transfer" participants significantly decreased from a mean of 3.67 (S.D.=1.07) to 3.50 (S.D.=1.04) three months after training ($p<.05$). No significant change occurred for high "transfer" participants. Utilizing one-way analysis of variance, the difference scores resulted in a statistically significant F ratio of 4.96 ($p=.02$). Correlational analysis revealed a low, but, statistically significant relationship between "transfer" and the perceived learning difference scores ($r=.23$, $p<.001$).

When comparing TPQ item 68 and HSTEP item 1 ("Overall, I am very satisfied with the workshop"), paired t-test analysis resulted in a significant difference. The mean rating at the immediate conclusion of training was 4.38 (S.D.=.88). This compared to a mean rating of 4.20 (S.D.=.89) three months later on the HSTEP. The Pearson Correlation between the two items was .62 ($p<.001$).

When comparing high vs. low "transfer" participants, both groups resulted in a significant decrease in satisfaction ratings. The low "transfer" group decreased from 3.97 (S.D.=1.11) to 3.70 (S.D.=1.08), ($p=.001$). The high "transfer" group decreased from 4.59 (S.D.=.63) to 4.48 (S.D.=.59), ($p=.018$). Utilizing one-way analysis of variance, the difference scores in the high and low "transfer" groups resulted in a nearly significant F ratio of 3.62 ($p=.06$). There was smaller decrease in score three months after the training event for participants in the high transfer group. Correlational analysis resulted in a low significant correlation between the satisfaction difference scores and "transfer" ($r=.20$, $p<.001$).

Paired t-test analysis of the identical items within the two by two matrix (high and low "transfer potential" and high and low "transfer") revealed additional information. Generally, the scores were somewhat lower three months after training. However, unexpected high levels of training utilization was associated with an increased perception of learning and satisfaction.

Participants with high "transfer potential" but low "transfer" scores decreased in both satisfaction and perceived learning. The mean score for TPQ item 68 was 4.72 (S.D.>=.45) contrasted with a mean score of 4.22 (S.D.=.68) for HSTEP item 1 ($p<.001$). The mean score for TPQ item 1 was 4.28 (S.D.=.61) compared to a mean score of 4.06 (S.D.=.63) for HSTEP item 2 ($p=.06$). This was close to significance at the $P=.05$ level. Participants

with both high "transfer potential" and high "transfer" scores significantly decreased in satisfaction but not perceived learning. The mean rating for TPQ item 68 was 4.81 (S.D.=.40) compared to an HSTEP mean rating of 4.57 (S.D.=.51), ($p<.001$). The mean rating for TPQ items 1 was 4.51 (S.D.=.55) compared to an HSTEP mean rating of 4.47 (S.D.=.52), ($p=.33$).

Those with both low "transfer potential" and low "transfer" scores had a mean TPQ rating of 3.67 (S.D.=1.15) for item 68 compared to a mean rating of 3.46 (S.D.=1.18) for HSTEP item 1 ($p=.06$). This was not quite significant at the .05 level. However, TPQ item 1 had a mean rating of 3.40 (S.D.'=1.08) compared to a 3.26 (S.D.=1.13) rating for HSTEP item 2 ($p=.03$).

Mean ratings for those with low "transfer potential" but high "transfer" scores revealed a different pattern from the other groups. A significant increase in perceived learning occurred for this group. A mean rating of 3.97 on TPQ item 1 (S.D.=.77) increased to a mean rating of 4.19 (S.D.=.63) for item 2 on the HSTEP. Ratings also increased for satisfaction from a mean score of 4.19 (S.D.=.80) for TPQ item 68 to a mean of 4.34 (S.D.=.64) for HSTEP item 1 ($p=.15$). See figure 1. Although the satisfaction increase score was not statistically significant at the .05 level, it did not significantly decrease as in the other three groups.

The below data indicates that the CPS social workers assessed their training experiences differently three months after the training compared to their assessments immediately following training. Apparently, time, along with the participants' application experiences, affected the participants' perception of learning from, and satisfaction with, the training.

The HSTEP provides additional evaluative data which cannot be obtained during the training event. The next two sections will provide additional examples of information on utilization of training. Obviously, this information cannot be gathered until the participants have had an opportunity to try out newly learned skills.

The cost (time, effort, money) to administer the HSTEP is greater than most level-one (reaction/satisfaction) evaluations. However, it is relatively inexpensive when compared to other indicators of training effectiveness, which go beyond level one evaluation.

Figure 1. Difference scores (ratings assessed at the immediate conclusion of training on the TPQ subtracted from ratings assessed three months after training on the HSTEP).

	LOW "transfer"		HIGH "transfer"	
	Satisfaction	**Perceived Learning**	**Satisfaction**	**Perceived Learning**
HIGH "TRANSFER POTENTIAL"	**-.50**	**-.22**	**-.24**	**-.04**
LOW "TRANSFER POTENTIAL"	**-.21**	**-.14**	**+.15**	**+.22**

Training Effectiveness Data

Table 4 indicates the percentage of agree/disagree responses for all five items on the HSTEP. The data showed that a large majority of the participants were very satisfied with the training and perceived that they learned a substantial amount of information, used the what they learned on the job, and viewed themselves as more effective workers as a result of the training. Over 38% of the participants perceived that client progress was a result of the participant's application of learning from the training. The results clearly indicate that the majority of workers perceived the training to have had a positive impact on the job.

As previously stated, the transfer items 3, 4, and 5 were combined to create a composite variable named "transfer." Eleven factors were identified which significantly correlated with "transfer." Similarly, by assigning values from one to five (strongly disagree=1, disagree=2, undecided=3, agree=4, strongly disagree=5), all five items can be combined and/or averaged to create a user-friendly, single score index of training effectiveness. As stated previously, the results from the previously cited Principal Components Analysis of the five Evaluation Postcard items support the appropriateness of using a single composite score index.

Since the HSTEP can be used with almost any type of training, workshops can be compared with each other, based upon a common outcome indicator which incorporates Kirkpatrick's four evaluation levels. A common indicator can help to answer training effectiveness questions such as: (1) does the amount of time spent in training influence utilization? (2) using a standardized curriculum, which trainers are more effective? and (3) which training program is most cost effective?

A common indicator also would encourage experimental research. For example, the effect of specific training interventions such as the use of action planning could be compared across different content areas and different training populations. Although there is a need for training-specific outcome indicators, a generic training effectiveness index could serve many purposes.

Table 5 displays workshop titles and their corresponding HSTEP means (average of five items for all participants in the workshop) and standard deviation scores from the data from the child protection social worker study. The overall mean score for all of the listed workshops was 3.77 with a standard deviation of .79. One-way analysis of variance indicated significant differences among the workshops ($F= 3.60$, $p<.001$). Clearly, the workshops had varying degrees of perceived impact on the job.

Table 4
Percentage of Agreement/Disagreement on Items on the Evaluation Postcard

Item#	S. Disagree	Disagree	Undecided	Agree	S. Agree
1.	1.6%	3.3%	5.7%	49.3%	38.9%
2.	1.4%	5.7%	7.7%	55.5%	29.8%
3.	3.0%	8.2%	16.2%	56.5%	16.2%
4.	4.4%	14.2%	42.8%	31.3%	7.3%
5.	1.8%	6.8%	18.9%	56.4%	16.0%

Note. N=441

Table 5

Evaluation Postcard Means and Standard Deviation Scores by Workshop

Workshop Title	Mean	S.D.
Diagnosis and Treatment of Adult Emotional Disorders	3.66	.65
Secondary Trauma	4.21	.39
Family Assessment in Sexual Abuse Cases	3.24	.71
Hispanic/Latino Culture	4.02	.71
Adolescent Suicide	3.77	.57
Reality Therapy	3.66	.69
Dealing with the Death of a Client	4.03	.54
Crisis Intervention	3.47	.62
Adolescent Sexuality	3.62	.74
Leading Support Groups	3.64	.63
Seeking Closure: Secondary Trauma	4.11	.59
Family Centered Practice	3.81	.72
Reducing the Risk of Liability	3.59	.81
After Placement: Working with the Foster Family	3.91	.55
Understanding Childhood Psychopathology (DSM IV)	4.10	.39
Transcultural Placement	3.83	.31
The Mother/Daughter Relationship	2.99	1.03
Psychopharmacology for the Caseworker	4.25	.55
Attention Deficit Disorder	3.38	.85
Therapeutic Issues in Child Sexual Abuse Cases	4.17	.76
Maximizing Your Effectiveness (Time and Stress Management)	3.42	.85
Effective Casework with Gay and Lesbian Clients	4.31	.86
Understanding the Dynamics of Family Violence	4.29	.52
Assertive Communication for Effective Communication	4.00	.52
Street Drugs	3.74	1.06
Overview of Child Sexual Abuse	3.80	.36
Working with the African-American Family	2.89	1.28

Qualitative Analysis

Another purpose of the HSTEP instrument is to provide a qualitative assessment of factors which influence a participant's application of learning to the job. In the study, one hundred and forty-nine participants (34% of those who completed the HSTEP) listed at least one factor which promoted transfer. One hundred and twenty-six participants listed barriers to transfer (29% of those who completed the HSTEP). The results indicated that the three most frequently listed factors which facilitated the participant.

The above data indicate that the CPS social workers assessed their training experiences differently three months after the training compared to their assessments immediately following training. Apparently, time, along with the participants' application experiences, affected the participants' perception of learning from, and satisfaction with, the training.

The HSTEP provides additional evaluative data which cannot be obtained during the training event. The next two sections will provide additional examples of information on utilization of training. Obviously, this information cannot be gathered until the participants have had an opportunity to try out newly learned skills.

The cost (time, effort, money) to administer the HSTEP is greater than most level-one (reaction/satisfaction) evaluations. However, it is relatively inexpensive when compared to other indicators of training effectiveness, which go beyond level one evaluation.

Training Effectiveness Data

Table 4 indicates the percentage of agree/disagree responses for all five items on the HSTEP. The data showed that a large majority of the participants were very satisfied with the training and perceived that they learned a substantial amount of information, used the what they learned on the job, and viewed themselves as more effective workers as a result of the training. Over 38% of the participants perceived that client progress was a result of the participant's application of learning from the training. The results clearly indicate that the majority of workers perceived the training to have had a positive impact on the job.

Table 6
Factors Which Helped Transfer

Factor	Frequency	Percentage
Perceived Learning	50	34%
Trainer Adult Learning and Transfer Strategies	36	26%
Training Relevance and Applicability of Training	32	22%
Handouts	7	05%
Co-worker support	5	03%
Self motivation	5	02%
Self strategies (metacognition)	3	02%
Time and size of caseload	3	02%
Value of previous training	1	00.7%
Other	4	03%

Table 7
Factors Which Hindered Transfer

Factor	Frequency	Percentage
Training Relevance and Applicability of Training	51	40%
Adult Learing and Transfer Strategies	17	14%
Time and Size of Caseload Demands	17	14%
Perceived learning	11	09%
Administrative Support	7	05%
Client Resistance	7	05%
Other	16	13%

Since assessment of training effectiveness in human services is still a new frontier, the inclusion of qualitative methodology along with quantitative

approaches is essential. However, the trend toward an outcome-based managed care approach to human services administration is likely to affect the delivery and evaluation of human services training as well. This may increase pressure for "hard" quantitative data in order to demonstrate the training's impact.

Although the development of quantitative indicators of transfer is long overdue, the importance of qualitative approaches must not be underemphasized. Factors which influence training impact can be understood better with the inclusion of qualitative strategies. The questions of how (the process of learning and transfer) as well as how much (the extent of learning and transfer) can be more effectively addressed with an integrative methodological approach to training evaluation and research in human services.

Discussion

The results provide support for the use of the HSTEP as a tool to help determine both the process of how training is utilized (factors which help or hinder) as well as a worker's perception of the extent that training has improved one's competence on the job.

The paired t-test analysis of the identical learning and satisfaction items indicated that a participant's experiences after the training can change one's perception of satisfaction and perceived learning from the training. Generally, the satisfaction ratings lowered three months after training. However, the decrease appeared to be buffered by positive application experiences. The perceived learning score did not change significantly. However, there was a significant difference between the high and low "transfer" groups. Positive application experiences resulted in highter perceived learning scores. Apparently, successful experiences with application of learning, positively affected the satisfaction and perceived learning ratings.

The comparison of the HSTEP with the TPQ ratings on the identical satisfaction and perceived learning items, clearly showed that there were significant differences in outcome ratings. Use of the HSTEP three months after training can provide additional useful information, particularly when combined with other tools such as the TPQ.

Of particular interest is the finding that individuals who initially perceived low "transfer potential" but later had high rates of "transfer" increased their perceived learning and satisfaction ratings while the other groups' ratings lowered or did not change. In addition to the participant's application experiences, their expectations of application potential may also influence their satisfaction and perception of learning three months after training. Unexpected application of learning may increase later satisfaction and perceived learning scores. To the contrary, the concept of high or low expectation of application may also explain why participants with high "transfer potential" but low "transfer" scores had the largest decrease in satisfaction and perceived learning scores.

In addition to the quantitative information, the HSTEP also provides information on factors which helped or hindered training application. This additional quantitative and qualitative information provided by the HSTEP cannot be obtained at the time of training by any existing instrument.

Study Limitations and Need for Additional Research

Limitations of the HSTEP primarily focus around the following four areas:

1. Its reliance upon self-report data.
2. Its need for additional concurrent validation with other outcome indicators.
3. It does not provide training-specific information concerning what content is applied.
4. Its need for utilization with different human service populations.

Probably the most important concern has to do with the tool's reliance upon self-report ratings. This concern questions the validity of tool itself. As stated previously, the HSTEP correlated with the overall score and the eleven factors on the Transfer

Potential Questionnaire. However, concurrent validation with other training effectiveness indicators would increase one's confidence in the tool's value. The lack of other existing "do-able" indicators complicates this problem. However, since the HSTEP is a generic indicator, it can be used easily in conjunction with other common indicators (e.g., completion of action plan objectives), or instruments/methodology designed to assess outcomes of specific training content.

Although HSTEP provides useful information on the extent of transfer, it does not provide information about which specific training content transfers and which does not. It is sometimes essential to know which specific training content is being utilized/not utilized, as well as the quality of the application. Not only is the amount of transfer important, but also the timing of when training content is utilized and when not. New learning, applied in the wrong situations, can have undesirable results. For example, the use of newly learned physical crisis control techniques could be used when verbal de-escalation interventions would be more appropriate. The use of the physical intervention could escalate into a situation where someone might get hurt. Knowledge of application and misapplication of training content is important evaluative information. When training content-specific application information is needed, the HSTEP is of limited value without the use of additional content-specific indicators.

A fourth limitation of the HSTEP is that its usefulness with a variety of human service professionals has not been documented. Although the NEORTC has used it with social workers, child and youth care workers, supervisors, and social service aides in public child welfare settings, it has limited documented usage with other groups and settings (Curry, 1997). (See Figure 2 for a comparison of HSTEP functions and limitations.)

Figure 2. Functions and limitations of the HSTEP

Function	Empirical support	Limitation
Incorporates Kirkpatrick's four levels into a single assessment tool. Can provide information on each of the four levels or can combine items to create a single index.	Results of Principal Components Analysis in NEORTC study indicates the presence of one factor. Supports the single index approach.	Relies on self-ratings. A need for concurrent validation with other training effectiveness indicators.
Provides a common indicator of training effectiveness which can be used to evaluate training across different topics, trainers, and populations. Can also promote research across topics, trainers, and populations.	NEORTC CPS study found significant differences among different workshops. Also, HSTEP correlates with overall transfer potential index on the TPQ as well as eleven transfer factors which comprise the TPQ.	It does not provide training specific information regarding which content transfers and which does not. Also, it has been used with child and youth care workers, supervisors, and case aides at NEORTC. However, there is no documented usage with populations other than public child welfare human service personnel.
Provides information on the process of training. Factors which help or hinder the application of learning.	NEORTC study categorized HSTEP application factors with an 83% inter-rater agreement rate. Concurrent validation between the top transfer factors on the HSTEP and TPQ.	Participants are often not aware of the factors which help or hinder transfer (metacognition).
Brief postcard questionnaire designed to promote high return rate for busy human service professionals with high work loads. Can have more confidence conducting statistical procedures which assume a normal population.	NEORTC study resulted in a 74% return rate for CPS social workers. Those who were less motivated to attend training were also less motivated to return the postcard. However, no overall differences found on the Transfer Potential Questionnaire between those who returned and those who did not return the card.	There may be different return rates for different populations.

Conclusion

The HSTEP questionnaire is a useful tool for assessing human service workers' perception of increased competence on the job as a result of training. The tool is designed to require little completion time from busy human service personnel. Utilization of the questionnaire resulted in an acceptable return rate.

The questionnaire can provide a common indicator which can be used with a wide variety of training topics. Workshops can be compared with each other based upon data collected after the participants have had an opportunity to use the training on the job. In addition to its usefulness as a training effectiveness indicator, it provides information on factors which help or hinder effective training application.

Similar to most reaction evaluations which are usually completed at the end of training, the tool relies upon participant perception. However, it goes beyond level-one reaction evaluation and requires the participant to provide a general assessment of all four levels of Kirkpatrick's training evaluation model, after the participant has had an opportunity to use the training on the job.

Although there is a need for additional validation of the tool (it does not provide training specific information regarding which content transfers and which does not), it appears to be useful as a common outcome indicator of training effectiveness as well as a means to explore factors which influence training application.

References

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Bell, J.D. & Kerr, D.L. (1987). Measuring training results: Key to managerial commitment. Training and Development Journal, 41, 70-73.

Bickerstaffe, G. (1993). Measuring the gains from training. People Management, 25, 48-51.

Birnbrauer, H. (1987). Evaluation techniques that work. Training and Development Journal, 41, 53-55.

Bramely, J. (1991). Evaluating training effectiveness: Translating training into practice. London: McGraw-Hill Co.

Clark, R. & Voogel, A. (1985). Transfer of training principles for instructional design. Educational Communication & Technology Journal, 33, 113-123.

Curry, D.H. (1997). Factors affecting the perceived transfer of learning of child protection social workers. Unpublished doctoral dissertation, Kent State University.

Curry, D.H., Caplan, P. & Knuppel, J. (1994). Transfer of training and adult learning (TOTAL). Journal of Continuing Social Work Education, 6, 8-14.

Garavaglia, P.L. (1993). How to ensure transfer of training. Training & Development, 47, 63-68.

Groth-Marnat, G. (1990). Handbook of psychological assessment. New York: John Wiley & Sons, Inc.

Jones, J.F., Stevenson, K.M., Leung, P. & Cheung, K.M. (1995). Call to competence: Child protective services training and evaluation. Englewood, CO: American Humane Association.

Kirkpatrick, R. (1975). Evaluating training programs. Madison, WI: American Society for Training and Development.

Krein, T.J. & Weldon, K.C. (1994). Making a play for training evaluation. Training & Development, 48, 62-67.

Pines, J. & Tingley, J.C. (1993). ROI of soft skills training. Training, 30, 55-60.

Rackham, N. (1979). The coaching controversy. Training and Development Journal, 39, 12-16.

VanderVen, K. (1990). The international directory of training and education activities in child and youth care work. Washington, DC: Education and Training Committee, National Organization of Child Care Worker Associations, Inc.