A comprehensive model for child welfare training evaluation ☆

Becky F. Antle *, Anita P. Barbee, Michiel A. van Zyl

University of Louisville, Kent School of Social Work, United States

Abstract

The purpose of this research was to develop and test a comprehensive theoretical model for child welfare training evaluation. Drawing upon the theoretical work in training evaluation as well as empirical research in child welfare, this study proposed a model of training evaluation that included individual and organizational predictors of outcomes; training satisfaction, learning and transfer; as well as federally mandated organizational outcomes of safety, permanency, and well-being. The model was tested through an experimental–control group pre- and multiple-post test design with 72 supervisors and 331 case workers in public child welfare. Supervisors and workers in the experimental group participated in a five-day training on skills for effective casework practice and federally mandated outcomes for child welfare. Subjects completed a number of standardized scales to measure the constructs in the model pre-training, immediately post-training, and two months post-training. The data were analyzed using structural equation modeling. Results indicate that individual learning readiness, supervisor support of learning, and knowledge gain are predictive of training transfer. Recommendations are provided to enhance the effectiveness of child welfare training delivery systems.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: Child welfare; Training evaluation; Learning readiness; Supervision

1. Introduction

A review of the child welfare training evaluation literature reveals that the majority of evaluation studies have focused on intermediate outcomes such as the reactions of trainees and increases in knowledge (for example, see Cauble & Thurston, 2000). There have not been a substantial number of research studies on the transfer of training to the workplace or organizational outcomes for training. Yet, the corporate literature clearly indicates that the transfer of skills from training and return of investment for the organization are the most important indicators of training success (Burke, 1997; Shelton & Alliger, 1993).

Child welfare training evaluation has historically been guided by the Kirkpatrick taxonomy of training evaluation. Kirkpatrick (1959) identified four levels at which training should be evaluated: reactions, learning, transfer, and

organizational outcomes. The child welfare training evaluation literature can be organized according to these levels of the Kirkpatrick Model.

2. Reactions

Attitude changes, an increase in sense of competence, and enhanced level of comfort are important reactions to learning in child welfare. Pecora, Delewski, Booth, Haapla, and Kinney (1985) found that training for family services workers resulted in statistically significant positive changes in 23 out of 27 key attitudes. These attitudes were considered to be essential for the implementation of this casework practice. Cauble and Thurston (2000) evaluated the use of child welfare interactive multimedia training with social work baccalaureate students. These researchers identified significant increases in sense of competence in content area and level of comfort in use of skills (reactions). By combining the outcomes in both the reaction and learning categories in this study, they also identified an increase in knowledge of key content areas (learning) and helped us to understand the relationship between different levels of learning outcomes.

3. Learning

The relationship between learning or knowledge gains following training and changes in attitudes or behaviors varies. Beckman and Mays (1985) conducted an evaluation of substance abuse training and found that workers retained knowledge from the substance abuse training at a six-month follow-up assessment, but they had experienced no significant change in attitudes or practice behaviors related to training. In their meta-analysis of the training evaluation literature, Alliger, Tannenbaum, Bennett, Traver, and Shotland (1997) reached a similar conclusion that knowledge gains following training are not necessarily associated with changes in attitudes or behaviors.

McCowan, McGregor, and LoTempio (1989) conducted a competency-based evaluation of child welfare training and found significant gains in knowledge from pre- to post-training. These researchers also identified that workers with less experience in child welfare had greater knowledge gains, which was attributed to differences in motivation for learning. Another key finding was that despite significant gains in knowledge, none of the trainees reached the target or "mastery" level for child welfare practice. This research emphasizes the importance of clinical/practical vs. statistical significant in training evaluation.

Cheung and Stevenson (1991) evaluated training of case management skills for sexual abuse interventions. Although this study focused on skill acquisition, these skills were measured through vignettes, which is more similar to knowledge than true training transfer. The results indicated that there were significant improvements in three out of four skill areas: formulating goals, setting appropriate objectives, and defining the key components for contract negotiation.

Gregoire (1994) conducted a study that combined reactions, learning and transfer outcomes. He found that worker attitudes changed in 10 out of 13 categories, and one-half of workers had complete or partial implementation of practice goals set at the end of training (similar to Delewski & Pecora's, 1986 methodology and Smith, Schinke and Springer (2000) findings).

4. Transfer

There have been several other studies on the transfer of training material to the job in child welfare. Rooney (1988) compared an experimental and control group in their use of task-centered practice. The experimental group received training in this approach, while the control group received no such training. Trained workers showed a significant increase in their use of this technique.

Kessler and Greene (1999) compared the findings of two studies that utilized a case study methodology. The first study used individual training, while the second study used group training. When individual training was used, there was limited transfer of casework skills for supervised visitation in child welfare casework (use of skills ranged from 3–20%). There were higher rates of transfer in the group training approach. However, in both studies, trainees experienced an increase in transfer between pre- and post-tests of training.

A recent study by Curry, McCarragher and Dellman-Jenkins (2005) building on his previous work on training transfer (Curry, 2001, Curry, Caplan, & Knuppel, 1994) attempted to link training transfer to retention in child welfare.

Through this research, transfer support variables were identified. These variables included supervisory support, co-worker support, application planning, and caseload size. The results of this study demonstrated that all variables affected training transfer and subsequently staff retention. Finally, studies by Yankeelov, Barbee, Barber and Fox (2000) and Patterson (2004) conducted experimental studies and found significant differences between control and experimental training groups in empathy, learning and practice behaviors.

5. Organizational impact

Studies on the organizational impact of training are much less common throughout the training literature (Shelton & Alliger, 1993). There is one study in the child welfare training evaluation literature that documents the organizational impact of training. This study by Jones and Biesecker (1980) evaluated training for permanency planning skills. The result of this training was a reduction in the caseload sizes of foster care workers, as well as fiscal savings associated with better permanency planning.

Although there is much value in the previous theoretical and research literature on training evaluation, the present research project proposed a new theoretical model for child welfare training evaluation. There were several rationales for this new model. First, the existing research in child welfare is primarily based upon the work of Kirkpatrick, and there is a need for a more comprehensive model of training evaluation. Second, such a model must consider the unique organizational and practice constraints of child welfare. While many of the variables identified in earlier models may be relevant to child welfare, there are variables that have not been considered and other variables that are not useful. Third, this new theoretical model was based upon the unique outcomes of the field of child welfare. This model combined many of the predictor variables of previous research and theoretical models.

6. Meta-analysis and specification

A significant contribution to our understanding of the Kirkpatrick model was made by Alliger et al. (1997) in their meta-analysis of the training literature. The purpose of this meta-analysis was to assess the degree of relationship between levels of the Kirkpatrick taxonomy. Yet, in this process, the authors greatly increased the specificity of Kirkpatrick's constructs. Alliger et al. (1997) operationalized level one trainee reactions as affective reactions, utility reactions, and combined reactions. Affective reactions refer to whether the trainees liked the training material, while utility reactions refer to whether the trainees perceived training material to be useful for their work.

Alliger et al. (1997) operationalized level two learning outcomes as immediate recall, long-term retention and behavioral demonstration. Immediate recall is measured through a post-test administered immediately following training. Long-term retention is measured through a post-test that is administered at a specified time in the future (e.g. six months following training). Behavioral demonstration refers to the use of role play or some other in-training observation of skills being taught.

These authors did not modify the existing definition of level three evaluation or training transfer. However, they did specify level four organizational outcomes as productivity, customer satisfaction, cost savings and morale.

In addition to the contribution made by their specification of Kirkpatrick's four levels, Alliger et al. (1997) also contributed to the training evaluation literature through the results of their meta-analysis. This meta-analysis found that the correlations within Kirkpatrick's levels are much stronger than the correlations between levels. See Fig. 1 on the next page. Previous research has assumed that there is a linear progression of training outcomes in the Kirkpatrick model—that positive reactions lead to learning; learning leads to transfer; and transfer promotes positive organizational outcomes (Alliger & Janak, 1989). However, the Alliger et al. (1997) analyses do not confirm these assumptions. Instead, these authors found that there are strong relationships between types of reactions (utility and affective) and types of learning (recall, retention, demonstration).

Another finding of this meta-analysis was that the correlation between training outcomes is strongest between variables of similar behavioral specificity (Alliger et al., 1997). For example, there was a strong correlation between utility reactions and transfer. These authors also found that utility reactions are much more predictive of other outcomes than affective reactions. This may be due to the fact that utility reactions force trainees to consider situational constraints (Peters, O'Connor, Eulberg, & Watson, 1988). This is a particularly informative outcome since much training evaluation research focuses on trainee reactions. Despite these numerous contributions to the training
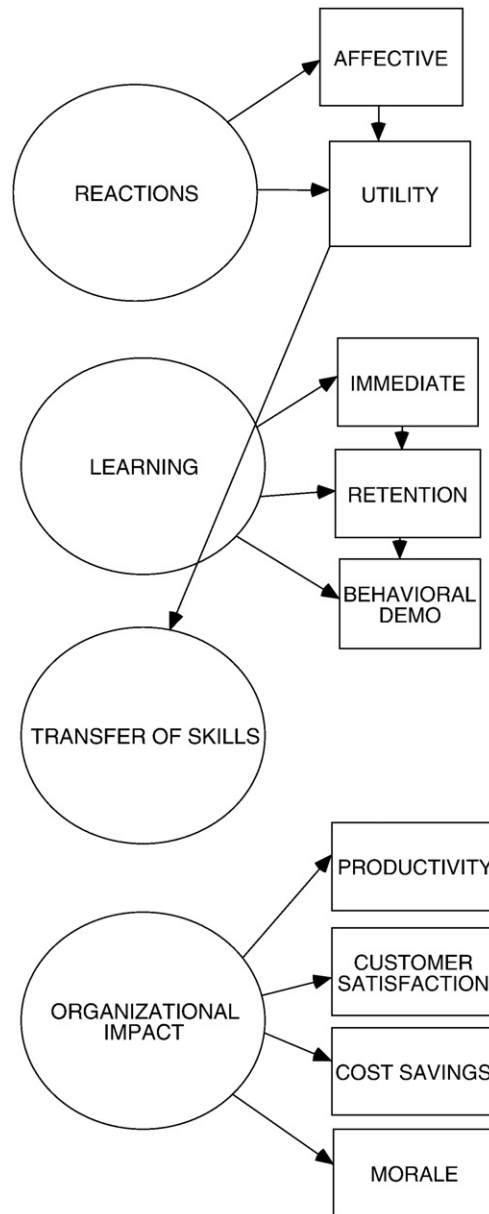
Fig. 1. Alliger et al. (1997) specification of Kirkpatrick's (1959) model and relationships.

evaluation research base, the work of Alliger et al. (1997) is also susceptible to the criticisms of Kirkpatrick, as well as objections to meta-analysis in general.

7. Child welfare training evaluation: theoretical model

Given the limitations of previous training evaluation in child welfare, the purpose of this research was to evaluate the effectiveness of training for child welfare supervisors in the state of Kentucky. This training targeted supervisors because of the critical role these leaders play in promoting positive casework outcomes (ASFA, 1997; Moore, Rapp & Roberts, 2000). Supervisors must monitor and reinforce the practice skills of caseworkers that are directly linked to such outcomes. This research evaluated the effectiveness of training for the acquisition of knowledge and transfer of supervisor skills such as social support, reinforcement of skills and evaluation of worker practice. This research also

evaluated the impact of supervisor training on the national child welfare organizational outcome of child safety (ASFA, 1997).

The first aim was to develop a comprehensive model for supervisor training evaluation in child welfare. The second aim of this research was to test the appropriateness and predictive capability of this training evaluation model for child welfare supervisors. The appropriateness of the model was tested by identifying the strength of relationships between variables—whether variables identified in the model were truly relevant to training outcomes. The predictive capability of the model was assessed by linking predictor variables to training and organizational outcomes to establish the degree to which the training and associated predictor variables predicted positive outcomes.

The key assumption of this new model for child welfare training evaluation is that there are predictor variables, such as individual and organizational characteristics, that predict training outcomes, such as trainee reactions, learning, and transfer. These training outcomes predict larger organizational outcomes such as child safety, permanency, and well-being, as the use of key skills from training promotes best practice. See Fig. 2 for the full Louisville Child Welfare Training Evaluation Theoretical Model.

7.1. Predictor variables

Due to sample size constraints and the need to test certain pathways of prediction, a more focused version the model was tested in this research and is described below. See Fig. 3 for the Louisville Child Welfare Training Evaluation Tested Model. There are three units of analysis for the predictors of child welfare training outcomes that were studied in the present research: individual learners, teams and organizations. Within the first level of analysis, there are the following constructs: learning readiness and personality type. Previous research by Ford, Quinones, Sego and Sorra (1992) identified the importance of the individual's learning readiness for training outcomes such as reactions and
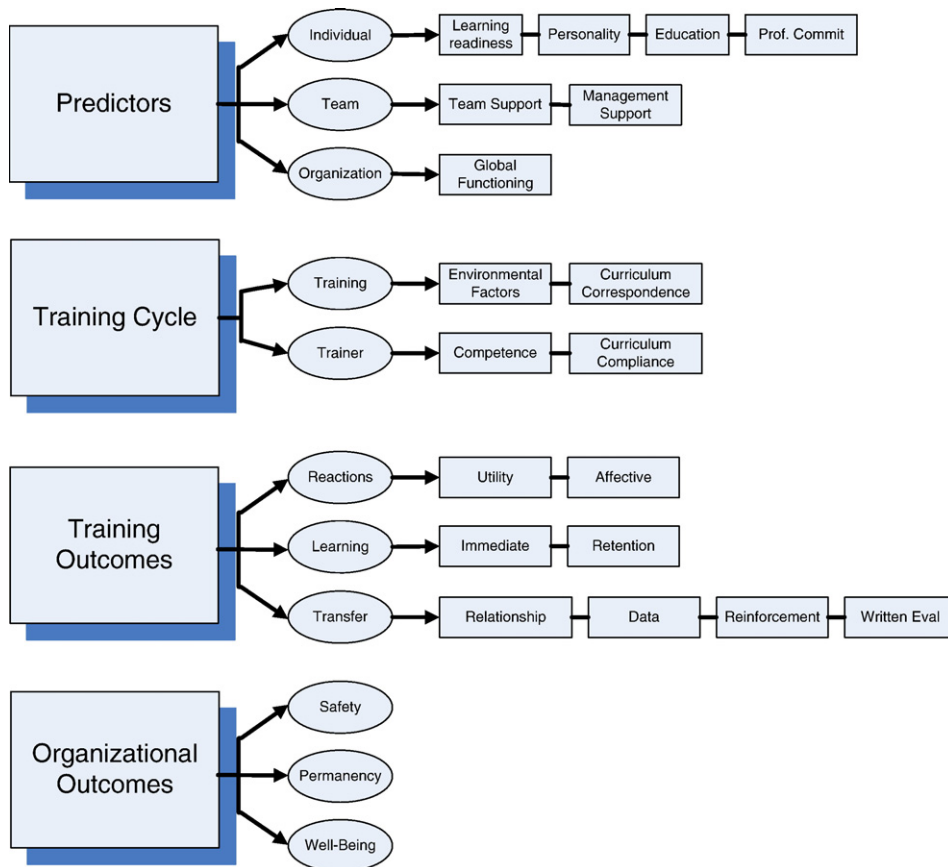

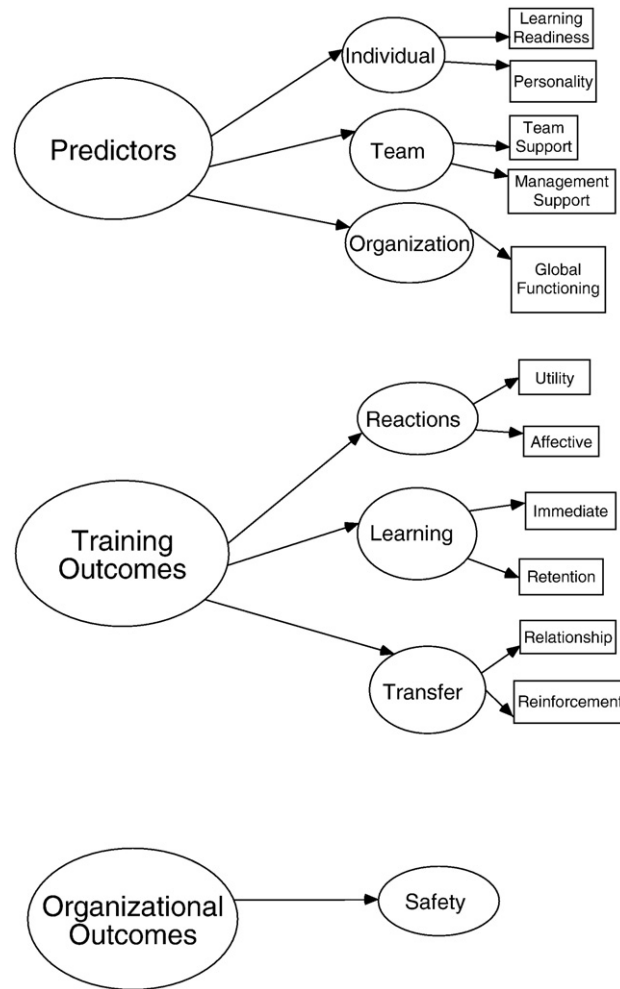Fig. 2. Louisville Child Welfare Training Evaluation Theoretical Model.

Fig. 3. Antle Child Welfare Training Evaluation Tested Model.

learning. Research by Barbee, Bledsoe, Antle and Yankeelov (1999) found that there are significant relationships between the education level, personality type and job satisfaction of workers and transfer of training. For example, workers who have the personality trait of conscientiousness (on the Big Five Personality Inventory) are significantly more likely to transfer training to the job.

Within the second level of analysis, teams, there are the following constructs: team attitude and management support (Ford et al., 1992). Team attitude may be measured by the team's reaction to training material. Management support for child welfare supervisors is provided by regional administrators.

For the final level of analysis, organizations, the primary construct is organizational support. Organizational support includes organizational cohesion, policy and procedure concordance, and other dimensions measured by the Global Scale of Organizational Functioning (Coetsee, 1998). An organizational culture that supports learning and outcome achievement is essential to quality child welfare practice (Moore et al., 2000).

7.2. Training outcomes

The training outcomes included in the Louisville Child Welfare Training Evaluation Model are trainee reactions, learning, and transfer. Trainee reactions are classified as utility reactions (is training material useful) and affective reactions (do trainees like the material) (Alliger et al., 1997). Changing attitudes is an important precursor to changing practice (Pecora et al., 1985). Learning may be categorized as immediate, retention, and behavioral demonstration

(Alliger et al., 1997). Although behavioral demonstration is part of the theoretical model (Fig. 2), this research only measured immediate learning and retention in the tested model (Fig. 3).

Finally, transfer refers to the use of new skills on the job. The transfer of training to the workplace is a key consideration in evaluating the effectiveness of training (Delewski & Pecora, 1986). There are a number of skills identified from the supervisor training curriculum that can be tracked for transfer on the job, including supervisor relationship and casework reinforcement.

7.3. Organizational outcomes

There are organizational outcomes for child welfare agencies that have been identified by federal policy (ASFA, 1997). These outcomes include child safety, permanency, and well-being. Child safety is defined as the absence of subsequent incidents of child maltreatment, while permanency refers to the length of time in out of home care or number of placements. Child well-being is generally measured through standardized scales that assess functioning across domains. While supervisors often do not provide the direct services that may promote these positive outcomes, they are responsible for ensuring that workers utilize appropriate skills for these outcomes. These outcomes may be classified as the organizational impact of training. Organizational outcomes were evaluated through a separate study and will be reported elsewhere.

8. Research questions

The primary set of research questions centered on the relationship between the above predictor variables and training outcomes for child welfare training. Specific questions included the following:

(1) What was the relationship between individual variables such as learning readiness and training outcomes such as learning and transfer?
(2) What was the relationship between team support and training outcomes?
(3) What was the relationship between the organizational climate and reactions, learning, or transfer?
(4) What was the relationship between training outcomes of reactions, learning, and transfer?

The current research will address the gaps in the literature in three domains: training outcomes at all levels, evaluation of supervisor performance, and organizational outcomes. Although there are a few studies on the evaluation of training in child welfare, the majority of the child welfare training literature has focused on needs assessment and methodology. Miller and Dore (1991) explain that training evaluation is often not conducted in child welfare because of limited resources. Child welfare agencies tend to direct these resources toward new training instead of the evaluation of existing training. The current research will address the gaps in the child welfare training evaluation literature by evaluating training outcomes at multiple levels. Following supervisor training, data will be collected on attitudes, learning, and transfer of skills associated with training.

The second way that this research will add to the child welfare literature is through the evaluation of supervisor training outcomes. There has been a needs assessment study on supervisory skills (Denning & Verschelden, 1993), as well as an informal study on the use of a reporting package for supervisor management of outcome data. However, there has not yet been any formal evaluation of supervisor training. The current research will focus on the evaluation of child welfare supervisor training.

9. Methodology

9.1. Research design

The study was conducted in the state of Kentucky and utilized a pre-post experimental–control group design. The experimental group in this research participated in the first wave of supervisory training. The control group participated in the second wave of supervisory training, which took place one year later. Hence, this design utilized the "waiting list" version of the control group (Rubin & Babbie, 1997). The experimental and control groups were compared in their knowledge, use of supervisory skills, and the organizational outcome of child safety during the same time period. The

collection of pre–post-training data allowed for analysis of change in key knowledge, skill and organizational outcome areas.

9.2. Training intervention

The training intervention consisted of five days of training. Days one and five contained content for supervisors only, including the parallel process of supervision, coaching and mentoring, and supervision to promote ASFA outcomes. Days two through four of the training were provided to both supervisors and their workers. These days covered material on Kentucky's child welfare practice model, Solution-Based Casework (Christensen, Todahl, & Barrett, 1999), as well as practice skills for assessment, case planning, ongoing case management, and work with community resources. This training was provided through a federally funded grant to enhance casework skills to promote positive ASFA outcomes of child safety, permanency, and well-being. There were two core trainers who trained the vast majority of the material for all cohorts. The fidelity of the training was assessed through live observation and videotaping of training events. This live observation and video tape was reviewed for trainer compliance with the curriculum (fidelity of the intervention).

9.3. Sample

There were 72 supervisors sampled for this study. There were 42 supervisors in the experimental group and 30 supervisors in the control group. These supervisors were drawn from six geographically representative regions: three urban and three rural regions. Several variables in this theoretical model required data from the actual caseworkers (e.g. reinforcement of casework practice, supervisor–worker relationship). In these cases, all members of the supervisor's team were included in the study. There were 195 experimental group caseworkers and 136 control group caseworkers. Supervisors and their teams across the state were encouraged to participate in this training in order to prepare for Kentucky's Child and Family Service Review, and there was strong administrative support for this project. However, participation was voluntary. The subsequent sample was representative of the urban and rural regions across the state, each of which has unique service delivery and resource dynamics.

9.4. Variables and instruments

9.4.1. Predictor variables

9.4.1.1. Individual.   All predictor variables were measured pre-training only. The first individual variable measured was *learning readiness*. Learning readiness incorporates such concepts as transfer of learning skills, use of feedback, learning as a life skill, support for learning, and self-directedness in learning (van Zyl & van Zyl, 2000). Learning readiness was measured using the Learning Benefit Inventory, developed and validated by van Zyl and van Zyl in the 2000 study. This scale contains 70 items to which subjects respond on five-point Likert scales, ranging from none of the time to all of the time. The internal consistency reliability of the scale was determined to be satisfactory, with the Cronbach alpha scores of factors or sub-scales ranging from 0.75 to 0.89. The construct validity (unidimensionality) of the scale was established using structural equation modeling. Four of the six factors were able to be combined into a super factor (Bentler–Bonnet nornnormed fit index = 1.002, comparative fit index = 1.000), supporting the measurement of a single construct (learning readiness) by the scale.

The second individual variable was *personality*. The short version of the Big Five Questionnaire was utilized to measure personality. The five personality traits that are measured by this scale include extraversion, conscientiousness, openness to experience, agreeableness, and neuroticism (Caprara, Barbaranelli, Borgogni, & Perugini, 1993). Although only the conscientiousness sub-scale of this inventory was included in the research model, all five sub-scales were administered. This allowed for embedding the target sub-scale within the other sub-scales and maintain reliability/validity associated with this instrument. This short version contains 40 adjectives, and respondents use five-point Likert scales to rate the degree to which that adjective accurately describes them. The reliability and validity of this scale were established through large-scale studies comparing Italians and Americans. The average weighted mean coefficient for the five sub-scales is 0.75 (Viswesvaran & Oanes, 2000). Construct validity of this scale was supported through high correlations with the similar NEO-PI scale (Barbaranelli, Caprara, & Maslach, 1997).

9.4.1.2. Team.   Team variables included both team support and supervisor support of the supervisor. Team support was measured using the Team Learning Conditions sub-scale of the Training Transfer Inventory (Coetsee, 1998). The Training Transfer Inventory was validated by Coetsee through a study of 2810 mine workers, supervisors, and managers in South Africa. The theoretical foundations and previous scales upon which this instrument was based include the works of Rouiller and Goldstein (1993) and Deschant and Marsick (1993).

The Team Learning Conditions sub-scale measures the degree to which the team is open to or supportive of new information and processes in the workplace. This scale contains 30 items to which subjects respond on five-point Likert scales, ranging from strongly disagree to strongly agree. This sub-scale has acceptable internal consistency reliability, with a Cronbach alpha of 0.778. The construct validity of the sub-scale was tested using structural equation modeling. The items on this sub-scale emerged as a single factor and were not able to be combined with any other sub-scales. This sub-scale was highly correlated with an associated sub-scale measuring Team Learning Processes. However, these sub-scales emerged as independent factors in the analysis.

Supervisor support of training/learning was measured using the Supervisor Sub-Scale of the Training Transfer Inventory (Coetsee, 1998). This sub-scale measures the degree to which supervisors support new learning or training material. Supervisors completed this scale in reference to their immediate managers, the Service Region Administrative Assistants (SRAAs). This sub-scale contains 14 items to which respondents rate degree of agreement on five-point Likert scales. The internal consistency reliability of the scale is acceptable, with a Cronbach alpha score of 0.884. Construct validity was supported through structural equation modeling analysis in which this sub-scale emerged as an independent factor.

9.4.1.3. Organization.   Organizational support was measured using the Organizational Learning Conditions and Support Sub-Scale of the Training Transfer Inventory (Coetsee, 1998). This sub-scale assesses the degree to which the organization supports or maintains an environment of learning. The scale contains 17 items to which subjects respond on five-point Likert scales, ranging from strongly disagree to strongly agree. This sub-scale has acceptable internal consistency reliability, with a Cronbach alpha of 0.789. The construct validity of the sub-scale was tested using structural equation modeling. The items on this sub-scale emerged as a single factor and were not able to be combined with any other sub-scales.

9.4.2. Training outcomes

Training outcomes were measured pre- and post-training. Trainee reactions were measured only post-training. Learning of training material and transfer of skills/behaviors from training were measured both pre- and post-training.

9.4.2.1. Reactions.   Trainee reactions were measured along two dimensions: utility and affective reactions. Utility reactions refer to the degree to which trainees find the training material useful. Affective reactions refer to the degree to which trainees like the training. Both of these reactions were measured using a scale adapted for this study—the Level One Training Evaluation Scale. This scale contains 12 items. For each item, respondents indicated their degree of agreement on five-point Likert scales. A similar scale was previously used for the evaluation of substance abuse training in child welfare (Barbee & Barber, 1995).

9.4.2.2. Learning.   Learning was measured using a test of the training curriculum. This knowledge-based test was developed specifically for this research and consisted of both multiple choice questions and open-ended questions. There were 39 items on this test that covered material from each of the key content areas of the training.

9.4.2.3. Transfer.   There were two primary skill areas that were assessed for transfer following training. The first skill area was the supervisor–worker relationship. The supervisor–worker relationship was measured with the Consideration sub-scale of the Supervisory Behavior Description Questionnaire (Fleishman, 1957). The Supervisory Behavior Description Questionnaire (SBDQ) contains two sub-scales: consideration and initiating structure. However, the psychometric properties of the second sub-scale are much weaker than the first. Also, the focus of this study is on supervisory support, which more closely matches the items on the Consideration sub-scale. There have been a number of studies on the reliability and validity of this scale. One study by Szilagyi and Sims (1974) found that the internal consistency reliability coefficients were very high (0.93). Furthermore, the construct validity of the sub-scales was confirmed through factor analysis, which found factor congruency indices (phi-coefficients) of 0.95. This scale contains 28 items to which respondents rate degree of agreement on five-point Likert scales.

The second skill area was supervisor feedback. A sub-scale of the Training Transfer Inventory was developed that assessed the degree of reinforcement or demonstration of specific skills/behaviors from the training. A similar procedure was used by Coetsee (1998) when he added a sub-scale based upon specific goals and curriculum from the training being evaluated.

9.5. Procedure

Per the requirements of the Institutional Review Board that reviewed this research, all participants in the research (supervisors and workers) completed full informed consent forms prior to the completion of these surveys. The pre-training measures of predictor variables and knowledge of training content were distributed to supervisors in the experimental group on the first day of supervisor training. Supervisors in the control group received these measured from a Regional Administrator responsible for delivery of tests. The pre-training measures of the supervisor–worker relationship and reinforcement/demonstration of curriculum-based skills were administered to workers in the experimental group on the first day of team training. These measures were delivered to workers in the control group at the same time. The response rate for pre-training measures were as follows: 73% experimental group supervisors, 95% experimental group workers, 81% control group supervisors, and 88% control group workers.

The post-training measures of training outcomes were collected at several points in time. First, the experimental group supervisor reactions to training and immediate learning of material were measured on the final day of training. Regional Administrators delivered learning tests to control group supervisors at the same time. The retention of material learned in training was measured one month after the completion of training. Tests were mailed to both experimental and control group supervisors. Other transfer variables, such as supportive supervisor relationship and supervisor reinforcement, were measured at this one-month time period to allow time for transfer. These tests were also mailed to experimental and control group workers. The response rate for immediate post-training measures were as follows: 78% for experimental group supervisors, 40% for experimental group workers, 58% for control group supervisors, and 22% for control group workers. The response rate for one-month post-training measures were as follows: 41% for experimental group supervisors, 31% for control group supervisors.

10. Results

10.1. Criteria for modeling

10.1.1. Criteria for fitness

Structural equation modeling involves two steps: validating the measurement model and fitting the structural model. The former is accomplished primarily through confirmatory factor analysis, while the latter is accomplished primarily through path analysis with latent variables. Three criteria were utilized to evaluate the "fitness" of these models through structural equation modeling (Tabachnick & Fidell, 2001). These criteria are used for research with small sample sizes. The first criterion is the chi-square statistic. This statistic should not be significant, and the data must meet the assumptions of the chi-square test (no less than five cases per cell). The second criterion is the ratio of the chi-square value to the degrees of freedom. This ratio should be less than 2:1. Finally, the best fitness measure available given the constraints of the data should be utilized. Given the sample size, that fitness measure is the Yuan-Bentler Fit Index. (This index is recommended for sample sizes of 60–120). For each of the following models tested, only the first two criteria could be applied.

In accordance with the first criterion (chi-square assumptions), there were no less than five cases per cell. The significance of the chi-square statistic will be reported for each model. Data for the second criterion, chi-square to degrees of freedom ratio, will also be reported for each model. The Yuan-Bentler Fit Index could not be calculated due to the missing data in this data set. The Yuan-Bentler Fit Index utilizes the ADF (asymptotically distribution free) estimator, but this estimator cannot be used if there is missing data. There was missing data due to the failure of some participants to complete all tests. In order to account for missing data in the structural equation modeling process, means and intercepts were estimated.

10.1.2. Numbers of variables

Due to the small sample size for supervisors ($N = 72$), an effort was also made to limit the number of variables in the models to three. The rationale for this number of variables was derived from the general rule regarding the ratio of the

number of parameters estimated to the sample size (Tabachnick & Fidell, 2001). For large samples, this ratio of number of parameters estimated to sample size should be 10:1. However, for small samples, the general ratio is 7:1. With the sample size of 72 and three variables in the model, the number of parameters to be estimated was eight. This maintained a ratio of sample size to parameters estimated of 9:1.

10.1.3. Marker variables

The concept of "marker variables" is derived from factor analysis and is commonly used in multivariate statistics (Tabachnick & Fidell, 2001). A marker variable is a pure measure of a factor. The marker variables for these models were selected based upon their representativeness as single factors within total domains and their high reliability scores. The marker variables selected were learning readiness and management support of training. Learning readiness represents the domain of individual characteristics. The alpha levels for the sub-scales of the Learning Benefit Inventory were as follows: Life Skills, 0.94; Self-Direction, 0.78; Support, 0.83. Management support of training represents the training support domain of the model. The alpha level for the Management Training Support Sub-Scale of the Training Transfer Inventory was 0.93. See Table 1 for reliability analysis on all scales. The following section summarizes additional reasons for the exclusion of other variables from the models.

10.1.4. Excluded variables

10.1.4.1. Predictor variables. The variables of conscientiousness and Team Learning Conditions were not included as marker variables due to very low reliability for these sub-scales (Conscientiousness: alpha=0.64; Team Learning Conditions: alpha=0.65). Organizational Learning Conditions was not included because it was determined that there was not enough variability in this data for the given group of supervisors. In the initial planning of this research, teams were selected from different regions of the state in order to provide greater variability at the organizational level. The relationship between supervisors/teams and the organization vary by region. However, these differences in relationships did not result in different ratings of Organizational Learning Conditions.

10.1.4.2. Utility reactions. Although there was a significant positive correlation between utility reactions to training and learning, r (22)=51, $p<.05$, reactions were not included in this model. The mean rating on this Level 1 Utility Reactions sub-scale was 3.01. See Fig. 4 on the next page for a frequency distribution of average scores across all 12 items of this Level 1 Reactions scale. Trainee reactions to training tend to be high, but these scores are average. This mean rating indicates that trainee reactions did not inflate scores for the remaining variables in the model. These models also focused on other variables because the training evaluation literature has already established the relationship between trainee reactions and learning (Alliger et al., 1997). A final reason that trainee reactions were not included in the models was that reaction data was only collected for the experimental group. Therefore, the inclusion of this variable would have limited the sample size to 41 for the modeling process.

The first model that was attempted included the organizational outcome variable of child safety. This variable was measured through a recidivism report generated by the state. However, there were no successful models generated to explain this outcome variable. Regression analysis confirmed that there were no significant predictor variables for the safety variable. Learner characteristics of conscientiousness and learning readiness and organizational support of

Table 1
Reliability analysis for all scales

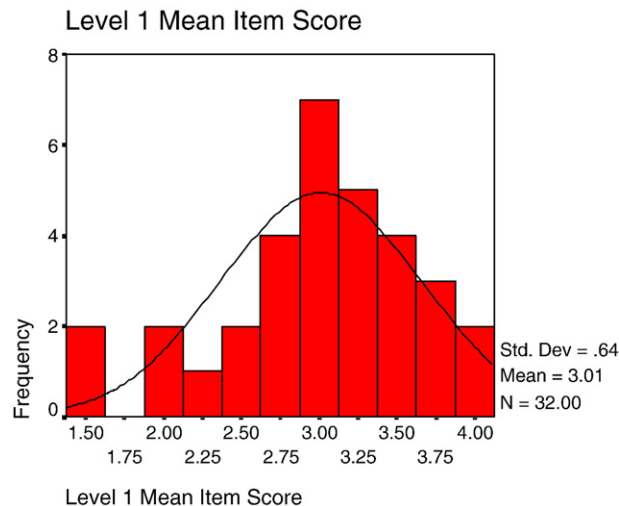| Scale	Alpha coefficient
Learning Benefit Inventory	0.93
Big 5 Conscientiousness	0.64
Team Learning Conditions	0.65
Organization Learning Conditions	0.87
Supervisor Training Support	0.93
Level 1 Utility	0.75
Level 1 Affective	0.87
TTI Curriculum Sub-Scale	0.94
SBDQ	0.91

Fig. 4. Distribution of average supervisor reaction scores.

training each explained only 1% of the variance in child safety. Reactions to training and learning following training each explained 7% of the variance in safety, while transfer of learning explained only 0.4%. Numerous tests were performed to confirm the reliability of the scales being used as predictors in these models. All predictors included had very high reliability in this study and previously established validity. Given the limited relationship between these psychometrically sound measures and the child safety data, the decision was made to focus subsequent models on the training outcomes of learning and transfer instead of this organizational outcome variable of child safety.

Finally, the variable of supervisor reinforcement of casework skills was used as the outcome in the transfer models. The other transfer variable, supervisor relationship skills, was excluded from the model due to training issues. Although the training curriculum included components on the supervisor–worker relationship, the primary emphasis of the training was casework skills. Therefore, it was determined that reinforcement of casework skills would be a better measure of supervisor training transfer.

10.2. Modeling

Five models were generated and tested according to the aforementioned criteria. One model was developed for immediate learning and one for retention of learning. Learning readiness and management support of training were the predictor variables for each of these outcomes. Three models were developed for transfer, with two of the following predictor variables in each: learning readiness, management support of training and immediate learning. The two models developed for learning did not meet the criteria for fitness. Similarly, the transfer model that included the predictors of learning readiness and management support of training did not meet these criteria. However, both of the transfer models that included the predictor of immediate learning did meet these criteria for a sound model.

10.2.1. Learning

10.2.1.1. Immediate learning. The model developed for immediate learning included the predictor variables of learning readiness (LBI total score) and management training support. The outcome variable was immediate learning, as measured by the Supervisor Level 2 Knowledge Test completed pre-training and immediately post-training. There was a significant change in knowledge (learning) between pre-training and immediately post-training, $F(1,31)=5.66$, $p<.05$. The mean total score on the Supervisor Level 2 Knowledge Test pre-training was 31.04 (SD=2.58, Range=25 to 37), and the mean on this test immediately post-training was 32.30 (SD=3.10, Range=25 to 37) (These knowledge test scores are reported as raw versus percentage scores.). A difference score was computed and entered into the model. This model did not meet the criteria for a sound model (χ^2 (1)=3.52, $p=.06$; and ratio of the chi-square to degrees of freedom is 3.52:1.
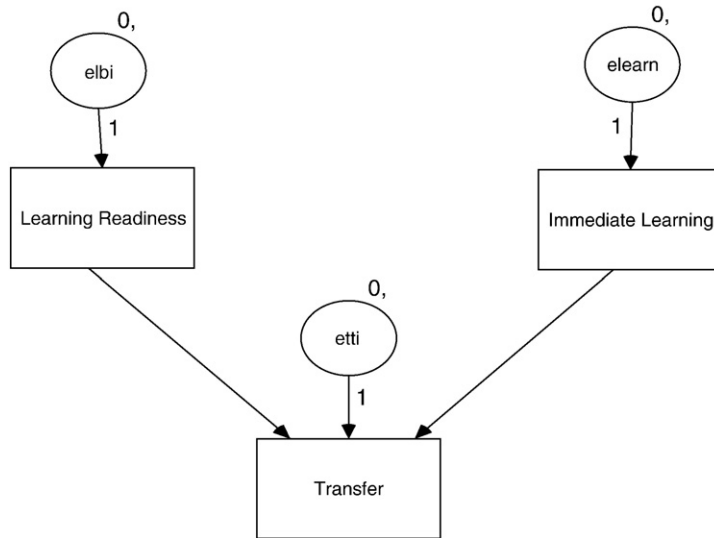
Fig. 5. Learning Readiness-Learning-Transfer Model.

10.2.1.2. Retention. A second model was developed to explain the training outcome of retention. Again, the marker variables of learning readiness and management training support were entered into the model as predictors. The outcome variable was retention of learned material, as measured by the Supervisor Level 2 Knowledge Test completed pre-training and one-month post-training. A difference score was entered into the model. This model did not meet the criteria of plausibility (χ^2 (1)=3.52, p=06 and the ratio of the chi-square to degrees of freedom is 3.52:1).

10.2.2. Transfer models

10.2.2.1. Predictors to transfer. First, a model was developed that used the marker variables of learning readiness and management training support as predictors. The outcome variable was transfer, as measured by the worker ratings of supervisor reinforcement of casework skills covered in training. Workers evaluated their supervisors' reinforcement of casework skills pre-training and one-month post-training. A difference score was used, and data was aggregated across workers in order to maintain supervisors as the unit of analysis. A second training transfer variable was included, focusing on the supervisor–worker relationship and any changes in those skills related to training. The overall focus remained the supervisor reinforcement of casework skills because this was the primary emphasis of the training curriculum. This model did not meet the criteria for a sound model (χ^2 (1)=3.52, p=.06; and the ratio of the chi-square to degrees of freedom is 3.52:1).

10.2.2.2. Predictors and learning to transfer. A second model was generated for transfer that included learning as a predictor variable. This research was interested in whether learning of training material impacts the transfer of trained skills. Therefore, immediate learning (pre-training to immediate post-training) was included as a predictor. Learning readiness was also included as a predictor. The outcome variable was the supervisor reinforcement of casework skills (difference score). See Fig. 5 for the Learning Readiness-Learning-Transfer Model. The "elbi," "elearn," and "etti" circles refer to the error variables for the measures of each construct in the model (e.g. "elbi" refers to error for the Learning Benefit Inventory).

Table 2
Maximum likelihood estimates for Learning Readiness-Learning-Transfer Model

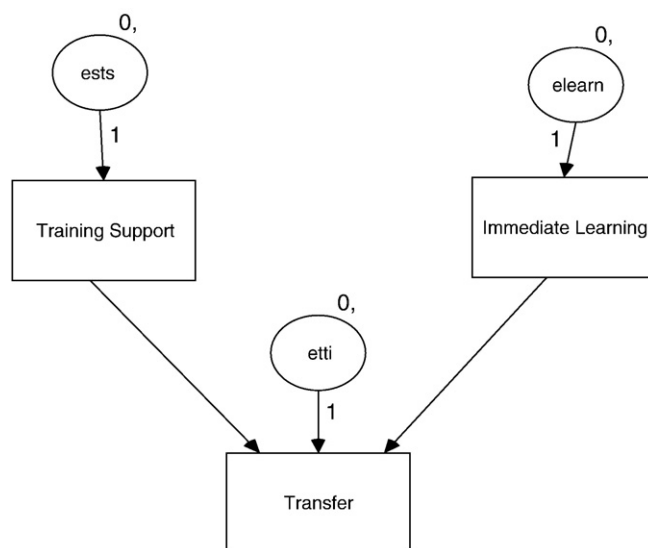| Regression weights	Estimate	Standard error	Critical ratio
Learning Readiness-Transfer	−0.038	0.132	−0.287
Learning-Transfer	−0.679	0.941	−0.722

Fig. 6. Training Support-Learning-Transfer Model.

This model did meet the criteria for a sound model, χ^2 (1)=0.426, p=0.514. This model meets criterion one because the chi-square statistic is not significant, and there are adequate numbers of cases in the cells. This model meets criterion two because the ratio of the chi-square to degrees of freedom is 0.4:1 (less than 2:1). See Table 2 for maximum likelihood estimates for this model.

A final model was developed for transfer that included learning and management training support as predictors, and supervisor reinforcement of casework skills as the outcome variable. See Fig. 6 for the Training Support-Learning-Transfer Model. This model did meet the criteria for a sound model, χ^2 (1)=0.406, p=0.524. This model meets criterion one because the chi-square statistic is not significant, and there are adequate numbers of cases in the cells.

This model meets criterion two because the ratio of the chi-square to degrees of freedom is 0.4:1 (less than 2:1). See Table 3 for maximum likelihood estimates for this model. Although both learning readiness and training support were predictive of transfer with learning, these variables were not entered into a model together due to the constraints of sample size in this study.

11. Discussion

There were several significant findings from this research that contribute to the child welfare training evaluation literature. Alliger et al. (1997) suggested that utility reactions were a better predictor of transfer than actual learning due to similar levels of behavior specificity. On the contrary, this research found that immediate learning is predictive of transfer. Immediate learning was a predictor in both of the transfer models. Learning was predictive of transfer in conjunction with learning readiness and management support of training.

Another major finding of this research is that the individual trainee characteristic of learning readiness predicts transfer of training material. Supervisors who viewed learning as more important and were open to learning new material were more likely to use and reinforce the practice skills from this training. This finding is consistent with the model of Holton (1996) that includes motivation to learn as a mediator of training outcomes, as well as the research of Durr, Guglielmino and Guglielmino (1996) on learning readiness of workers. Ford et al. (1992) also suggest that

Table 3
Maximum likelihood estimates for Training Support-Learning-Transfer Model

Regression weights	Estimate	Standard error	Critical ratio
STS-Transfer	0.060	0.171	0.351
Learning-Transfer	−0.766	0.937	−0.818

individual characteristics such as ability and self-efficacy impact transfer of training. The identification of differences in learning readiness in the child welfare organization was significant given the rigid organizational structure (Hopkins, Mudrick, & Rudolph, 1999). Child welfare organizations change so rapidly that people are often not open to learning because they are overwhelmed by this frequent change (Gleeson & Philbin, 1996).

Another significant finding was that organizational support of training (in this research, management support of supervisors) predicts learning and transfer. This finding builds on that of Ford et al. (1992), who reported that supervisor attitudes toward trainees were one of the strongest predictors of transfer. Numerous theoreticians in the field of training evaluation have emphasized the importance of organizational support to maximize the impact of training. Gleeson and Philbin (1996) asserted that supervisors need organizational support for innovation, and Xiao (1996) identified that organizational factors such as modern management concepts facilitate change in knowledge and skills following training. Alvarez, Salas, and Garofano (2004) found that 10 variables consistently influence training outcomes including trainee variables such as self-efficacy, experience, and post-training mastery orientation, training variables such as training performance, and post-training variables such as learning and post-training interventions. These findings are in line with our theoretical model and the results of this study.

11.1. Limitations

One primary limitation of this research is the above failure of the organizational outcome data and subsequent requirement that only training outcomes be used in the models. It was the hope of this research to link training interventions and outcomes to organizational outcomes in order to address the gap in the literature in this area. However, due to the anticipated problems in isolating the effects of training on casework, as well as the possible systems problems, this relationship could not be established. A related problem with measurement was the small difference detected between pre- and post-training knowledge test scores. This small difference may be attributed to the difficulty of the test and the need to develop instruments that are more sensitive to change. Despite this limitation, knowledge change scores could still be used in modeling processes.

A second area of limitation was the research design and lack of random assignment for a true experimental design. The funding of this project required that the research team reach as many child welfare teams as possible. This goal was strengthened further by the emphasis of this grant on preparing child welfare supervisors and workers for the upcoming federal review and promoting positive outcomes through sound casework practices. Other barriers to random assignment to conditions in child welfare training research include the general difficulty with engaging participants in new training initiatives (as child welfare workers are constantly exposed to such initiatives), as well as the bureaucratic nature of the state child welfare system. The quasi-experimental design utilized in this study was adequate, as the control group was matched to the experimental group along a number of key dimensions.

Another issue related to the research design was the length of time between training and post-training measurement. This organizational outcome data was collected for the two months pre-training and two months post-training for each group. However, this time period may not have been sufficient for the impact of training to be detected at the organizational level. Similarly, the one-month time periods to post-training data collection may not have been sufficient to detect changes in learning and transfer.

In addition to this limitation of length of time to measurement, there were also logistical problems related to the training. Due to the organizational constraints on scheduling training and the different numbers of teams participating in each region, this training was provided over different time periods for each region. For example, in some regions the training was completed in three weeks, while in other regions the training was completed in six weeks. The number of days of training and the training content remained the same for all groups, but the length of time between training days varied. This difference in the length of time in which training was provided resulted in differences in the length of time between pre- and post-tests for these groups. The post-tests were always administered on the last day of training and one month following training, but for some regions, these time periods were much farther removed from the pre-training measures. Although tremendous efforts were made to overcome these problems, the large child welfare bureaucracy and limited control of the researchers resulted in these limitations. The magnitude of these problems was not such that it jeopardized the study in totality.

Another limitation was the relatively small sample size. Although there were 72 supervisors and 331 workers involved in this study, there were gaps in the data that presented problems in data analysis. The aforementioned numbers represent the total number of persons who completed *any* evaluation measure in the research. However, there

were many more respondents for the pre-test measures than the post-test measures. The response rate for post-tests ranged from 41%–95% for the experimental group and 22%–88% for the control group (depending on the measure). In addition, some respondents completed only the pre-tests, while others completed only the post-tests. An effort was made to gather this missing data by adding the incentive of a free lunch for teams who completed the training. Although this effort resulted in a higher response rate than originally received, missing data and the subsequent sample size remained challenging for data analysis. Similar problems with obtaining post-training data were experienced by Highsmith and Illian (2005). The successful completion of the present study can be attributed to the large sample size and multiple data sources (supervisors, case workers, and administrative data).

The final major limitation of this research was the intervention itself. There were many other trainings that teams could have participated in before, during or after this training that may have impacted outcomes. The learning organization theories postulate that learning is an ongoing process and not limited to a specific time period (e.g. this training and research process; Chawla & Renesch, 1995). Furthermore, the training curriculum was refined during the research process and new trainers were added to the two core trainers. Although core content and competencies remained the same, training exercises and activities were modified. The short duration of this training (five days) also raises questions about the strength of the intervention.

11.2. Future research

One way that future research could address the aforementioned problems with isolating the effect of the training is to perform repeated measures panel studies to capture the continuous flow of learning. These repeated measures could detect changes in participant knowledge and skills over time and relate these changes to the numerous influences on learning (including training, policy change, formal education, etc.). This method is particularly relevant to child welfare because workers in this field are constantly experiencing new training, practice innovations and policy changes that require learning and skill transfer.

Also, a number of slight modifications to the design of the present research could be useful. For example, this research measured transfer of training through worker evaluations of supervisors. This method was preferable to supervisor self-report, but perhaps the best method would be direct observation of supervisor behaviors. Another modification could be the manner in which control group teams are matched to experimental group teams. These controls should be matched along key characteristics such as urbanicity and resource availability but geographically distant in order to limit potential contamination. Lastly, future research could allow a longer time period between training and post-training measures. The current research allowed one month for retention of learning and transfer of skills and two months for organizational outcome variables.

11.3. Implications for practice

There are four primary implications of the present research. One implication is the importance of explaining the utility or relevance of training to trainees. This research found that higher utility reactions (perceptions that trained material was useful) were related to greater knowledge gain (learning). Learning was, in turn, related to training transfer. Therefore, one step that trainers or child welfare organizations may be able to take to promote positive training outcomes is to enhance trainee perceptions of the utility of training. This is particularly true for child welfare because of the frequency of training and high levels of training burnout in this field.

Another implication of this research is the application of learning readiness data to training practice. This research found that learning readiness is related to higher levels of training transfer. One future training approach might be to identify learners who are ready to learn (through a learning readiness scale) and provide training to these individuals first. Then efforts can be made to facilitate learning readiness in others before they are trained. This offers a more cost effective approach, as those who are most ready to learn and benefit from training are trained first. Others are prepared for learning so that training resources are used when they will have the greatest impact. Another potential use of this variable is in personnel selection. Given that child welfare workers are constantly being trained in new skills and concepts, openness to learning is an invaluable characteristic for success in the field.

This research also highlights the importance of training reinforcement following formal training programs. Solidifying this knowledge gain through reinforcement is important given the finding that learning promotes training transfer. The child welfare training evaluation literature has often stated this need for training reinforcement (Miller &

Dore, 1991; Leung, Cheung, & Stevenson, 1994). An additional component of the current training program is "training refresher courses" for each of the supervisors/teams who complete training. These refresher courses were scheduled to begin after the evaluation was complete. Evaluation of learning and transfer of training following the refresher courses will provide additional support for this type of reinforcement.

A final major implication of this research is the need to maximize organizational support of training. The models generated by this research identify organizational support, specifically management support of supervisor, as a primary predictor of training transfer. This finding suggests that there would be a significant return on investment to get supervisor/management buy-in for training and ensure that organizational support components (e.g. forms, computer systems, policies) are in place before training.

In conclusion, this research has made a significant contribution to the child welfare training evaluation literature. These findings have added empirical support specific to child welfare on the importance of learning readiness and management support of training. The relationships between "levels" of training evaluation, such as reactions, learning and transfer, have been clarified, as reactions were identified as a mediator of learning and learning as a predictor of transfer. Finally, this research has generated potential models of prediction for training transfer, an area of training evaluation that has been previously difficult to measure. As child welfare continues to face challenges of policy change and multi-problem families, such research on child welfare training evaluation will be essential for the adequate preparation of professionals in the field.

References

The Adoption and Safe Families Act of 1997. (1997). Pub. L. No. 105-89.

Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, *42*(2), 331−341.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of relations among training criteria. *Personnel Psychology*, *50*, 341−358.

Alvarez, K., Salas, E., & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review*, *3*, 385−416.

Barbaranelli, C., Caprara, G. V., & Maslach, C. (1997). Individuation and the Five Factor Model of personality traits. *European Journal of Psychologica Assessment*, *13*(2), 75−84.

Barbee, A. P., & Barber, G. (1995). *A multidisciplinary training approach to substance abuse as it relates to child abuse and neglect.* Final evaluation report for grant #90-CP-0059 to National Center on Child Abuse and Neglect.

Barbee, A. P., Bledsoe, L., Antle, B. F., & Yankeelov, P. A. (1999). *An evaluation of the Virtual Office pilot project for the Cabinet for Families and Children.* Final evaluation report for child welfare training assessment grant (part).

Beckman, L. J., & Mays, V. M. (1985). Educating community gatekeepers about alcohol abuse in women: Changing attitudes, knowledge and referral practices. *Journal of Drug Education*, *15*, 289−309.

Burke, L. A. (1997). Improving positive transfer: A test of relapse prevention training on transfer outcomes. *Human Resource Development Quarterly*, *8*(2), 115−128.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "Big Five Questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences*, *15*(3), 281−288.

Cauble, A. E., & Thurston, L. P. (2000). Effects of interactive multimedia training on knowledge, attitudes, and self-efficacy of social work students. *Research on Social Work Practice*, *10*(4), 428−437.

Chawla, S., & Renesch, J. (1995). *Learning organizations: Developing cultures for tomorrow's workplace.* Portland, OR: Productivity Press.

Cheung, K. M., & Stevenson, K. M. (1991). *Competency-based evaluation of case-management skills in child abuse intervention.*

Christensen, D. N., Todahl, J., & Barrett, W. G. (1999). *Solution-based casework: An introduction to clinical and case management skills in casework practice.* New York: Aldine DeGruyter.

Coetsee, W.J. (1998). An evaluation model for human resources interventions. Unpublished dissertation. Rand Afrikaans University, Johannesburg.

Curry, D. (2001). Evaluating transfer of learning in human services. *Journal of Child and Youth Care Work*, *15–16*, 155−170.

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

Curry, D., McCarragher, T., & Dellmann-Jenkins, M. (2005). Training, transfer, and turnover: Exploring the relationships among transfer of learning factors and staff retention in child welfare. *Children and Youth Services Review*, *27*, 931−948.

Denning, J. D., & Verschelden, C. (1993). Using the focus group in assessing training needs: Empowering child welfare workers. *Child Welfare*, *72*, 569−580.

Delewski, C. H., & Pecora, P. J. (1986). Evaluating CPS training: The participant action plan approach. *Child Welfare*, *65*(6), 579.

Deschant, K. C., & Marsick, V. (1993). *Team learning survey, facilitator guide.* Pennsylvania: Organization Design and Development.

Durr, R., Guglielmino, L. M., & Guglielmino, P. J. (1996). Self-directed learning readiness and occupational categories. *Human Resources Development Quarterly*, *7*(4), 349−358.

Fleishman, E. A. (1957). A leader behavior description for industry. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* : Columbus Bureau of Business Research, Ohio State University.

Ford, J. K., Quinones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. *Personnel Psychology*, *45*(3), 511−524.

Gleeson, J. P., & Philbin, C. M. (1996). Preparing caseworkers for practice in kinship foster care: The supervisor's dilemma. *The Clinical Supervisor*, *14*(1), 19−34.

Gregoire, T. K. (1994). Assessing the benefits and increasing the utility of addiction training for public child welfare workers: A pilot study. *Child Welfare*, *73*, 69−79.

Highsmith, R., & Ilian, H. (2005). Aiming at a moving target: A multidimensional evaluation of supervisory training—Two years later. *Proceedings of Eighth Annual National Human Services Training Evaluation Symposium* Berkeley, CA: California Social Work Education Center.

Holton, E. F. (1996). The flawed four level model. *Human Resource Development Quarterly*, *7*(10), 5−19.

Hopkins, K. M., Mudrick, N. R., & Rudolph, C. S. (1999). Impact of university/agency partnerships in child welfare on organizations, workers, and work activities. *Child Welfare*, *78*(6), 749−773.

Jones, M. L., & Biesecker, J. L. (1980). Training in permanency planning: using what is known. *Child Welfare*, *59*, 481−489.

Kessler, M. L., & Greene, B. F. (1999). Behavior analysis in child welfare: Competency training caseworkers to manage visits between parents and children in foster care. *Research on Social Work Practice*, *9*(2), 148−170.

Kirkpatrick, D. L. (1959). Techniques for evaluating programs. *Journal of the American Society of Training Directions*, *13*(11), 3−9.

Leung, P., Cheung, K. M., & Stevenson, K. M. (1994). A strengths approach to ethnically sensitive practice for child protective service workers. *Child Welfare*, *73*, 707.

McCowan, R. J., McGregor, E. N., & LoTempio, S. J. (1989). Competency-based evaluation of social services training. *Journal of Continuing Social Work Education*, *2*, 11−31.

Miller, J., & Dore, M. M. (1991). Innovations in Child Protective Services in-service training: Commitment to excellence. *Child Welfare*, *70*, 437−450.

Moore, T. D., Rapp, C. A., & Roberts, B. (2000). Improving child welfare performance through supervisory use of client outcome data. *Child Welfare*, *79*, 475−497.

Patterson, G. (2004). Evaluating the effects of child abuse training on the attitudes, knowledge and skills of police recruits. *Research on Social Work Practice*, *14*, 273−280.

Pecora, P. J., Delewski, C. H., Booth, C., Haapala, D. A., & Kinney, J. (1985). Comparing intensive family preservation services with other family-based service programs. In E. M. Tracy, D. A. Haapala, J. Kinney, & P. J. Pecora (Eds.), *Intensive family preservation services: An instructional sourcebook* (pp. 117–142). Cleveland, OH: Mandel School of Applied Social Sciences, Case Western Reserve University.

Peters, L. H., O'Connor, E. J., Eulberg, J. R., & Watson, T. W. (1988). An examination of situational constraints in Air Force work settings. *Human Performance*, *1*, 133−144.

Rooney, R. H. (1988). Does in-service training make a difference? Results of a pilot study of task-centered dissemination in a public social service setting. *Journal of Social Service Research*, *8*, 33−50.

Rouiller, J. Z., & Goldstein, I. L. (1993). The relationship between organizational transfer climate and positive transfer of training. *Human Resource Development Quarterly*, *4*(4), 377−390.

Rubin, A., & Babbie, E. (1997). *Research methods for social work.* Pacific Grove: Brooks/Cole Publishing.

Shelton, S., & Alliger, G. (1993). Who's afraid of level four evaluation? *Training and Development*, *47*(6), 43−50.

Smith, T. E., Schinke, S. P., & Springer, D. W. (2000). Single-system evaluation of child protective services training. *Professional Development: The International Journal of Continuing Social Work Education*, *3*, 33−39.

Szilagyi, A. D., & Sims, H. P. (1974). Cross-sample stability of the Supervisory Behavior Description Questionnaire. *Journal of Applied Psychology*, *59*(6), 767−770.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics.* Needham Heights, MA: Allyn & Bacon.

Van Zyl, K. & van Zyl, M.A. (2000). Re-conceptualizing learning readiness and standardizing the Learning Benefit Inventory. Unpublished manuscript.

Viswesvaran, C., & Oanes, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, *8*(4), 216−226.

Xiao, J. (1996). The relationship between organizational factors and the transfer of training in the electronics industry in Shenzhen, China. *Human Resources Development Quarterly*, *7*(1), 55−57.

Yankeelov, P. A., Barbee, A. P., Barber, G., & Fox, S. (2000). Timing isn't everything, but it can be important. *Training and Development in Human Services*, *1*, 67−81.