International Journal of Training and Development 10:4
ISSN 1360-3736

The suitability of computer-based training for workers with limited formal education: a case study from the US agricultural sector

W. Kent Anger, Jeff Stupfel, Tammara Ammerman, Alys Tamulinas, Todd Bodner and Diane S. Rohlman

The suitability of computer-based instruction (CBI) for workers with limited education was evaluated in an Hispanic orchard workforce that reported little computer experience and 5.6 mean years of formal education. Ladder safety training was completed by employees who rated the training highly (effect size [d_gain] = 5.68), and their knowledge of ladder safety improved (d_gain = 1.45). There was a significant increase ($p < 0.01$) in safe work practices immediately after training (d_gain = 0.70), at 40 days post training (d_gain = 0.87) and at 60 days (d_gain = 1.40), indicating durability. As in mainstream populations, reaction or affective ratings correlated well with utility ratings, but not with behavior change. This demonstrates that an agricultural workforce with limited formal education can learn job safety from CBI and translate the knowledge to work practice changes, and those changes are durable.

❒ W. Kent Anger, Jeff Stupfel, Tammara Ammerman, Alys Tamulinas, and Diane S. Rohlman are from Oregon Health & Science University, Center for Research on Occupational Safety and Health, 3181 SW Sam Jackson Park Rd. L606, Portland, OR 97239, USA. Email: anger@ohsu.edu; jeffstupfel@hotmail.com; amme8169@pacificu.edu; tamulial@ohsu.edu; rohlmand@ohsu.edu. Todd Bodner is from Portland State University; Psychology Department, Cramer Hall, Portland, Oregon 97207, USA. Email: tbodner@pdx.edu.

Introduction

The vast majority of training research has been conducted in well-educated, white-collar workers. For example, approximately 90 per cent of the references from two widely cited meta-analyses of the training effectiveness literature (Alliger *et al*., 1997; Arthur *et al*., 2003) reported research on managers, supervisors and college students, with the remaining 10 per cent on blue-collar workers. Indeed, scant attention has been paid to teaching safe work practices to workers on the 'shop floor'. Cohen and Colligan (1998) found only 80 publications on workplace safety and health that were scientifically sound from 1980 to 1996. Burke *et al*. (2006) found only 95 published quasi-experimental (field) studies of safety and health training interventions between 1971 and 2003 with sufficient information to calculate effect size (d).

While blue-collar workers have only infrequently been the subjects in training effectiveness research (Alliger *et al*., 1997; Arthur *et al*., 2003), Hispanic workers appear to be entirely absent from that research. Yet Hispanic workers from Central and South America are the largest segment of immigrant workers in the United States (Marotta & Garcia, 2003; Mines *et al*., 1997), and some of those workers in agriculture report that they receive limited occupational training (e.g. Arcury *et al*., 1999; Salazar *et al*., 2005). This may be one reason that the relative risk for occupational fatalities is 22 per cent higher for Hispanic men than the risk for all men and for any other gender or race/ethnicity group (US National Research Council, 2003). A deeper look suggests that the risk may be much higher because of a lack of injury and illness reporting (Salazar *et al*., 2005).

Training is the key to occupational safety, yet the educational tools designed for training US workers may not be effective in immigrant Hispanic workers because of: (1) their different cultural background (Gomez & Fassinger, 1994; Hansen, 2003; Harrison & Hopkins, 1967); (2) limited education (e.g. respectively 4.2 and 5.4 mean years in Hispanic agricultural workers, reported by Kamel *et al*., 2003 and McCauley *et al*., 2001); and (3) the lack of work experience in a country with a strong governmental regulation of the workplace, which is limited in Latin countries (Brunette, 2005). It is impossible to assess the combined impact of *missing* experience in an industrialized and regulated workplace *and* of *missing* 5–10 years of education, both within the context of the urgent daily need to feed and house one's family.

Computer-based instruction (CBI) is an efficient and effective training technology (Arthur *et al*., 2003; Burke *et al*., 2006) that is increasingly the only available option (Brown, 2001). However, CBI has been studied with only one Hispanic, blue-collar workforce (Anger *et al*., 2004), and that was a feasibility study. It may seem doubtful that anyone with limited education and computer experience would respond positively to and learn content presented in CBI, so it is not a choice that most employers are likely to make.

After meeting with an orchard company to describe a CBI format shown feasible to use with a contemporary Hispanic agricultural workforce including those with limited education (Anger *et al*., 2004), the company asked the authors to provide a CBI-based safety training to reverse an increasing number of ladder-related accidents. The company would likely have addressed this safety issue with a lecture-format approach had they not been offered a no-cost CBI option. We structured a research study to answer the question, 'Can CBI be used to train an Hispanic workforce with limited formal education?'

Methods

Ladder safety training was provided to the Hispanic workforce of an agricultural and food products company that farmed 1093 hectares of orchards. The project consisted of a needs assessment that identified an increase in ladder accidents (the most frequent type of accident in orchards, according to Salazar *et al*., 2005) in the prior two years, development of training content in a CBI format, observation of work practices in the

orchards, delivery of training with pre and postknowledge tests and reaction ratings, and observation of work practices after training.

Participants

The orchard company required their entire orchard staff of 109 men to take ladder safety training in a CBI program in January 2004. This report describes data collected from 51 of those orchard workers who signed an IRB-approved consent form translated into Spanish. All 109 worked in the orchards, and even the field supervisors in the sample performed tree-pruning work during much of the workday.

Participant demographics are listed in Table 1. Most participants were orchard workers (73 per cent), although two crew leaders and eight lead workers also participated. All but four orchard workers identified themselves as Hispanic, and all respondents identified their language as Spanish (10 people reported that they spoke some English). Virtually all employees had worked in orchards and used ladders for several years (mean = 10.5). The mean participant age was 46.6 (SD = 9.2) years, and they reported a mean of 5.6 (SD = 3.2) for years of education, all completed before entering the United States. The distribution of years of education is depicted in Figure 1. Table 1 also reveals that most participants reported little contact with technology. Of the 44 responding, 11 per cent reported using a computer in their work (these were the crew leaders and some lead workers, but computer use time was minimal from our observations), and 7 per cent indicated that they used a computer 2 or more hours per week at home. CBI was new to 86 per cent of the participants (Table 1).

Equipment and training software

Training was provided in the front half of a 10.4 meter long recreational vehicle (RV) outfitted for training. The multimedia ladder safety training (including pictures, movies and sound) was presented on laptop computers with a user-designed 9BUTTON response unit (Anger *et al*., 2001) placed over the keyboard. The 9BUTTON units consisted of nine large (20 × 15 mm) buttons arranged in an arc on a gray metal body that covered the keyboard.

Computers were placed approximately 0.5 m apart. Headphones were required, and reading glasses were offered (approximately 20 per cent of workers accepted the offer). The CBI software was cTRAIN (Anger *et al*., 2001). The 10 system user instruction screens had been specifically developed in an Hispanic agricultural workforce with limited formal education (Anger *et al*., 2004). These screens taught the trainee to navigate the program and to answer two-option multiple-choice questions. Computer-generated spoken text was an option that could be elected on each screen throughout the training. When selected, the text was spoken by the voice 'Carlos high quality' in 'Mexican Spanish' by the Macintosh OS 9 operating system.

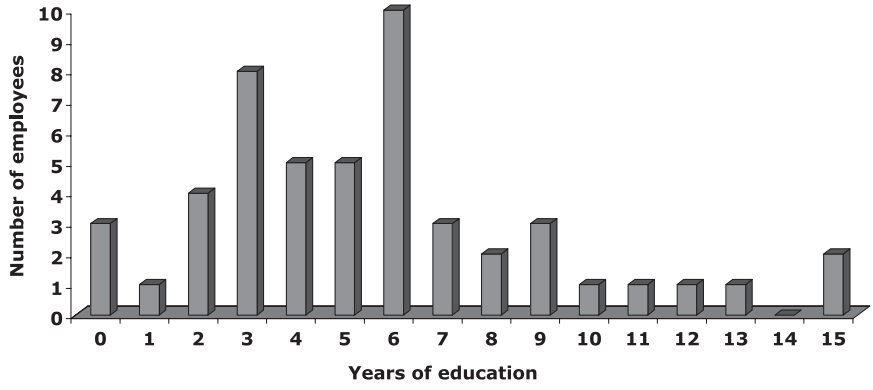

Figure 1: Distribution of years of education according to orchard worker self-report.

Table 1: Participant (volunteer) demographics (maximum N = 51)

Occupation	*N* (%)	Ethnicity	*N* (%)
Orchard worker	37 (73)	Hispanic	47 (92)
Equipment operator	4 (8)	African-American and Hispanic	2 (4)
Lead worker	8 (16)	No response	2 (4)
Crew leader	2 (4)		

	Mean (SD)
Years in current occupation	10.5 (7.2)
Age	46.6 (9.2)
Years of education	5.6 (3.2)
Years in US	20.5 (7.3)

Country of origin	%	Primary language	%
Mexico	96	Spanish	100 (of those responding)
El Salvador	2	No response (6)	
Honduras	2		

Activity	% (of 44 responding)
Watch TV (>2 h/week)	42
Read magazines (>2 h/week)	5
Read books (>2 h/week)	9
Play video games (>2 h/week)	0
Use Internet (>2 h/week)	9
Use a computer (>2 h/week)	7
Use a computer at work (per cent yes)	11

	Yes	No
Previous safety or job training on a computer (of 50)	7 (14%)	43 (86%)

SD = standard deviation.

The ladder safety training consisted of 33 information screens and 21 quiz screens presented in 7 'information sets'. Thus, there were a mean of 4.7 information screens and 3.0 quiz screens per information set. Many quiz questions presented a picture, and the text asked: Is the person in the picture working safely? The answers were either yes or no. Based on the structure of cTRAIN, participants would view all the information screens in an information set and then answer the quiz questions in the set. Correct answers to all quiz questions produced a stylized smiling face icon and a continuation to the next information set, whereas an incorrect answer caused the trainee to repeat the information screens and return to the same question. This continued until all questions were answered correctly, and the program then transitioned to the next information set, and so on.

Measures

Examiner rating of training difficulty

The training examiner (JS) provided a subjective rating of the ease or difficulty that employees had in completing the training, on a 4-point scale following the criteria in Table 2.

Table 2: Criteria for rating employees' training 'Easy, Somewhat easy, Difficult, Struggled'

Easy	Employee asked 0–1 question during the 10-screen system user instruction training and asked no questions during content delivery.
Somewhat easy	Employee asked two to three questions during system user instructions and gave the wrong answer to the 'practice' question at the end of the system user instruction screens, leading to repetition of the previous seven instruction screens (only one time).
Difficult	Employee asked four or more questions during user instructions, answered the 'practice' question at the end of the user instructions incorrectly one to two times (thus repeating the instructions), and asked up to five questions during content delivery.
Struggled	Employee asked questions throughout training, made repeated errors in the system user instructions, appeared frustrated and repeated the same mistake following an examiner intervention.

Reaction to training: employee's affective ratings (attitudes and expectations)
Following training, participants were asked to complete a series of written questions in Spanish[1]: (1) if they had previously completed any training on a computer; (2) how they rated the training (typically called the 'affective' rating in the research literature); (3) how this training compared with previous training; (4) what they expected from the training before they began; (5) how useful they thought the training would be; and (6) if the training would change the way they worked on ladders. Some participants elected to have the questions read to them by the examiner, an option offered when presenting the rating forms to avoid a possible embarrassment to nonreaders. Saying the ratings aloud to an examiner could bias responding. Participants filling out the questionnaire in writing were not required to speak their ratings directly to a member of the study team, reducing the potential bias.

Knowledge outcomes
The 21 two-answer quiz questions in the ladder safety training content were also presented by the CBI software, without feedback, as a pretest given before the training and a posttest taken immediately following the training.

Behavioral outcomes: work practices
Orchard workers used two pieces of equipment, ladders and loppers. The ladders were 2.4, 3.0 or 3.7 meters tall. They had steps on one side and a single pole or tongue on the other side. Ladder-related work practices can be described as a cycle: setting the ladder, climbing up, pruning, climbing down and carrying the ladder. A complete cycle was defined by 17 work practices, and these work practices were the behaviors recorded by examiners.

Behavioral outcomes: employee ratings of training impact
A structured questionnaire was administered orally in a conversational, informal manner. It sought participants' opinions on the impact of the training on their work and on specific examples about how the training had affected their work practices and those of their coworkers. The two primary questions, asked in Spanish, were: (1) Did

[1] Como califica el entrenamiento de hoy? (Excelente, Bueno, Así-así, No es bueno, Malo).
Cómo haga el entrenamineto de hoy **compara** con otro entrenamiento que ha tenido? (Lo de hoy fue **mucho mejor** que antes; **un poco major; igual calidad**; **un poco pejor**; **mucho pejor**).
Califica la información que aprendió de hoy. ¿Fue **útil o practical**? **(**Será **extramadamente** útil; **muy**; **un poco**; **No**; **No** sera útil **niguna** parte). Haga este entrenamiento un cambie de la manera que usted trabaja en escaleras? (Cambiará como trabajo en **todos lados** de mi trabajo; **muchas maneras; bastante un poco**; **un poco**; No se cambiará como trabajo).

you notice any difference in safety from previous years when workers did not get the computerized training; and (2) What is an example of a difference you noticed?

Results: injuries and workers' compensation claims
The company provided injury report summaries and workers' compensation claims for a four-year period, including the months after the training.

Procedures

Recruitment
Prior to training, employees were recruited to participate in a research study. Participating volunteers received a shirt embroidered with a pear, the name of the orchard company, the university conducting the study and 'Los trabajadores bien informados crean un lugar de trabajo sequro' (knowledgeable workers create a safe workplace).

Development of training content
Management had conducted a needs assessment by reviewing their workers' compensation data. They found an increase in ladder-related accidents in the past two years. The initial draft ladder safety training content was based on interviews with managers and supervisors. The draft content was 'pilot tested' with ten orchard workers and supervisors, and the feedback led to changes in roughly one-third of the original screens.

Delivery of training
The RV traveled to each of six 'orchard groups' during eight working days. Field supervisors transported groups of five employees from the orchard to the RV. All but one employee took the training in Spanish. An examiner with basic Spanish fluency initiated the training by saying 'press the 9 button to begin; the computer will show you what to do' in Spanish. He was also present to answer questions and to provide support as needed. Because the training was self-paced, people took between 40 min and 2.5 h to complete the training (approximation).

Behavioral measures: work practice observations
Two nonconcealed examiners observed volunteers' work practices during ladder work in the orchards. The examiners were cognizant of when training occurred and thus whether or not observations were being made pre or posttraining. Data were collected pretraining and in three posttraining observation periods: (1) the days immediately after training (post); (2) 40 days after training (post-40); and (3) 60 days after training (post-60).

Between 1 and 16 observations were made on each available employee during measurement periods. Before training, examiners recorded 178 observations of 45 workers; in post, 276 observations were recorded on 37 employees; in post-40, 145 observations were recorded on 27 employees; and in post-60 (when only a few employees were still pruning), 40 observations were recorded on 13 employees. In most cases, each observation included all 17 work practices. Duplicate observations were made on 74 (11.6 per cent) of the total 638 observations. The overall interobserver agreement of the paired observations was 90.3 per cent (SD = 9.3). At pre, the interobserver agreement of keeping loppers 12″ away from the face was very low while climbing up (47 per cent) and while climbing down (65 per cent). These low values were the result of an evolving procedural definition during the pretraining period. Excluding these two measures, the overall agreement was 95 per cent at pre, 90 per cent at post and 91 per cent at post-40.

Behavioral measures: employee ratings of training impact
At the post-40 period, the examiner who oversaw the training contacted each trainee individually while they were pruning in the orchards to administer the training impact questionnaire.

Results

The results address the effects of the training on employee affective ratings or attitudes toward training, knowledge of safe work practices and safe work practice behaviors in a context where training may be difficult because of language and education limitations. All 109 employees of the company's orchard staff completed the ladder safety training program (pretest, training, posttest). The results are drawn from the 51 employees who signed consent forms. Results were analysed and evaluated with RMANOVAs, ANOVAs, *t*-tests, Pearson correlations and effect size. Effect size (d_gain for the pre vs. post design) was calculated using the pooled standard deviation (SD) adjusted for the correlation between the two measures: SD of (post minus pre) scores divided by the square root of (2 * [1 – *r*]) (Lipsey & Wilson, 2001).

Evaluation of training difficulty

Of the 51 participants, all but two completed the CBI training with little assistance. The examiner was occasionally called on to provide help with the integral CBI user instructions, but he only rarely was asked to assist while completing the training content. The examiner did read the final 10–12 content screens for one participant who he judged to be having considerable difficulty with the training. Based on the criteria in Table 2, the examiner rated 2 participants (5 per cent of those rated) as having 'struggled' to complete the training, 6 (15 per cent) were rated as finding the training 'difficult', 20 (49 per cent) were rated as finding the training 'somewhat easy' and 13 (32 per cent) were rated 'easy'. Difficulty ratings were not assigned to 10 participants because of an insufficient observation time. The only trainees who could complete training without a rating were those who had minimal difficulty, thus biasing the data to include all those with difficult or struggled ratings at the expense of those who did not need help.

The Pearson correlation of the training difficulty ratings with years of education was $r\,(41) = -0.23$ (i.e. lower education was associated with greater difficulty). Of those who received a 'difficult' rating, five out of the six had 6 or fewer years of education. The correlation between difficulty and age was higher ($r\,[40] = 0.31$). The two participants with 'struggled' ratings were over 50 years of age. There was no correlation between difficulty ratings and years in the United States ($r\,[39] = 0.01$).

Reaction to training: employee's affective ratings (attitudes and expectations)

Each rating scale had five options, although the lower two options were not selected in response to four of the five questions. The employees' affective evaluation ('how do you rate today's training') was 'excellent' (62 per cent) to 'OK-neutral' (50 provided ratings) (see Figure 2, top left panel). When asked how the CBI training compared to other training taken in the past, 84 per cent of employees rated it as 'much better than previous training' (see Figure 2, top right panel). Regarding their expectations of the training, 79 per cent expected that the training would be 'much better than previous training' (47 providing a rating). As to the utility of the training, 39 per cent reported that it would be 'extremely useful', 53 per cent selected 'very useful', 4 per cent selected 'useful' and 4 per cent selected 'not at all [useful]' (51 provided ratings). When asked if the training would change the way they worked, 56 per cent selected 'in every way'.

In order to calculate an effect size for the employees' affective rating ('how do you rate today's training'), a 'pre' measure was derived by reversing the employee's rating of prior training. Thus, if the CBI training was 'much better' than prior training, then the prior training was assigned a rating of 'much worse', 'better' was assigned a rating of 'worse' and 'OK-neutral' was not changed. The difference between these ratings and the participant ratings of the current training was significant by a paired *t*-test ($t(48) = 24.67$, $p < 0.01$, d_gain = 5.68).

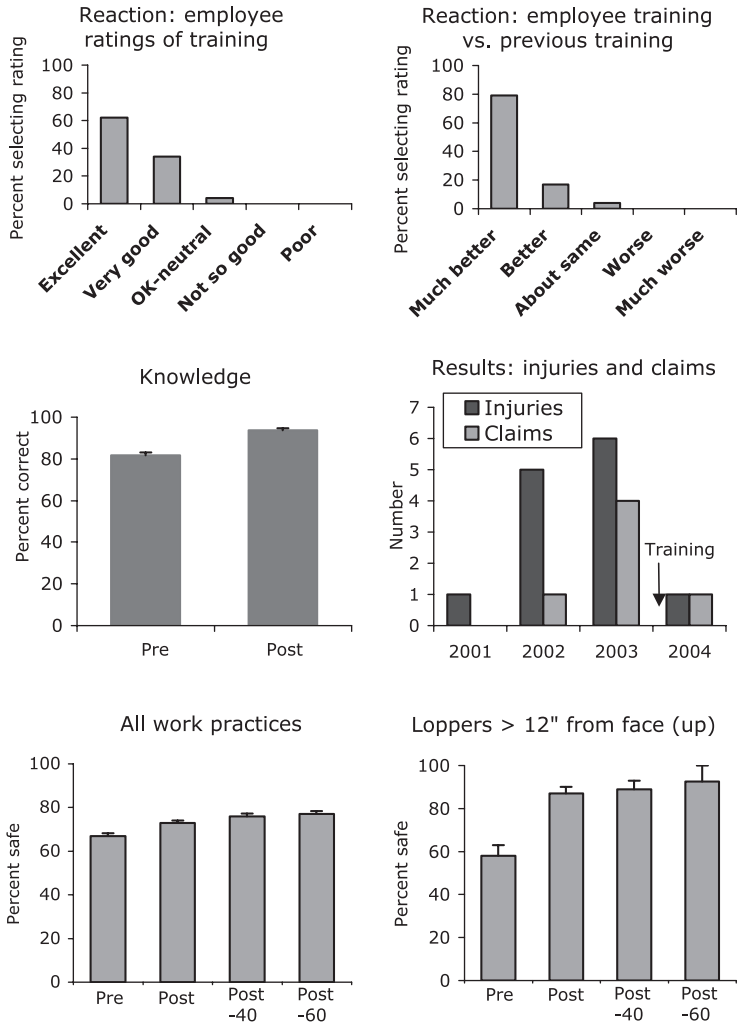

Figure 2: Depiction of main study findings: Reaction (top panels); Knowledge (middle left), Results: injuries and claims (middle right); and Behavior: all work practices (bottom left) and holding loppers more than 12″ from the face (bottom right). Reaction measures consist of the per cent selecting the employee ratings for the training and by comparison with previous training experience. Pre and posttraining knowledge or learning is presented as per cent of correct performance on two-answer questions. The number of injuries and claims are shown before (2001–2003) and after (2004) training. Behavioral changes are depicted as per cent of safe work practices for all measures (lower left) and holding loppers more than 12″ from the person's face while climbing up the ladder at pre and posttraining and post-40 and post-60 days after training (lower right panel).

Knowledge outcomes

Most of the orchard staff had worked in the orchards for several years and were familiar with safe work practices on ladders. The mean score on the 21 ladder safety questions given before training was 81.8 per cent (SD = 9.5, median = 85.7) (50 per cent was chance), and the mean score on the same test given immediately after the training

was 93.8 per cent (SD = 6.3; median = 95.2) (see Figure 2).[2] A paired samples *t*-test revealed that knowledge improved significantly from pre to posttraining ($t(50) = 9.59$, $p < 0.01$, 1-tailed, d_gain = 1.45, r (51) = 0.42 between pre and posttraining scores).

Behavioral outcomes: work practices aggregated across workers

The mean per cent safe work practices were aggregated across all participants. In sum, three work practices became safer following training, ten work practices were safe before and remained safe after training, and four work practices were unsafe before training and remained unsafe after training. More specifically, three work practices improved from 47–67 per cent before training to 84–86 per cent after training: keeping the loppers more than 12″ from their face while carrying the ladder, climbing up the ladder and climbing down. Ten work practices were performed safely 87–99 per cent of the time at pre and posttraining, and four safe work practices were performed only 9–15 per cent of the time at pre and posttraining. The same trends were seen at post-40 and post-60, suggesting that the changes were durable for the 2–3 months of pruning.

Behavioral outcomes: work practices by worker

To normalize work practice observations across workers, each orchard worker's data were reduced to the per cent correct performance on each work practice at each time period (pre, post, post-40, post-60). The mean per cent correct across all 17 work practices for each participant was aggregated to provide an overall change indicator. Figure 2 (bottom panels) depicts the per cent changes in all work practices (bottom left panel) and keeping loppers more than 12″ from the face (bottom right panel) at pretraining through post-60.

For each measure and the aggregate indicator, the difference between the per cent safe behaviors at pretraining was compared to post, post-40 and post-60 training periods by repeated measures ANOVA and by paired *t*-tests. The two analytic approaches produced similar results. Overall, compared to pretraining work practices, there was a significant increase in positive (safe) work practices across post, post-40 and post-60 by RMANOVA (F [3, 24] = 8.05, $p < 0.01$). Posthoc LSD tests were significant between pre and post ($p = 0.02$), post-40 ($p < 0.01$) and post-60 ($p < 0.01$).

As the lopper-close-to-face measures had a low interobserver agreement (see Methods), the RMANOVA and posthoc *t*-tests were repeated without the lopper data. The overall differences that were in a positive or safer direction remained significant (F (3,24) = 3.90, $p = 0.01$). Posthoc paired *t*-tests indicated that the changes from pre were not significant at post (t [8] = 0.52, $p = 0.31$, 1-tailed, d_gain = 0.70) but were significant at post-40 (t [8] = 2.65, $p = 0.02$, 1-tailed, d_gain 0.87) and at post-60 (t [8] = 2.51, $p = 0.019$, 1-tailed, d_gain 1.40).

A repeated measures analysis was also conducted on each of the 17 individual work practices. This analysis demonstrated that keeping loppers further than 12″ from the face decreased significantly while climbing up ($p < 0.01$), down ($p < 0.01$) and carrying the ladder ($p = 0.06$) overall and at every time point for climbing up and down ($p < 0.01$–0.05).

The per cent of workers who used safe lopper work practices more frequently after training than before training are depicted in the columns above the 0 line in Figure 3, while the per cent of workers with more unsafe practice observations after training compared to before are below the 0 line. The d_gain at post, post-40 and post-60, respectively, was 1.15, 1.24 and 1.09 for climbing up; 1.11, 1.63 and 1.47 for climbing

[2] Included in the data are revised scores for two people who had substantially lower scores on the posttest than on the pretest. At post-40, an examiner met individually with these two participants in their orchard, away from the other workers, and the participants repeated the posttest, orally. Their retest scores were 86 and 95 per cent compared to 52.4 and 66.7 per cent, respectively, immediately posttraining. The higher revised scores are included in the data analysed.

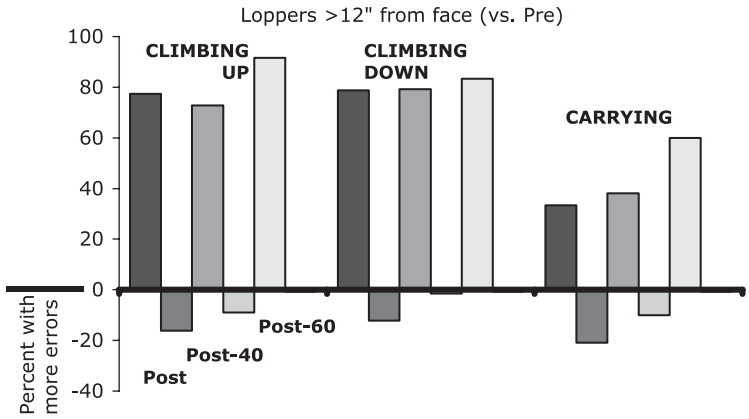

Figure 3: Per cent change from pretraining work practices to those in post, post-40 and post-60. The work practice depicted is workers holding loppers 12″ or more away from their face while climbing up, climbing down and carrying ladders. Columns above the '0' line reflect the per cent of workers with increased safe work practices ('per cent of workers improved') and columns below the '0' line reflect the per cent with less safe work practices ('per cent with more errors') for each work practice.

down and 0.37, 0.43 and 0.62 for carrying the ladder. These effect sizes while climbing up and down were large (0.8 or above per Cohen, 1992) but only medium while carrying the ladder.

Behavioral outcomes: employee responses about training impact

When queried at post-40 if the training had 'made a difference in safety', 84 per cent of participants responding (38) said yes. When asked to name a specific example of something that had changed, 16 provided an answer: 'better ladder placement' (3 employees), 'not carrying loppers on their shoulders' (2 employees), not overreaching, better balance (1 employee). Ignoring the request to identify *specific* changes, 5 responded that workers were safer, and 3 indicated that there had been no accidents following the training while 2 had 'learned' or 'remembered more'.

Results: injuries and workers' compensation claims

Accident/injury data and workers' compensation claims during the pruning season of January–March in 2001–2004 did not provide sufficient data for analysis. There were two injuries (no claims) in 2004 following the ladder and pruning safety training. This was an improvement over the two previous years when there were five and six injuries leading to one and four claims, respectively. However, there was only one injury and no claims in 2001 (see Figure 2, middle right panel). Participant consent was not requested for release of accident data; anonymous injury and claims data were provided by the company and of course reflected all orchard workers, not just the 51 participants. As noted, all 109 orchard workers completed ladder safety training at the same time in January 2004 at the outset of the pruning season.

Discussion

This study demonstrates that an agricultural workforce from an Hispanic culture, including those with limited or no formal education: (1) liked their CBI training experience compared to previous training (d = 5.68); (2) learned job safety information from a CBI presentation (d_gain = 1.45); (3) translated the knowledge to work practice

changes (d_gain = 0.71); and (4) those changes were durable over 40 days (d = 1.41). This bodes well for an efficient training of workplace safety in immigrant populations from a different culture with a different work experience and where access to education is limited. It is likely, however, that this positive outcome depended on using the 9BUTTON response unit instead of a keyboard (Rohlman *et al*., 2003), the use of behavioral education principles in the training program (Anger *et al*., 2001; Edgar & Sulzbacher, 1992), the availability of spoken text in Spanish, the extensive work to create effective system user instructions for those with limited formal education (Anger *et al*., 2004), and careful attention to the design and pilot testing of the training content. Thus, the training was specifically designed to train *all workers effectively*, the essential objective of workplace safety training. This may be contrasted with the disturbing direction taken by some, selecting only trainees who are likely to perform at a high level (Brown, 2001) instead of devoting time to make the training as effective as possible for *all* employees.

Knowledge outcomes: learning changes

Effect size (d) has been used to compare training effectiveness among different studies. Arthur *et al*. (2003) calculated the mean effect size d = 0.63 for learning or knowledge across all studies. Arthur *et al*. (2003) also reported that d = 0.40 for studies ($N = 7$) that employed 'computer assisted instruction' and d = 0.38 for studies ($N = 4$) that used 'computer assisted instruction and programmed instruction'. Arthur (personal communication, 2004) identified 26 articles reporting pre vs. post comparisons for the same group, the same design as in the present study that contained sufficient information to calculate effect sizes. Twenty-two data points from the articles using pre and post learning measures had a mean d = 0.98 (Arthur, personal communication, 2004), compared to a d_gain = 1.45 in the present study (the d and d_gain effect size measures are comparable[3]). In sum, this suggests that the training described in the present study of Hispanic workers was at least as effective in conveying knowledge as the other training reported in the literature.

Behavioral outcomes: work practice changes

Following training, there was a significant overall increase in safe work practices that improved in follow-up visits compared to pretraining work practices (d = 0.71 – 1.41). By comparison, Arthur *et al*. (2003) reported a mean d = 0.62 ($N = 122$ data points from the 165 studies) for behavior change in his meta-analysis of independent groups, and d = 1.69 ($N = 5$ data points) in his dependent groups' meta-analysis (Arthur, personal communication, 2004). Thus, the effect sizes found in the present study were higher than the mean effect sizes found in the literature for changing behavior in independent group studies but were below the mean effect sizes found in pre vs. posttraining comparison studies. Of course, the ladder safety work practices reported in the present study are very different from the largely academic or cognitive skill training in the studies analysed by Arthur *et al*. (2003) and Arthur, personal communication (2004).

Measures of training effectiveness

Training is rarely evaluated thoroughly. Focusing on Kirkpatrick's (1994) widely recognized four levels of evaluation, the American Society for Training & Development (Thompson *et al*., 2002) reported that 78 per cent of organizations assessed training effectiveness with reaction or affective rating, 32 per cent used learning or knowledge

[3] Arthur *et al*. (2003) used Dunlap *et al*.'s (1996) correction in his meta-analysis of independent group studies for those d's calculated from test statistics (e.g. ANOVA, *t*) rather than raw data (which used uncorrected d's). Arthur, personal communication's (2004) meta-analysis of pre vs. post designs also used Dunlap *et al*.'s (1996) correction, which takes the correlation between the pre and post scores into account, and is thus equivalent to the d_gain used in our study.

change, 9 per cent used behavioral change and 7 per cent employed results (cost savings) measures to assess training (the total sums up to 116 per cent because some organizations surveyed used more than one measure). Arthur *et al*.'s (2003) meta-analysis of 165 articles on training effectiveness revealed that relatively few articles incorporated more than one evaluation level, only 10 of the studies included three effectiveness measures and none included all four. The present study employed statistically reliable measures of three of the four and included data on the fourth measure of effectiveness.

The results of this study allow an examination of how relationships between measures of training effectiveness (Alliger *et al*., 1997) apply to the present population. Overall, the strongest correlations were between the pretraining and posttraining knowledge tests ($r = 0.42$), between the affective rating and utility ($r = 0.40$), between the affective rating and the prediction of changing work practices ($r = 0.43$), between pretraining and posttraining work practices ($r = 0.47$) and post-40 work practices ($r = 0.54$), and between post and post-40 work practices ($r = 0.42$). These correlations are presented in Table 3. Thus, two relationships suggested by Alliger *et al*.'s (1997) meta-analysis are supported: (1) affective ratings and utility ratings correlate highly; and (2) affective ratings are poor predictors of behavior (work practice) change. These relationships are thus extended to safety training in an Hispanic population with limited formal education.

Correlations were low between affective ratings and work practices at post ($r = 0.02$) and post-40 ($r = 0.08$), and the correlations were similarly low between workers' utility ratings and work practices at post ($r = –0.13$) and post-40 ($r = 0.04$) (Table 3). It may be that the restricted range of the ratings in the present study (virtually all ratings selected were in a 3-point range, despite the 5-point scales) and the limited option for improvement between the high pretraining learning scores and posttraining scores offered little potential for larger correlations (Nunnally & Bernstein, 1994). It should thus be noted that these data support calls for the use of learning and behavior change measures as direct indicators of training effectiveness, rather than affective or utility ratings (e.g. Alliger *et al*., 1997; Arthur *et al*., 2003).

Limitations

This study has limitations. The pre vs. post design does not allow us to fully address all threats to internal validity (Cook & Campbell, 1979; Shadish *et al*., 2002), the most widely described being the 'Hawthorne effect' that would predict an across-the-board improvement. Of course, the interpretation of this effect has been questioned (Adair, 1973; Bramel & Friend, 1981; Gillespie, 1988). That some behaviors changed while others did not argues against an interpretation that the presence of the experimenters or examiners and the attention stimulated the changes in work practices. In addition, the study is limited by the N of 51. Recruitment was compressed into a single day because of the constraints of the start-up of pruning, so there was no time to establish a relationship or even a benign presence. Thus, only about half of the orchard staff consented to participate in the study. This limited the size though probably not the representativeness of the sample. However, the statistical significance, which is of course adjusted for the N, and the effect size together suggest that the significant and substantial differences were produced by the training, not by the presence and interest of the examiners.

Training for Hispanic workers

Recent meta-analyses and literature reviews reveal that there is relatively little scientifically sound research on training effectiveness in the workplace and that there is virtually no research on the use of CBI in the workplace to teach employees with limited education (Alliger *et al*., 1997; Arthur *et al*., 2003; Burke *et al*., 2006; Cohen & Colligan, 1998). On the other hand, there are data that indicate that the Hispanic

Table 3: Correlations between levels of evaluation, knowledge and behavior before and after training

	Reaction		Knowledge		Behavior: work practices		
	Utility	Prediction	Pretest	Posttest	Pre	Post	Post-40
Affective rating	0.40 (50)	0.43 (48)	0.06 (50)	0.17 (50)	0.02 (43)	0.02 (43)	0.08 (26)
Utility rating		0.29 (48)	0.14 (50)	0.25 (50)	0.24 (43)	–0.13 (37)	0.04 (28)
Prediction rating			0.05 (47)	–0.01 (47)	0.00 (42)	–0.06 (35)	–0.23 (26)
Pretest rating				0.42 (51)	–0.05 (44)	0.17 (37)	0.17 (37)
Pretest % correct					–0.11 (44)	–0.10 (37)	0.02 (28)
Posttest % correct						0.47 (34)	0.54 (25)
Posttraining work practices							0.42 (26)

Note: The number of pairs for each correlation is in parentheses to the right of the correlation. The work practices measures reflect the summary per cent of safe work practices calculated from all work practices for each participant.

worker sustains more fatal and nonfatal injuries than other ethnic groups in high-hazard industries (Brunette, 2004; US National Research Council, 2003), and there have been calls for improved training for this group (Anonymous, 2001). Self-report data suggest that there are companies that do not provide sufficient training at the workplace for the Hispanic worker (O'Conner *et al*., 2005), including orchard workers (Salazar *et al*., 2005). Although the trend for training is slowly increasing for all work population segments, those with less education report less on-the-job and other training than those who are well educated (Stern *et al*., 2004). Although there is research supporting the effectiveness of workplace training in Hispanic workers (e.g. Forst *et al*., 2004), the training involves relatively expensive individual trainers that are not realistically within the means of many companies that employ such workers. There is research reporting the use of CBI in the workplace, but prominent examples of that research specifically exclude those with limited computer literacy (e.g. Simon & Werner, 1996), and one review recommends the selection of those likely to be successful as the *sole* candidates for CBI (Brown, 2001). However, the goal of safety and health training must be to teach every worker, as was the goal of the present study.

Summary of findings

The present study demonstrated positive, statistically significant changes in affective ratings or reaction (d_gain = 5.68, based on comparison with prior training), knowledge (d_gain = 1.45) and behavior or work practices (d_gain with [and without] loppers = 0.71 [0.17], 0.88 [0.33] and 1.41 [1.00], respectively, at post, post-40 and post-60), and a trend in reduced accidents (results) from the previous two years. By comparison, Arthur *et al*. (2003) reported that mean effect sizes in these four evaluation levels were: 0.60 (reaction), 0.63 (learning or knowledge), 0.62 (behavior) and 0.62 (results). Eckerman *et al*.'s (2004) study of food services workers concluded that CBI could be used with a population that included members with limited formal education (mean = 12.4; SD = 2.7, range 3–16 years of education). The present study extends this conclusion downward to those with no formal education and from a different culture. Although one format cannot be optimal for everyone, a format that is broadly effective would respond to the pressure for cost-efficient training to teach basic workplace skills and safety. Thus, calls for improved safety training materials for Hispanic workers (Brunette, 2005) can be answered by computer-based instruction, and clearly, the need is great for orchard workers such as those studied here and, by implication from Salazar *et al*. (2005), for all Hispanic workers in dangerous occupations.

Acknowledgements

This project was conducted at Bear Creek Corporation in Medford, OR. The initial training content was developed through interviews with Keith Emerson, Annika Forester and Benito Barajas of Bear Creek Corporation in Medford, OR. Appreciation is extended to the orchard management and employees of Bear Creek Corporation, who were extremely supportive and accommodating. Dr. Winfred Arthur of Texas A&M University generously provided data on pre- versus posteffect sizes collected for but not published as part of his analysis of training effectiveness (Arthur et al., 2003). Dr. Arthur and Dr. George Alliger supplied the references used in their meta-analyses of training effectiveness (Alliger et al., 1997; Arthur et al., 2003). Appreciation is also due to Nick Gorman of Salem, OR's Willamette University, for carrying out the statistical analyses, to Heather Fercho for data verification and spreadsheet creation, and to Dr. David Eckerman of the University of North Carolina at Chapel Hill, Donald Truxillo and to Robert Sinclair of Portland State University for counsel on the manuscript. The use of a Palm database for collecting data was suggested at a workshop by Dr. Raymond Romanczyk, Ms. Jennifer Gillis and Dr. Stephanie Lockshin at the Association for Behavior Analysis in 2002, 'Effective Supervision and Data Collection in Applied Settings: Maximizing Efficiency by Going Paperless'.

This research complied with APA ethical standards for human participant research, was performed according to the protocol approved by the Oregon Health & Science University (OHSU) Institutional Review Board (IRB), and all participants were provided with Spanish-language OHSU IRB-approved consent forms which we offered to read to them and gave to them to read, after summarizing the main points. Funding for this research was provided by outreach funds of the Center for Research on Occupational and Environmental Toxicology, NIOSH RO1 OH04193, UO1 OH 008108 and NIEHS R41 ES013088. This report is the sole responsibility of the authors and does not imply endorsement by agencies funding the research. OHSU and Drs. Anger and Rohlman have a significant financial interest in Northwest Education Training and Assessment, LLC, a company that may have a commercial interest in the results of this research and technology. This potential conflict was reviewed, and a management plan, approved by the OHSU Conflict of Interest in Research Committee, was implemented.

References

Adair, J. (1973), 'The Hawthorne effect: a reconsideration of a methodological artifact', *Journal of Applied Psychology*, **69**, 334–45.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H. and Shotland, A. (1997), 'A meta-analysis of the relations among training criteria', *Personnel Psychology*, **50**, 341–58.

Anger, W., Rohlman, D., Kirkpatrick, J., Reed, R., Lundeen, C. and Eckerman, D. (2001), 'cTRAIN: a computer-aided training system developed in SuperCard© for teaching skills using behavioral education principles', *Behavior Research Methods, Instruments and Computers*, **33**, 277–81.

Anger, W., Tamulinas, A., Uribe, A. and Ayala, C. (2004), 'Computer-based training for immigrant Latinos with limited education', *Hispanic Journal of Behavioral Sciences*, **26**, 373–89.

Anonymous (2001). 'Hispanic worker fatalities rise; improved training is needed', *Professional Safety*, **46**, 12, 6.

Arcury, T. A., Austin, C., Quandt, S. A. and Saavedra, R. (1999), 'Enhancing community participation in intervention research: farmworkers and agricultural chemicals in North Carolina', *Health Education & Behavior*, **26**, 563–78.

Arthur, W., Bennett, W. J., Edens, P. S. and Bell, S. T. (2003), 'Effectiveness of training in organizations: a meta-analysis of design and evaluation features', *Journal of Applied Psychology*, **88**, 2, 234–45.

Bramel, D. and Friend, R. (1981), 'Hawthorne, the myth of the docile worker, and class bias in psychology', *American Psychologist*, **36**, 867–78.

Brown, K. G. (2001), 'Using computers to deliver training: which employees learn and why?' *Personnel Psychology*, **54**, 271–96.

Brunette, M. (2004), 'Construction safety research in the United States: targeting the Hispanic workforce', *Injury Prevention*, **10**, 244–8.

Brunette, M. (2005), 'Development of educational and training materials on safety and health', *Family and Community Health*, **28**, 253–66.

Burke, M., Sarpy, S., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. and Islam, G. (2006), 'Relative effectiveness of worker safety and health training methods', *American Journal of Public Health*, **96**, 2, 315–24.

Cohen, A. and Colligan, M. (1998), *Assessing Occupational Safety and Health Training* (Atlanta, GA: DHHS [NIOSH]).

Cohen, J. (1992), 'A power primer', *American Psychologist*, **112**, 155–9.

Cook, T. and Campbell, D. (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Chicago: Rand-McNally).

Dunlap, W., Cortina, J., Vaslow, J. and Burke, M. (1996), 'Meta-analysis of experiments with matched groups or repeated measures designs', *Psychological Methods*, **1**, 170–7.

Eckerman, D. A., Abrahamson, K., Ammerman, T., Fercho, H., Rohlman, D. S. and Anger, W. K. (2004), 'Computer-based training for food services workers at a hospital', *Journal of Safety Research*, **35**, 317–27.

Edgar, G. and Sulzbacher, S. (1992), 'Influences and effect of the behavioral paradigm in special education', in R. West and L. Mamerlynck (eds), *Designs for Excellence in Education: The Legacy of B. F. Skinner* (Longmont, CO: Sopris West), pp. 187–217.

Forst, L., Lacey, S., Chen, H. Y., Jimenez, R., Bauer, S., Skinner, S., Alvarado, R., Nickels, L., Zanoni, J., Petrea, R. and Conroy, L. (2004), 'Effectiveness of community health workers for

promoting use of safety eyewear by Latino farm workers', *American Journal of Industrial Medicine*, **46**, 6, 607–13.
Gillespie, R. (1988), *The Hawthorne Experiments and the Politics of Experimentation* (New Haven, CT: Yale University Press).
Gomez, M. and Fassinger, R. (1994), 'An initial model of Latina achievement: acculturation, biculturalism, and achieving styles', *Journal of Counseling Psychology*, **41**, 205–15.
Hansen, C. (2003), 'Cultural myths in stories about human resource development: analysing the cross-cultural transfer of American models to Germany and the Côte d'Ivoire', *International Journal of Training and Development*, **7**, 16–30.
Harrison, R. and Hopkins, R. (1967), 'The design of cross-cultural training: an alternative to the University model', *Journal of Applied Psychology*, **3**, 431–60.
Kamel, F., Rowland, A. S., Park, L. P., Anger, W. K., Baird, D. D., Gladen, B. C., Moreno, T., Stallone, L. and Sandler, D. P. (2003), 'Neurobehavioral performance and work experience in Florida farmworkers', *Environmental Health Perspectives*, **111**, 14, 1765–72.
Kirkpatrick, D. L. (1994), *Evaluating Training Programs: the Four Levels* (San Francisco, CA: Berret-Koehler).
Lipsey, M. and Wilson, D. (2001), *Practical Meta-Analysis* (Thousand Oaks, CA: Sage Publications).
Marotta, S. and Garcia, J. (2003), 'Latinos in the United States in 2000', *Hispanic Journal of Behavioral Sciences*, **25**, 13–34.
McCauley, L., Lasarev, M., Higgins, G., Rothlein, J., Muniz, J., Ebbert, C. and Phillips, J. (2001), 'Work characteristics and pesticide exposures among migrant agricultural families: a community-based research approach', *Environmental Health Perspectives*, **109**, 5, 533–8.
Mines, R., Gabbard, S. and Stierman, A. (1997), *A Profile of U.S. Farm Workers. Demographics, Household Composition, Income and Use of Services*. Based on data from The National Agricultural Workers Survey (NAWS) (Washington, DC: The Commission on Immigration Reform).
Nunnally, J. and Bernstein, I. (1994), *Psychometric Theory* (New York: McGraw-Hill).
O'Conner, T., Loomis, D., Runyan, C., dal Santo, J. and Schulman, M. (2005), 'Adequacy of health and safety training among young Latino construction workers', *Journal of Occupational and Environmental Medicine*, **47**, 3.
Rohlman, D. S., Gimenes, L. S., Eckerman, D. A., Kang, S. K., Farahat, F. M. and Anger, W. K. (2003), 'Development of the Behavioral Assessment and Research System (BARS) to detect and characterize neurotoxicity in humans', *NeuroToxicology*, **25**, 523–32.
Salazar, M., Keifer, M., Negrete, M., Estrada, F. and Sydner, K. (2005), 'Occupational risk among orchard workers', *Family and Community Health*, **28**, 239–52.
Shadish, W., Cook, T. and Campbell, D. (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (New York: Houghton-Mifflin).
Simon, S. and Werner, J. (1996), 'Computer training through behavior modeling, self-paced, and instructional approaches: a field experiment', *Journal of Applied Psychology*, **81**, 648–59.
Stern, D., Song, Y. and O'Brien, B. (2004), 'Company training in the United States 1970–2000: what have been the trends over time?' *International Journal of Training and Development*, **8**, 191–209.
Thompson, C., Koon, E., Woodwell, W. H., Jr. and Beauvais, J. (2002), *Training for the Next Economy: an ASTD State of the Industry Report on Trends in Employer-Provided Training in the United States* (Alexandria, VA: American Society for Training & Development).
US National Research Council (2003), *Safety is Seguridad* (Washington, DC: US National Research Council).