An Item Response Theory Analysis of Safety Knowledge in Colorado Farm Residents

Cheryl L. Beseler, PhD, and Lorann Stallones, PhD, MPH

Objective: This study was designed to identify safety knowledge questions that differentiated those with high and low vulnerability to injury and assess differences in safety questions by sex and age. **Methods:** Colorado farm operators and their spouses were recruited in 1993 from farm truck registrations ($n = 872$) and between 1993 and 1997 from property value assessment lists or rural directories ($n = 761$). Item Response Theory was used to assess ten safety questions. **Results:** Knowing fatigue causes injury best discriminated between high and low injury vulnerability; incorrectly answering the question about inattention causing injury was associated with highest vulnerability to injury. No measurement invariance was identified by sex, but those older than 65 years answered differently than those younger than 45 years ($P = 0.005$). **Conclusion:** Safety programs emphasizing the role of fatigue and inattention on injury risk might reduce injuries in the farming community.

Farming is a dangerous occupation in terms of work-related injuries. According to the Bureau of Labor Statistics,[1] 142 (2.7%) of the 5214 occupational fatalities occurred among agricultural workers, but agricultural workers accounted for less than 1% of workers. Fatalities among agricultural crop production workers showed an 18% increase in 2008 compared to 2007, while fatalities among animal production workers declined by 8%.[1] On farms with at least 11 employees, the incidence rate for fatal injuries was 32.5 per 100,000 full-time equivalent workers in crop production and 16.8 per 100,000 full-time equivalent workers in animal production, compared to 3.7 per 100,000 full-time equivalent workers for all industries combined.[1] As smaller farms are not included in these numbers and the seasonal and migrant workers are not likely to be included in these incidence rates, injuries on farms are most likely underreported.

Safety interventions designed to reduce injuries on farms have targeted increasing safety knowledge through training and increasing safe behaviors, but these efforts have not consistently resulted in significant reductions in injury rates.[2,3] Limited attention has been given to explaining the lack of effectiveness of the prevention efforts, but efforts to assess a farm worker's perception of safety knowledge may provide an improved understanding of the complex relationship between knowledge, behavior, and injuries. The relationship between farm safety knowledge and injuries has been studied.[3,4] Farm workers who have sustained a farming-related injury frequently attribute the injury to personal factors such as fatigue, hurry, carelessness, and inattention.[5,6] Even in the presence of adequate safety knowledge, personal factors may reduce the exercise of safe behaviors. In a sample of 87 older Illinois farmers, those with a high level of safety knowledge were significantly more likely to report an injury than those with a low level of safety knowledge, a result that was opposite the expected direction.[3] The investigators hypothesized that those who had experienced an injury were more knowledgeable about farm hazards than those who had not experienced an injury and that having more safety training made farmers better at recognizing injuries. The causal direction of this association could not be determined because the study was a cross-sectional design. In a study of 652 farm residents in Colorado, where part of the analysis was longitudinal, safety knowledge failed to predict injury except among those who were also depressed.[4] We previously showed that safety knowledge was significantly associated with only two of ten safety behaviors (wearing hearing protection and a dust mask).[4] Given the evidence that safety knowledge does not translate directly into exercising safe behaviors and personal factors may play an important role in injury, we used Item Response Theory (IRT) to develop a more comprehensive understanding of the properties of the safety knowledge scale used in previous studies. Furthermore, we seek to explore the functioning of items related to perception factors that may play a role in injury and explain the interaction with depression in our previous study. We hypothesized that a better understanding of the safety items might be used to identify perceptions that increase vulnerability to injury and could be targeted to improve farm safety interventions.

A two-parameter IRT model allows determination of items that discriminate between those with low and high safety knowledge and questions that are most difficult to answer correctly in those with a similar overall level of safety knowledge. The benefit of the IRT approach is that it allows for testing whether items are behaving differently by demographic group. In this report we address the following four questions using a large population of farm residents in Colorado: (1) Which safety items discriminated best between farm residents with high and low safety knowledge? (2) Which safety items were most difficult for farm residents to answer correctly given a certain level of safety knowledge? (3) Were there differences between how men and women answered the safety questions? (4) Were there differences in how the questions were answered by age groups with low rates of farm injuries and age groups with higher rates of injury?

METHODS

To provide sufficient power to examine differential item functioning (DIF) of safety knowledge questions by sex and age, we combined two samples of Colorado farm residents for a total sample size of 1628 after excluding five individuals with incomplete data. The two farming samples did not differ significantly by farm size or types of crops or animals produced.

Statewide Survey Participants

Farm operators and their spouses were recruited in 1993 from a list of farm truck registrations available from the Colorado Department of Motor Vehicles. A stratified probability sampling technique was used to select the number of farm operations to be included in the sample. A detailed description of the procedure has been previously

From the Colorado Injury Control Research Center, Psychology Department (Drs Beseler and Stallones) and Graduate Degree Program in Public Health (Dr Stallones), Colorado State University, Fort Collins, Colo.

Grant Sponsor: Centers for Disease Control and Prevention (CDC) National Institute of Occupational Safety and Health
Grant Number: U04/CCU806060
Grant Sponsor: CDC the National Center for Injury Prevention and Control
Grant Number: 1R49/CE001168

Address correspondence to: Cheryl L. Beseler, PhD, Research Faculty, Colorado Injury Control Research Center, Psychology Department, Colorado State University, 1879 Campus Delivery, Fort Collins, CO 80523; E-mail: cbeseler@lamar.colostate.edu.

DOI: 10.1097/JOM.0b013e31820f3843

published.[7] Farms were identified using the US Bureau of Agriculture definition requiring agricultural production sales of US $1000 or more in a normal year. A total of 876 primary operators and his or her spouse from 485 farms were enrolled. The initial response rate for farms based on total contacted minus refusals and procedural interview failures (could not understand the interviewer or questions) was 62%. The response rate based on contacted households minus outright refusals to participate was 70%. The survey was conducted by telephone and was designed to take approximately 20 minutes per respondent. Interviews were conducted during the slow months of agricultural production in the region: November to March of the study years.

Eight County Survey Participants

Principal farm operators and spouses ($n = 761$) in one of eight counties in Northeastern Colorado were enrolled between 1993 and 1997. Participants were selected from property value assessment lists or from rural directories. Farms were identified using the US Bureau of Agriculture definition requiring agricultural production sales of US $1000 or more in a normal year. A target sample of 500 farms was selected on the basis of power calculations done for outcomes of interest including farm injuries. All participants were interviewed in person at their farm by trained study personnel. Interviews took 45 minutes to 2 hours depending on details of personal histories. Interviews were conducted during the slow months of agricultural production in the region: November to March of the study years.

Study Variables

Study participants answered 10 questions concerning safety knowledge with possible responses being true, false, or do not know (Table 1). Answering a question correctly was coded as 0 and an incorrect or do not know response was coded as 1 to reflect increased vulnerability to injury if answering incorrectly. We included the covariates sex and age (categorized into <45, 45 to 64, and older than 65 years). Injury rates were calculated by totaling the number of injuries during the year prior to the interview, dividing by the number of hours worked and multiplying by 100 to obtain the injury rate per 100 hours of farm work.

Statistical Analysis

Descriptive statistics were used to summarize the demographic characteristics and to compare injury rates by gender and age group. Logistic regression was used to estimate odds ratios and 95% confidence intervals for associations between sex or age group and injuries. These analyses were done in SAS version 9.2 (SAS Institute, Cary, NC).

We previously showed that the safety knowledge questions loaded onto a single factor and fit a dimensional construct and could therefore meet the assumption of unidimensionality required by IRT.[4] Item response theory is used to describe the probability of observing a particular pattern of item endorsement given an individual's underlying knowledge or ability. The inputs for the IRT model are the factor loadings from a factor analysis and thresholds, which are analogous to means for continuous variables. A 2-parameter logistic model can be used to calculate the probability of a person with a given level of knowledge answering an item incorrectly. This probability is modeled using a slope parameter (discrimination or α) and a location parameter (difficulty or β) for each item.[8] The slope parameter measures the strength of the relationship between the safety item and the underlying latent factor from the factor analysis; a steeper slope means better item discrimination than a flatter slope for individuals at similar levels of knowledge. The difficulty of a question (the location parameter) is measured at the point on the vulnerability spectrum where the respondent has a 50% chance of answering incorrectly. The latent factor in this analysis represents vulnerability to injury because answers are coded to reflect endorsing an incorrect response.

TABLE 1. Percent of Incorrect Answers and Factor Loadings From Analysis of Ten Safety Knowledge Questions by 865 Men and 763 Women in a Total of 1628 Colorado Farm Residents, 1993–1996

Safety Question	Percent Incorrect	Factor Loading	Factor Loading: Men	Factor Loading: Women
More falls are connected with the use of farm machinery than any other farm work operation	49.5	0.57	0.63	0.51
Children are rarely injured in falls from farm equipment	27.3	0.61	0.58	0.64
Agriculture has one of the lowest injury rates of the major U.S. industries	18.2	0.77	0.75	0.80
Tractors are involved in many fatal farm injuries	14.7	0.66	0.60	0.71
Inattention is a leading cause of injuries	7.9	0.57	0.48	0.66
Farmers lung disease results from overexertion	21.7	0.43	0.40	0.46
Fatigue can lead to injuries	2.0	0.93	0.95	0.93
Hearing loss is common among farm workers	18.1	0.63	0.56	0.70
Animals are rarely responsible for injuries	23.3	0.53	0.53	0.53
Amputation injuries are more common when working with farm equipment than when involved in other agricultural work	13.0	0.64	0.63	0.66

We first factor analyzed the ten safety items in the full sample and tested the fit of one- and two-factor models using the comparative fit index (CFI) and root mean squared error of approximation (RMSEA). A model with a high CFI (>0.95) and low RMSEA (<0.05) is considered to adequately fit the data. Weighted least squares were used for the factor analysis and maximum likelihood for the IRT analysis with the sandwich estimator to generate standard errors because it is robust to non-normality. We used the item factor loadings and thresholds to calculate a two-parameter normal ogive item response model.[8,9] Item characteristic curves (ICCs) and test information curves (TICs) were used to display the results. Those with high safety scores are more likely to answer the more difficult questions incorrectly and would fall on the left side of the ICC graph; those with a low safety score would have a higher probability of answering the easier questions incorrectly and would fall on the right side of the ICC graph.

Item response theory models were tested for DIF to determine whether the safety questions performed differently for men compared to women, or for younger farm residents compared with older farm residents (measurement non-invariance). This analysis was conducted by estimating a model where all factor loadings and thresholds were allowed to vary across groups (the less restrictive model).[10,11] The less restrictive model was compared with more restrictive models, where thresholds or factor loadings were constrained to be equal across groups. Structural invariance (population heterogeneity) was examined by testing whether the factor mean in women differed from the factor mean in men, for example. The delta parameterization was used because the factor loadings were the parameters of interest, not the residual variances. Chi-square tests were used to assess whether significant differences were found after constraining the parameters of the nested model, which would indicate model non-invariance. If not significant, then measurement invariance holds and there are no differences between the groups. Results were considered significant at $P < 0.05$ for a two-tailed test, except in pairwise comparisons ($P < 0.02$) and when comparing the individual item parameters, which were considered significant at $P < 0.005$ after Bonferroni correction.[12] All analyses were done in MPlus version 5.[13]

RESULTS

Among the 1628 farm residents, 98.9% were white and 2.6% Hispanic. They were 53.0% men, 91.4% were married, 91.7% were high school graduates, 93.6% reported being in good, very good, or excellent health, 31.4% reported having financial problems, and they had experienced 154 injuries in the past 12 months. Of the 154 injuries, 131 individuals reported a single injury, 19 reported two injuries, and 4 reported more than two injuries (range, 3 to 6). Only about half of the farm residents correctly answered the question related to falls, but nearly everyone correctly answered that fatigue and inattention can lead to injury. The mean injury rates (±standard deviation) for women were 0.83 (±6.20) and for men 0.62 (±1.97). A total of 865 men reported 82 injuries (9.5%) and 767 women reported 72 injuries (9.4%). Farm residents younger than 45 years had the highest mean injury rate (1.13 ± 6.12), those 45 to 64 years of age had a lower mean injury rate (0.51 ± 2.08), and those older than 65 years had the lowest injury rate (0.18 ± 1.15). In the 680 who were under 45 years of age, there were 82 injuries (12.1%); in the 732 who were 45 to 64 years of age, there were 60 injuries (8.2%); and in the 221 who were 65 years of age and older, there were only 12 injuries (5.4%). Compared to those 45 to 64 years of age, the younger group had a significantly elevated odds of injury (OR = 1.54; 95% CI = 1.08 to 2.19) and those 65 years of age and older had a lower odds of injury (OR = 0.64; CI = 0.34 to 1.22).

In the full sample, all factor loadings were greater than 0.40 (Table 1), the eigenvalues were 4.62 and 1.06, indicating that nearly all of the variance loaded on one underlying latent factor. The CFI was 0.97 and the RMSEA was 0.04, which indicates a good fit of the data to the underlying latent injury vulnerability factor. The factor analysis showed that fatigue had the highest factor loading (0.93) and lung disease had the lowest (0.43). The questions concerning fatigue leading to injuries and agriculture having high injury rates best discriminated between those with high and those with low safety knowledge, although most farm residents answered correctly (Table 2 and Fig. 1). The remaining safety questions did not discriminate between different levels of safety knowledge (Table 2). Farm residents with low levels of safety knowledge, or greater vulnerability to injury, had the most difficulty answering the fatigue question and were less likely to answer correctly that inattention can lead to injuries (Table 2). Those with a higher safety score and lower vulnerability to injury were more likely to answer correctly whether falls were related to farm machinery (Fig. 1). Incorrectly answering questions about inattention and fatigue leading to injury identified those with a high vulnerability to injury. Overall, the difficulty parameter covered a broader range of the safety knowledge spectrum than the discrimination parameter (Fig. 1).

Differential Item Functioning by Sex

The item most highly correlated with a latent safety factor among men and women was fatigue. Among women, inattention and hearing loss were more strongly correlated with the safety factor than in men (Table 1). Nevertheless, a DIF analysis from the IRT showed that allowing the factor loadings to vary by sex did not significantly alter the model ($\chi^2 = 10.7$, $df = 6$, $P = 0.10$), nor did allowing the thresholds to vary by sex change the fit of the data to the model ($\chi^2 = 9.17$, $df = 7$, $P = 0.24$). Testing for structural invariance (population heterogeneity) by fixing the factor means to be equal in each sex group did not alter these results ($\chi^2 = 15.5$, $df = 9$, $P = 0.08$). Thus, the ten safety items did not function differently depending on whether the respondent was a man or women, nor was there evidence of differences in properties of the underlying latent factors comparing men and women.

Differential Item Functioning by Age Group

A model where the factor loadings were constrained to be equal across the three age groups did not change the fit of the data to the model ($\chi^2 = 9.15$, $df = 12$, $P = 0.69$). Nevertheless, allowing the factor loadings to vary and constraining the thresholds to be equal across groups resulted in a significantly poorer fitting model ($\chi^2 = 32.7$, $df = 15$, $P = 0.005$). Although this indicates some measurement invariance in the threshold parameters by age, no population heterogeneity, as measured by differences in factor means, was observed ($\chi^2 = 20.1$, $df = 19$, $P = 0.39$).

We further explored the evidence of measurement invariance by using three pairwise comparisons of the age groups to determine

TABLE 2. Results of Item Response Theory Parameters for the Total Sample and by Gender* for the Ten Safety Questions, Colorado Farm Residents, 1993–1996

	Total Sample (n = 1628)		Men (n = 865)		Women (n = 763)	
Safety Question	Discrimination	Difficulty	Discrimination	Difficulty	Discrimination	Difficulty
Falls	0.69	0.02	0.85	−0.04	0.58	0.10
Kids	0.77	0.98	0.71	0.98	0.85	0.98
Agriculture	1.24	1.16	1.10	1.21	1.42	1.11
Tractor	0.89	1.59	0.76	1.75	1.04	1.44
Inattention	0.67	2.59	0.56	2.96	0.82	2.30
Lung disease	0.45	1.87	0.42	1.90	0.49	1.85
Fatigue	1.57	2.55	1.77	2.75	1.56	2.32
Hearing loss	0.81	1.45	0.65	1.62	1.00	1.33
Animals	0.60	1.40	0.60	1.35	0.61	1.44
Amputation	0.82	1.79	0.79	1.76	0.85	1.83

*Differential item functioning analysis showed no significant differences by gender (discrimination: $\chi^2 = 10.7$ $P = 0.10$; difficulty: $\chi^2 = 9.17$, $P = 0.24$; factor means: $\chi^2 = 15.5$, $P = 0.08$.)

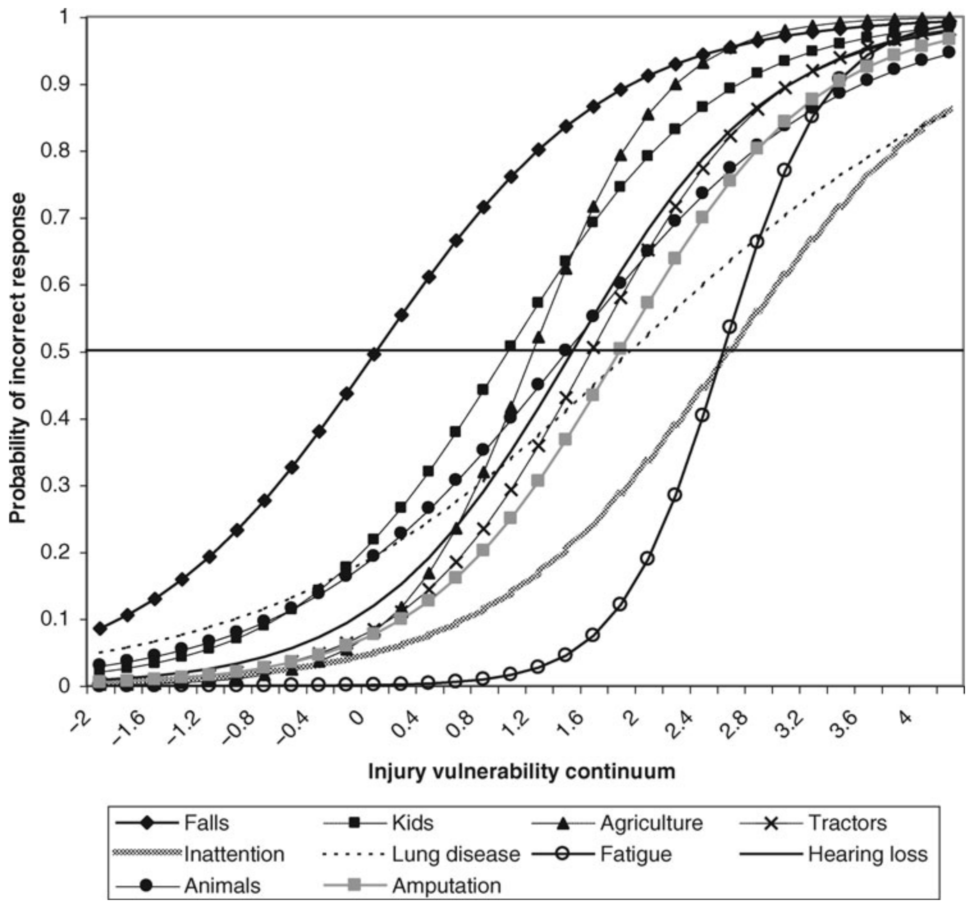

FIGURE 1. Probability of an incorrect answer to each of 10 safety questions in the entire sample of 1628 Colorado farm residents. (Horizontal line is the 50% point where difficulty is measured on the *x* axis. Those with a high level of safety knowledge answer the harder questions incorrectly and fall on the left side; those with a low level of safety knowledge answer the easier questions incorrectly and fall on the right side of the graph. Discrimination is the slope of the curve at the 50% mark. The greater the slope the better the question discriminates between those with high and low safety knowledge.)

which age groups differed from the others and what items differed by group. Significant measurement invariance was observed between farm residents younger than 45 years and older than 65 years ($\chi^2 = 18.7$, $df = 7$, $P = 0.009$). After Bonferroni correction, the difference between the 45- to 64-year group and those older than 65 years was not statistically significant ($\chi^2 = 15.1$, $df = 7$, $P = 0.04$); no difference was seen between those younger than 45 years and those 45 to 64 years of age ($\chi^2 = 13.0$, $df = 7$, $P = 0.07$). Thus, the difficulty in answering each of the items systematically and significantly differed in those younger than 45 years compared with those older than 65 years (Table 3 and Figs. 2 and 3). Although all items showed measurement non-invariance at $P < 0.05$, after adjustment for multiple comparisons, significant measurement non-invariance was identified in the thresholds for tractors being involved in many fatal farm injuries and farmer's lung disease resulting from overexertion comparing those younger than 45 years of age with those older than 65 years ($P < 0.005$). Comparing the middle group (ages, 45 to 64 years) to those older than 65 years showed that nearly all items (excluding amputation, $P = 0.16$) were answered differently between the groups, but none of the items reached significance after Bonferroni correction (P values 0.02 to 0.05).

Summing up item information curves to produce the TICs showed a very high peak for the older group. This indicates that DIF impacted the safety test as a whole and that the safety questions contained much more test information for those older than 65 years compared with the other age groups (Fig. 4). The safety knowledge test had better discrimination over a narrower range of items for older farm residents at the same knowledge level. Two peaks can be seen in the curve for older-than-65-year group. The smaller peak is at a lower knowledge level than the other groups and a second, much higher and narrower peak, is located at a higher knowledge level. This provides evidence that the scale is not unidimensional in this age group and this lack of dimensionality might point to a protective factor that accounts for the low injury rate in this group.

DISCUSSION

The fatigue item best discriminated between low and high safety knowledge in the sample and in each subgroup and was strongly correlated to the injury vulnerability factor. Although most farm residents answered this question correctly, some who otherwise showed high safety knowledge did not always answer the fatigue question correctly. This aspect of safety knowledge is a potential target for identifying farm residents who may be at high risk of injury. Future work should focus on developing more sensitive measures of attitudes toward fatigue and farm work and factors that might influence perceptions about fatigue. The questions concerning inattention and fatigue leading to injuries were the most difficult for the farm residents to answer correctly if they scored low on safety knowledge.

TABLE 3. Results of Separate Item Response Theory Models for Ten Safety Items in Three Age Groups,* 1628 Colorado farm Residents, 1993–1996.

	Age < 45 (n = 678)		Age 45–64 (n = 730)		Age > 64 (n = 220)	
Safety Question	**Discrimination**	**Difficulty**	**Discrimination**	**Difficulty**	**Discrimination**	**Difficulty**
Falls	0.57	−0.25	0.75	0.17	1.06	0.16
Kids	0.66	1.19	0.91	0.88	0.70	0.78
Agriculture	1.14	1.28	1.26	1.16	1.35	0.89
Tractor	0.97	1.51	0.91	1.57	0.63	1.90
Inattention	0.62	2.78	0.81	2.34	0.48	3.00
Lung disease	0.46	1.67	0.48	2.02	0.38	1.88
Fatigue	1.47	2.64	1.46	2.67	2.50	2.06
Hearing loss	1.03	1.26	0.73	1.62	0.67	1.41
Animals	0.59	1.35	0.59	1.48	0.67	1.31
Amputation	0.69	2.07	0.99	1.69	0.69	1.51

*Differential item functioning analysis showed no significant differences by age group for the discrimination parameters but difficulty parameters significantly differed in the oldest group compared with the youngest group.

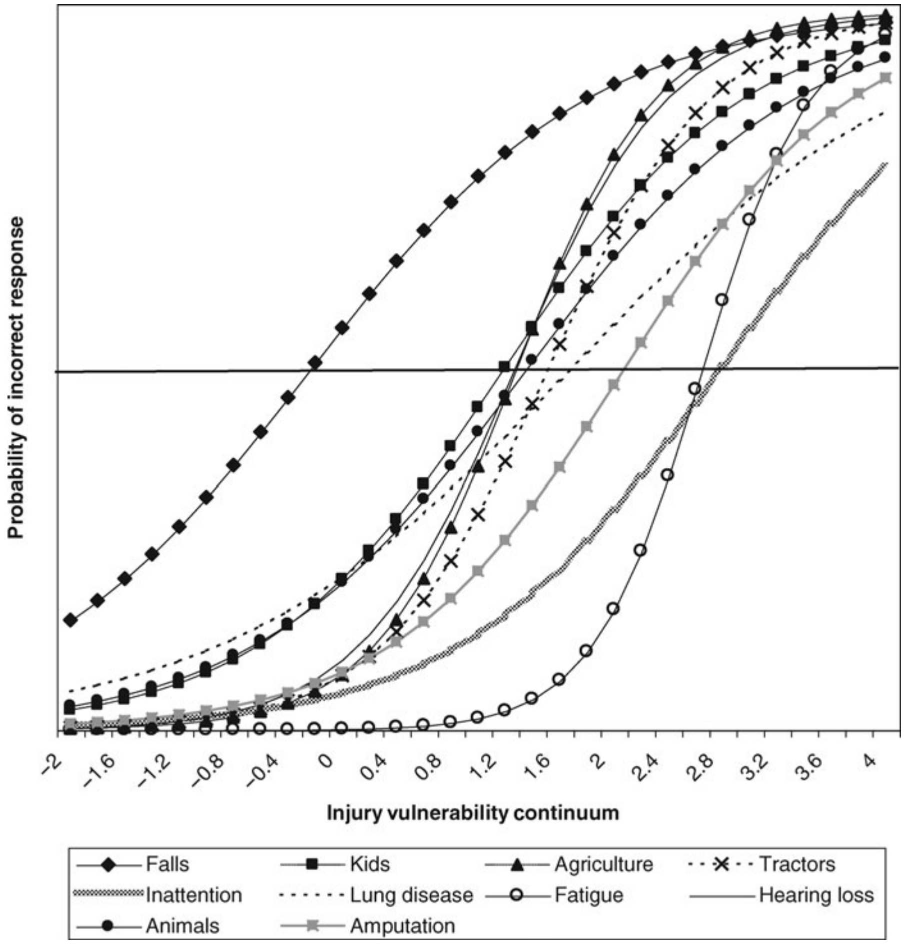

FIGURE 2. Probability of an incorrect answer to each of 10 safety questions in those younger than 45 years ($n = 678$). (Horizontal line is the 50% point where difficulty is measured on the *x* axis. Those with a high level of safety knowledge answer the harder questions incorrectly and fall on the left side; those with a low level of safety knowledge answer the easier questions incorrectly and fall on the right side of the graph. Discrimination is the slope of the curve at the 50% mark. The greater the slope the better the question discriminates between those with high and low safety knowledge.)

This was true for males and females, as well as by each of the age groups. The fatigue question was least difficult for those 65 years and older.

These results support previous reports that farmers often attribute injury to fatigue, hurry, carelessness, and inattention.[5,6] Therefore, safety perception scales should be developed and tested in the farming community that assess perceptions of personal factors and used in farm safety programs, especially in young farmers. More research is needed to help farmers understand the importance of taking breaks, recognizing when their attention is beginning to wander and recognizing when they are in a hurry and need to slow down. Teaching safety education and safe behaviors will not prevent injuries without also addressing these personal factors. Farmers might believe that they can and must overcome fatigue and can prevent inattention because of the solitary and fiercely independent nature of farming culture and the pressing need to get the work done.

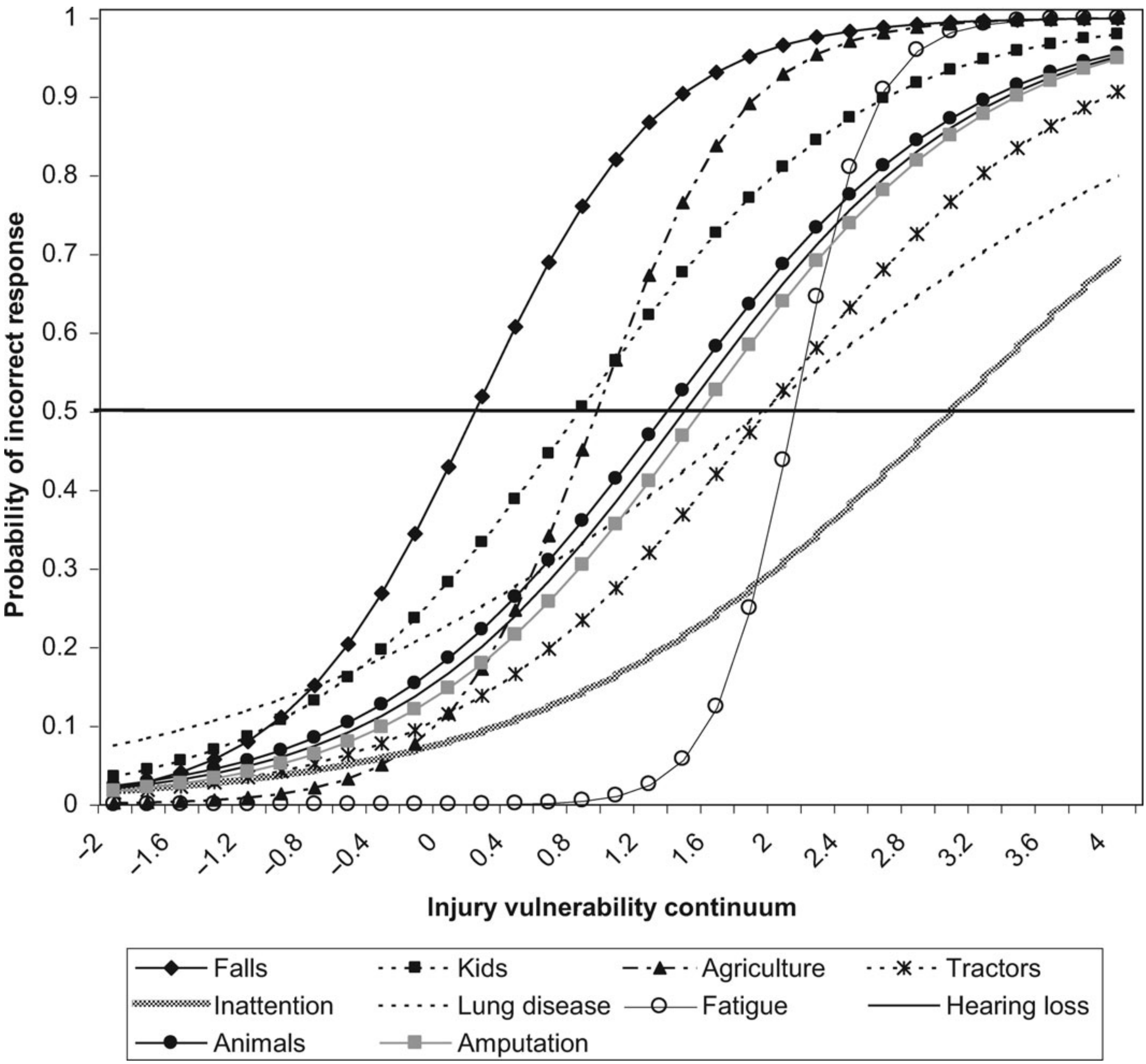

FIGURE 3. Probability of an incorrect answer to each of ten safety questions in those older than 65 years ($n = 220$). (Horizontal line is the 50% point where difficulty is measured on the *x* axis. Those with a high level of safety knowledge answer the harder questions incorrectly and fall on the left side; those with a low level of safety knowledge answer the easier questions incorrectly and fall on the right side of the graph. Discrimination is the slope of the curve at the 50% mark. The greater the slope the better the question discriminates between those with high and low safety knowledge.)

The ten safety items were shown to be measurement invariant across sex; they were similarly difficult and equally related to the latent safety knowledge factor in men and women. Although the differences were not statistically significant, factor loadings by sex showed that items such as tractors being involved in fatal farm injuries, inattention being the leading cause of injury, and hearing loss being common in farm workers were more strongly correlated to the safety knowledge factor in women than in men. With the exception of falls being related to farm machinery use, men generally showed equal or weaker item loadings on the safety factor. Given that the number of injuries did not differ significantly between men and women, these differences do not seem to translate into a lower injury prevalence among women farm residents. This is contrary to a Canadian study where men were reported to experience many more fatal and nonfatal injuries than women and a larger proportion of injuries in those older than 60 years.[14] Nevertheless, the study looked only at fatal injuries and those requiring hospitalizations. In the present study, there was no difference in the number of serious injuries defined as having sought medical attention or lost at least 4 hours of work between men and women.

Significant differences in the difficulty of the items were observed in those younger and older, indicating measurement

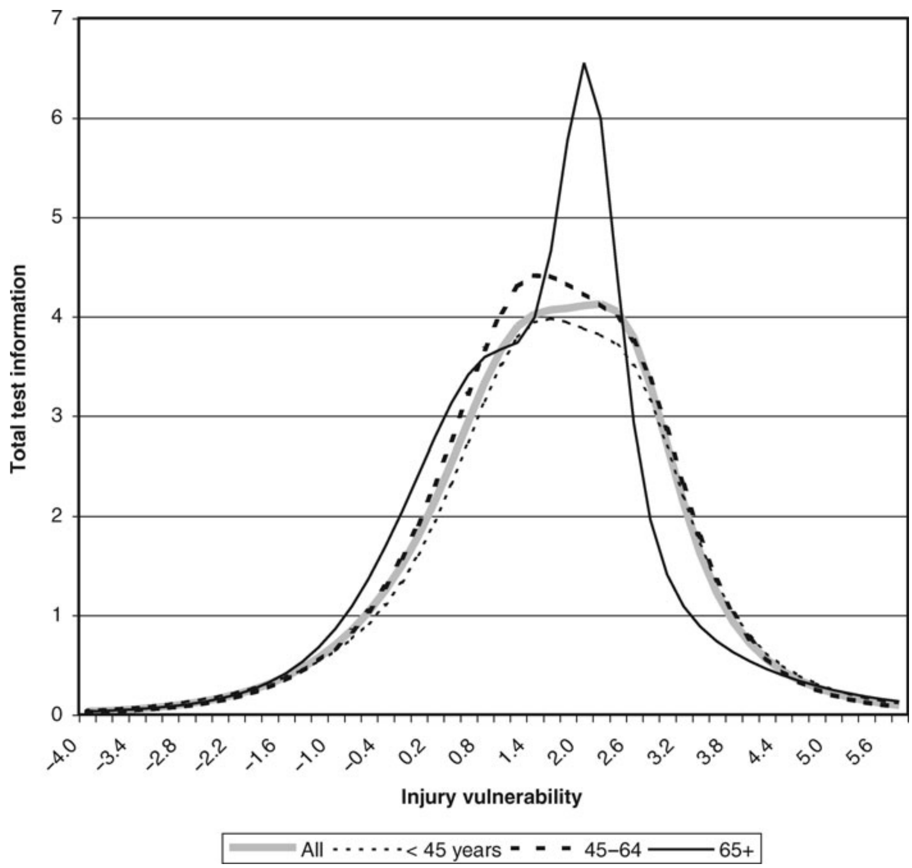

FIGURE 4. Total information curves comparing those under the age of 45 ($n = 678$), those 45 to 64 ($n = 732$), and those older than 65 years ($n = 220$). Taller curves represent greater information than shorter curves.

invariance and uniform DIF. Uniform DIF occurs when items are systematically more difficult for one group than for another even though they have the same level of safety knowledge. Uniform DIF occurs at all levels along the latent ability continuum. This type of DIF is analogous to confounding in epidemiology because a demographic characteristic associated with the item and with the safety factor act to change the effect estimate of the relationship between underlying ability and the response to an item.[15] Nonuniform DIF is analogous to effect modification and was not observed among the age groups in this sample.

Although the odds of an injury were highest for those younger than 45 years, there was no clear pattern of safety item endorsement that explained the differences in injury rates among the three age groups. Some items were more difficult for those in the younger group than those in the older group and some were not. Age differences were solely at the item level; no structural invariance suggesting population heterogeneity was identified. Nevertheless, the item level differences resulted in differences in the amount of information the safety items were providing among the age groups, as evidenced by the TIC. Comparing the Colorado sample of farm residents that were 55 years and older to the same age group in the study of Lizer and Petrea[3], where these safety questions were also administered, showed that three of the top four items that were answered incorrectly were identical. Comparing the Illinois sample with the Colorado sample, 45.2% of Colorado farm residents missed the question on falls causing injuries (Illinois, 29.2%), 31.2% missed the question on kids being injured by farm equipment (Illinois, 19%), 20.8% missed the causes of farmer's lung disease question (Illinois, 46.9%), and 22.3% missed animals causing injuries (Illinois, 12.9%). In those 55 years and older in Colorado, 21.2% missed the question about agriculture having a high injury rate. Of the four most frequently missed questions in the Illinois study, only lung disease was answered correctly more often by Colorado farm residents than by Illinois farm residents. It is interesting that of the ten questions, the same three questions were answered incorrectly in two very different farming states, but with different frequencies. Some of the discrepant results in the percentage of correct answers among states may reflect measurement non-invariance. Differential item functioning represents a lack of construct validity, and results may not be reproducible across farm resident populations of diverse ages. Further work on the psychometric properties of the safety knowledge scale and testing of additional questions to broaden how much of the safety knowledge continuum is covered would be valuable for safety training programs and evaluation of those programs.

A previous prospective study of respondents in the statewide survey showed that safety knowledge was only associated with injuries in the presence of depression, assessed using the Center for Epidemiological Studies Depression scale.[4] A possible explanation for this finding is that depression can cause feelings of fatigue and the inability to concentrate. If the safety questions are partly capturing a depressive state in addition to safety knowledge, then an interaction between these two risk factors could increase the risk of injury. Farm safety interventions may need to first address mood disorders in farmers, as well as focusing on safety behaviors and personal factors.

A limitation of IRT is the lack of methodology for validating the results by using external validators or risk factors known to be associated with injury. Such a validator might be hours of work per week. In this study, none of the specific safety items were significantly associated with injury, but there was limited power to detect an association and a lack of power to detect an interaction with depression in the subgroups. Despite this limitation, this study had ample power to detect differences in even the smallest age group of those 65 years and older.

In this study, we found that the ten-item safety knowledge test demonstrated significant measurement non-invariance by age, and therefore caution should be taken when applying this test across different age groups. The understanding or interpretation of the safety questions by age likely results in an overestimation of true mean differences when comparing farm residents of different ages.[16] There are many reasons why older farmers might answer these questions differently than their younger counterparts, including (1) length of exposure to farming and farm injuries; (2) farming cohort differences involving changes in farm practices; (3) stage of life differences (eg, being or having been a parent); and (4) access to and awareness of current farming knowledge and technologies (more likely in those who are younger). For example, those who have been farming longer may have greater awareness of the dangers of farm equipment to children, the hazards of agricultural work, the role of fatigue, and risk of amputation than those that have not been farming as long.[17] They also might underestimate the role of falls, tractors, and inattention in causing injury. No clear pattern emerged between the difficulty of the questions and the rate or number of injuries that occurred in different age groups, although the youngest group had significantly more injuries than the two older groups. There are likely to be many other risk factors that act in conjunction with an individual's level of safety knowledge that put a farm resident at risk of an injury.

In conclusion, further understanding of the role of safety knowledge in relation to farm injuries is needed to improve targeting prevention programs at those areas of knowledge that are most likely to influence the risk of injuries. On the basis of the results reported here, several areas should be targeted in safety education and training among farm residents, including the role of fatigue and inattention. In particular programs that address younger farm residents need to include modules that will increase awareness among these younger workers regarding the injury risk associated with working while fatigued or distracted. Programs in Colorado should also address some specific areas of safety knowledge including falls from equipment, children and equipment injuries, animal injuries, and the overall injury risk relating to working and living on farms.

REFERENCES

1. Bureau of Labor Statistics. Census of fatal occupational injuries summary, 2009. Available at: http://www.bls.gov/news.release/cfoi.nr0.htm. Accessed June 26, 2010.
2. Rautiainen RH, Lange JL, Hodne CJ, Schneiders S, Donham KJ. Injuries in the Iowa Certified Safe Farm Study. *J Agric Saf Health.* 2004;10:51–63.
3. Lizer SK, Petrea RE. Health and safety needs of older farmers. Part II. Agricultural injuries. *AAOHN J.* 2008;56:9–14.
4. Beseler CL, Stallones L. Safety knowledge, safety behaviors, depression, and injuries in Colorado farm residents. *Am J Ind Med.* 2010;53:47–54.
5. Lyman S, McGwin G, Enochs R, Roseman JM. History of agricultural injury among farmers in Alabama and Mississippi: prevalence, characteristics, and associated factors. *Am J Ind Med.* 1999;35:499–510.
6. Zhou C, Roseman JM. Agricultural injuries among a population-based sample of farm operators in Alabama. *Am J Ind Med.* 1994;25:385–402.
7. Stallones L, Leff M, Garrett C, Criswell L, Gillan T. Depressive symptoms among Colorado farmers. *J Agric Saf Health*. 1995;1:37–43.
8. Hays RD, Morales LS. Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care*. 2000;38:II28–II42.
9. Muthen B. A method for studying the homogeneity of test items with respect to other relevant variables. *J Educ Stat*. 1985;10:121–132.
10. Muthen B, Lehman J. Multiple group IRT modeling: applications to item bias analysis. *J Educ Stat*. 1985;10:133–142.
11. Langer MM, Hill CD, Thissen D, Burwinkle TM, Varni JW, DeWalt DA. Item response theory detects differential item functioning between healthy and ill children in QoL measures. *J Clin Epidemiol*. 2007;61:268–276.
12. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995;310:170.
13. Muthén BO, Muthén LK. *Mplus: Statistical Analysis With Latent Variables (Version 5.1)*. Los Angeles, CA: Muthén & Muthén Inc; 2008.
14. Dimich-Ward H, Guernsey JR, Pickett W, Rennie D, Hartling L, Brison RJ. Gender differences in the occurrence of farm related injuries. *Occup Environ Med*. 2004;61:52–56.
15. Crane PK, van Belle G, Larson EB. Test bias in a cognitive test: differential item functioning in the CASI. *Stat Med*. 2004;23:241–256.
16. Fleishman JA, Spector WD, Altman BM. Impact of differential item functioning on age and gender differences in functional disability. *J Gerontol: SOC SCI*. 2002;57B:S275–S284.
17. Lewis MQ, Sprince NL, Burmeister LF, Whitten PS, Torner JC, Zwerling C. Work-related injuries among Iowa farm operators: an analysis of the Iowa Farm Family Health and Hazard Surveillance Project. *Am J Ind Med.* 1998;33:510–517.