

Prediction of Vehicular Accident Involvement: A Meta-Analysis

Winfred Arthur, Jr.
Texas A&M University

Gerald V. Barrett and Ralph A. Alexander
The University of Akron

Previous attempts to summarize the vehicular accident involvement literature have been nonquantitative. Outcomes of these reviews have also reflected the equivocality of research in this area. In an attempt to synthesize the diverse research findings into a collective result, a meta-analysis procedure that controlled for sampling error was used. Four classes of variables were identified as predictors of vehicular accident involvement. These were information-processing, cognitive ability, personality, and demographic/biographical variables. Moderate-to-marginally favorable overall meta-analysis results were obtained for selective attention, regard for authority, locus of control, and cognitive ability as predictors of vehicular accident involvement. Suggestions and directions for future research are discussed.

The prediction of a person's likelihood of vehicular accident involvement has had a long, turbulent history; the arguments for and against the concept seem to swing from one extreme to the other (Shaw & Sichel, 1971). Early reviews of the literature, which considered a large assortment of individual variables, were rather pessimistic. For example, Goldstein (1962) only found small and nonsignificant correlations between categories of predictors and accident involvement. The correlations between reaction time and accident criteria in four studies using fairly large samples of drivers were found to be no higher than .17. Goldstein's review indicated that cognitive measures, paper-and-pencil, sensory-perceptual, and psychomotor tests all showed small correlations with accident criteria.

Some studies such as Barrett and Thornton (1968) and Barrett, Thornton, and Cabe (1969) identified measures exhibiting high relations with accidents in automobile simulators. These findings were subsequently

replicated in field studies (Mihal & Barrett, 1976). More recently, Avolio, Kroeck, and Panek (1985) obtained significant postdictive correlations of .13 to .43 between motor vehicle accident involvement and performance on six measures of information processing. Lower levels of performance on the information-processing measures were correlated with elevated levels of motor vehicle accident involvement. McKenna, Duncan, and Brown (1986), however, found small or no relations between performance on information-processing measures and accident involvement.

Given this background, the first objective of this article was to identify other classes of variables that might display greater usefulness as predictors of accident involvement from an applied-organizational perspective and a research perspective (particularly in reference to such personnel functions as selection and training). The means chosen to identify these variables was meta-analysis. *Meta-analysis* is a quantitative review technique that may be used to resolve conflicts between two or more studies. Unlike traditional reviews, the statistical analyses used in a quantitative review can find relationships and trends too subtle to be detected otherwise. Quantitative reviews also have the advantage of efficiently summarizing large volumes of literature. Thus, a quantitative review overcomes some deficiencies of traditional review procedures.

Meta-analysis procedures were used to resolve the conflicting results and conclusions drawn from the accident involvement research. Meta-analysis combines the results of previous discrete studies and uses them as a sample to reach a generalization. Schmidt and Hunter's (1977) validity generalization technique is one of several meta-analysis procedures. The Schmidt and Hunter procedure uses correlation coefficients as data points; in contrast, other techniques use effect sizes or differences among means (Hedges & Olkin, 1985). The Schmidt and Hunter procedure was the most appropriate technique because we were interested in the predictive validity of variables. The basic premise of the Schmidt and Hunter approach is that variance in test validities from one study or situation to another may be the result of certain statistical artifacts. To demonstrate the generalizability of test validities across situations, it is necessary to correct the variance of study results for the relevant statistical artifacts to obtain an overall corrected, observed test validity. If the percentage of variance accounted for is more than or equal to 75% (i.e., less than 25% unaccounted variance) after this corrective procedure, then the test validities generalize and are not considered to be situationally specific.

The second objective of this article was to apply the validity generalization technique to studies from the accident prediction literature to investigate whether the obtained validity coefficients were statistically meaningful. The variance in studies predicting accident involvement was hypothesized to be the result of statistical artifacts and errors identified by Schmidt and Hunter (1977). Finally, the results of the meta-analysis al-

lowed a comparative assessment of several classes of predictors used in this research domain and the determination of which one was "best" associated with elevated accident rates.

METHOD

A comprehensive and exhaustive search of the pertinent literature using *PsychLit* and *Dialog Computer Search* was undertaken. Non-English publications were excluded from the search. Upon reviewing the literature, we derived the following classification of variables used in the vehicle accident research: (a) information processing, (b) cognitive ability, (c) personality, and (d) biographical or demographic. Meta-analysis procedures ideally require many data points to obtain interpretable results; however, a very liberal minimum of 5 data points per predictor or variable was established as a cutoff for inclusion in this study. This relatively low minimum cutoff was necessary due to the small number of studies in each category. Using the classification just noted and the cutoff rule, the following specific predictors or variables were obtained.

Information-Processing and Cognitive Ability Variables

Three types of information processing variables—selective attention, perceptual style, and choice and complex reaction time—were analyzed. Selective attention was operationalized as scores on the Auditory Selective Attention Test (Arthur, Barrett, & Doverspike, 1990) or the Dichotic Listening Test (McKenna et al., 1986). Data points for selective attention comprised omission, intrusion, and switching errors. Perceptual style was operationalized using measures of field dependence–independence like the Group Embedded Figures Test (Avolio et al., 1985), the Portable Rod-and-Frame Test (Mihal & Barrett, 1976), and the Hidden Figures Test (e.g., Clement & Jonah, 1984). Cognitive ability was operationalized by tests such as the Culture Fair Intelligence Test (McKenna et al., 1986) and the Australian Council for Educational Research Higher Test (Smith & Kirkham, 1982). All correlations were transposed so that a positive correlation indicated better performance on both predictor and criterion.

Personality Variables

Whereas the literature search revealed a number of possible personality variables, we kept the distinct personality constructs separate (Day & Silverman, 1989). As a result, only four constructs had enough data points to be included in the meta-analyses. These constructs were level of distress, general activity level, regard for authority, and locus of control. Representing level of distress were the Depression scale and the Psychasthenia scale of the Minnesota Multiphasic Personality Inventory and the Neuroticism scale

TABLE 1
Number of Data Points (*rs*) per Predictor

<i>Predictor</i>	<i>Number of r's</i>
Selective attention	13
Perceptual style	12
Choice and complex reaction time	5
Cognitive ability	10
Age	8
Education	7
General activity level	40
Regard for authority, law, and norms	28
Level of distress	13
Locus of control	13
Total	149

of the Eysenck Personality Inventory (Loo, 1978). General activity level was represented by scales such as Aggression, Depression (Selzer & Vinokur, 1974), and Withdrawal (Mayer & Treat, 1977). Respect for Law (Clark, 1976), Juvenile Delinquency, and Antisocial Tendencies (Mayer & Treat, 1977) were scales representing regard for authority. Locus of Control was typically represented by scores on Rotter's Internality-Externality (I-E) Scale. Again, all correlations were transposed so that a positive correlation indicated better performance on both predictor and criterion.

Demographic and Biographical Variables

Although the literature search revealed many possible demographic and biographical variables, only age and education met the 5-data-point rule. As with the other variables, all correlations were transposed.

Using the just-noted classification scheme and 5-data-point cutoff, the literature search identified 149 usable data points from 32 articles. References for articles in which these studies were reported are presented in the Appendix. To be "usable," a study had to report sample sizes along with a correlation coefficient or some other statistic that could be converted to a correlation coefficient. Thus, studies that did not provide a numerical summary of their results were excluded. Also excluded were those studies that reported results in a format that could not be transformed into a correlation coefficient (e.g., studies that reported only a chi-square statistic).

Because the variable classifications were conceptually distinct, different studies employing these variables could not be combined for a common meta-analysis (Schmidt & Hunter, 1977). Consequently, separate meta-analyses were performed for each class of predictor variables. The numerical breakdown of data points by predictors is presented in Table 1. As has been true in other meta-analytic studies of validity coefficients, some of the

data points were nonindependent in the sense that several correlation coefficients were computed from data collected on a single group of subjects. When there were nonindependent correlation coefficients, we kept them separate rather than combining them. Averaging these correlations would have reduced the number of data points to such a small number that a meaningful analysis of several predictor categories would have been impossible. The potential effects of this decision are addressed later.

Our method initially required an overall meta-analysis (by predictor) of the pertinent vehicular accident prediction literature using the Schmidt and Hunter (1977) meta-analysis procedure. The Schmidt and Hunter procedure allows for a correction of variability in correlations due to: (a) sampling error, (b) sample-to-sample differences in reliability in predictor or criterion measures, and (c) sample-to-sample differences in range restriction. This study corrected for sampling error only because: (a) given the quality of microlevel reporting (Orwin & Cordray, 1985), there was no way of generating estimates to correct for either of the other two sources of artifactual error; and (b) past validity generalization research (Pearlman, Schmidt, & Hunter, 1980; Schmidt, Gast-Rosenberg, & Hunter, 1980; Schmidt & Hunter, 1981) demonstrated that sampling error alone accounts for at least 85% of the explained variance due to the just-mentioned artifacts.

To conduct the meta-analysis, correlation coefficients were first transformed to Fisher's- z values ($z_i = \tanh^{-1}r_i$; Alexander, Scozzaro, & Borodkin, 1989) and a variance-weighted average z was found by:

$$\bar{z} = \sum_{i=1}^k \frac{(n_i - 3)z_i}{\sum n_i - 3k}, \tag{1}$$

where r_i and n_i were the correlation coefficient and number of persons in the i th study and k = number of studies, respectively. The corresponding variance across studies was the frequency-weighted average squared error given by:

$$S_z^2 = \frac{k \sum n_i (z_i - \bar{z})^2}{(k - 1) \sum n_i}. \tag{2}$$

The sampling error variance was given by:

$$S_e^2 = \frac{\sum \frac{n_i}{n_i - 3}}{\sum n_i}. \tag{3}$$

A chi-square test for homogeneity of effect sizes was conducted as: $\chi^2_{(k-1)} = \sum (n_i - 3)(z_i - \bar{z})^2$.

Thus, the variance across combined studies was given by $S_o^2 = S_z^2 - S_e^2$. After correcting for sampling error variances across studies it was possible

to assess whether there was any true variance in the results across these studies (i.e., $S_e^2 \neq 0$). Finally, a 95% confidence interval was constructed on the \bar{z} value, and these were back transformed to r values ($r = \tanh z$).

When there were large amounts of variance across studies as evidenced by a statistically significant chi-square test, the possible influence of moderator variables was investigated as a plausible explanation of this variance. To test a hypothesized potential moderator variable, we used this variable to group the observed correlations into subsets. A moderator variable was identified by (a) a corrected variance that had a lower average in the subsets than for the data as a whole, and (b) an average correlation that varied from subset to subset. In brief, if large differences among subsets were found, then the hypothesized variable was considered to be a moderator variable. This meta-analysis began with the objective of classifying studies on several moderator variables, including gender, criteria time frame, and criteria type; but due to deficiencies in the quality of the microlevel reporting, this was not possible because the data were simply not available to permit these classifications. A moderator analysis was conducted only if there were at least 5 data points available for the given moderator. Unfortunately, we were simply unable to analyze as many moderator variables as originally planned.

RESULTS

The results of the meta-analysis are reported in Table 2. The second column of the table represents the total number of subjects across all studies used in the analysis. The third column represents the number of correlation coefficients used in the analysis, and the next column is the corrected mean of these correlation coefficients. The fifth and sixth columns are the range of observed correlation coefficients. The percentage of unaccounted variance is in the seventh column. The lower and upper limits of a 95% confidence interval about the mean correlation are presented in the eighth and ninth columns, respectively. Also, the percentage of unaccounted variance column contains the result of a chi-square test for homogeneity of sample correlations (Hunter, Schmidt, & Jackson, 1982). This test determines whether the unexplained variance in the correlations is significantly greater than zero. A significant chi-square value indicates that the unexplained variance is significantly greater than zero, suggesting that the remaining variance is due to additional moderators or statistical artifacts.

Note that, in Table 2, very favorable results (i.e., those that generalize) are indicated by a moderate to high corrected mean correlation, an unaccounted variance less than 25%, a nonsignificant chi-square, and a 95% confidence interval that did not include zero. Unfavorable results (i.e., those that do not generalize) are demonstrated by a low corrected mean correlation, an unaccounted variance greater than 25%, a significant chi-square value, and a 95% confidence interval that included zero.

TABLE 2
Results of Meta-Analysis for Predictors of Automobile Accidents

Predictor	N	Number of rs	Mean r	Range of Observed rs		% Variance Unaccounted for	95% Confidence Interval	
				Minimum	Maximum		Lower	Upper
Selective attention ^a	1,101	13	.257	-.07	.43	42.52*	.205	.317
Perceptual style (overall)	838	12	.151	-.04	.77	65.07**	.090	.225
Professional drivers ^b	525	9	.226	-.05	.77	63.50**	.152	.318
Choice and complex reaction time	5,406	22	.053	-.02	.39	53.33**	.026	.080
Cognitive ability	1,020	10	.117	.05	.40	0.00	.056	.179
Age (overall)	88,797	8	-.064	-.11	.39	72.34**	-.071	-.057
Professional drivers	515	5	.080	-.10	.39	69.87**	-.006	.167
Archival criterion data	3,312	5	.051	-.02	.40	69.52**	.017	.086
Education	1,001	7	.028	-.01	.07	0.00	-.035	.090
General activity level (overall)	7,137	40	.072	-.07	.42	50.28**	.049	.095
Nonprofessional drivers	6,463	30	.057	-.07	.42	43.52**	.032	.081
Professional drivers	674	10	.216	.00	.39	7.69	.145	.292
Archival criterion data	88,523	7	-.064	-.11	.39	75.03**	.071	.058
Self-report criterion data	1,035	14	.112	-.06	.42	33.61	.052	.174
Regard for authority (overall)	3,242	28	.155	-.07	.55	46.90**	.124	.192
Nonprofessional drivers	2,607	21	.124	-.07	.41	0.00	.086	.162
Professional drivers	635	7	.283	.04	.55	66.73**	.219	.364
Archival criterion data	2,208	14	.145	.04	.55	64.16**	.107	.190
Self-report criterion data	686	12	.175	-.07	.41	0.00	.103	.251
Level of distress	3,106	13	.023	-.04	.25	5.03	-.013	.058
Locus of control (overall)	1,909	13	.196	.02	.42	49.95*	.154	.242
Nonprofessional drivers	1,794	11	.193	.02	.42	56.79**	.151	.240
Archival criterion data	995	5	.273	.14	.32	0.00	.215	.331
Self-report criterion data	840	7	.102	.02	.42	10.67	.033	.169

Note. A significant chi-square indicates that the unexplained variance is significantly greater than zero.

^aThese studies all had a professional driver sample and used archival criterion data. ^bThese studies all used archival criterion data.

* $p < .05$. ** $p < .01$.

Each variable, along with potential moderators, is presented in the rows of the table. Nonprofessional drivers and professional drivers represent an analysis of those studies that used those types of drivers as subjects. Self-report criterion data studies, in contrast to archival criterion data studies, are those that used subjects' self-reports of accident criteria (e.g., having subjects report the number of accidents they had within the past 2 years).

Moderately favorable meta-analysis results were obtained for selective attention. As shown in Table 2, the corrected mean correlation for all selective attention studies was .257, with a 95% confidence interval that ranged from a lower limit of .205 to an upper limit of .317. These results also indicated that approximately 57% of the total variance was due to sampling error. Although a significant amount of unexplained variance remained (43%, and the chi-square was significant), the 57% of total variance accounted for was relatively high.

The meta-analysis results for perceptual style were not as favorable as those obtained for selective attention. For the overall analysis, sampling error accounted for only 35% of the total variance (65% of the variance was unaccounted). Moderator analyses improved the results only slightly (for professional drivers and archival data). Similar negative results were obtained for choice and complex reaction time.

Marginally favorable results were obtained for cognitive ability. Although a large percentage of variance was accounted for, the mean correlation was very small and the lower limit of the 95% confidence interval (.056) was extremely close to zero. For the demographic and biographical variables of age and education, the results were unfavorable and, therefore, did not support the validity generalization hypothesis. Even for education, which accounted a large amount of variance, the mean correlation was close to zero, and the 95% confidence interval included zero.

Of the personality variables, the most favorable results were obtained for general activity level among professional drivers (mean $r = .216$) and locus of control with archival data (mean $r = .273$). Marginally favorable results were obtained on the self-report-criterion-data moderator for general activity level, regard for authority, and locus of control. The results for all other personality variables were unfavorable and did not support the validity generalization hypothesis. The variance accounted for was generally minimal, coupled with significant chi-squares and/or 95% confidence intervals that included zero.

CONCLUSIONS

Several conclusions can be drawn from the results of this meta-analysis. First, the magnitude of the relation between valid predictors and accident criteria was consistent with other meta-analyses predicting job perform-

ance. These relations were obtained after the variances of coefficients were corrected for sampling error. Second, moderate to marginally favorable overall meta-analysis results were obtained for selective attention, regard for authority, locus of control, and cognitive ability. Better selective attention, higher regard for authority, an internal locus of control, and higher cognitive ability were associated with lower levels of accident involvement. Third, several significant moderators were present. Professional drivers, in contrast to nonprofessional drivers, significantly moderated the relations between accident involvement and the following predictors: perceptual style, general activity level, and regard for authority. The type of criterion data (i.e., self-report vs. archival) also significantly moderated the relation between automobile accidents and general activity level, regard for authority, and locus of control.

Although favorable results were obtained for some predictors, a general factor tended to weaken the results and bias them against validity generalization. Neither the predictors nor the criteria were identical across studies (a common predicament in the field of accident research). For the information-processing variables, three different measures of perceptual style (the Group Embedded Figures Test, the Portable Rod-and-Frame Test, and the Hidden Figures Test) and two measures of selective attention (the Auditory Selective Attention Test and the Dichotic Listening Test) were used in the primary studies. This diversity of measures acts as a source of irrelevant variance; therefore, the results of the meta-analysis are weakened because the same measures are not being used. This makes our results conservative.

Studies using information-processing measures had a decided advantage over those using other predictors. Standardized instruments are common in most information-processing research. This is in sharp contrast to the range and diversity of measures used in the assessment of personality variables, even when the analysis is limited to a single construct (Day & Silverman, 1989). Also, the measures of personality and demographic/biographical variables were more unstructured when compared to those used with information-processing variables. Personality researchers have used questionnaires, inventories, and tests specifically developed for the study in question (e.g., Clark, 1976; Schuster & Guilford, 1964).

A meta-analysis ideally calls for several hundred data points (e.g., Schmitt, Gooding, Noe, & Kirsch, 1984). This goal is, of course, difficult to meet; this study was no exception. The primary reason why less than the optimal number of data points were analyzed was that many studies failed to report sufficient data to permit their inclusion in the meta-analysis. Although this problem can be remedied if the necessary information is published in future studies, the results obtained in this meta-analysis need to be cautiously interpreted. Specifically, as the number of studies in a validity generalization study decreases, the likelihood of sampling error increases. Consequently, it becomes more likely to obtain favorable validity general-

ization results than to obtain unfavorable results. Because meta-analysis results obtained from a small number of studies are biased in favor of the Type I errors, they could be very misleading and must be carefully interpreted. This is particularly true when moderator analyses are conducted because the number of correlation coefficients are usually drastically reduced to a relatively small number (Schmidt & Hunter, 1978).

Another issue of concern in interpreting the results of this meta-analysis is the nonindependence of correlation coefficients. As previously mentioned, we decided not to average nonindependent data points because it would have reduced the number of correlation coefficients to a number too few to permit any meaningful analysis. The most likely result of this decision to not combine nonindependent correlation coefficients is that nonindependence of data points would reduce the observed variability of the correlations. Therefore, interpretations of the homogeneity of effect sizes (chi-square) must be more cautious. This outcome, however, does not affect other conclusions.

A comparison of our results with those of Goldstein (1962) is warranted. Goldstein concluded that "accident records are only slightly predictable from measures of other stable human characteristics such as visual acuity, reaction time, sensory, psychomotor, cognitive, and attitudinal measures" (p. 5). Subsequently, Goldstein's conclusions have been frequently cited in support of the nonpredictability of the likelihood of accident involvement (e.g., Harano, Peck, & McBride, 1975; Panek & Rearden, 1987; Pelz & Krupat, 1974). Cognitive ability and reaction time variables were common to both this article and Goldstein's. The meta-analysis results for cognitive ability corroborated Goldstein's (1962) conclusion. Although correcting for sampling error accounted for 100% variance (on only 10 data points), the corrected mean correlation was only .117, and the 95% confidence interval ranged from .056 to .179. In addition, the meta-analysis results for choice and complex reaction time did not favor generalization. It must be pointed out, however, that the studies reviewed by Goldstein used initial and simple reaction time. Unfortunately, there were not enough data points to permit a meta-analysis of those variables.

Goldstein (1962) did not specifically discuss personality variables *per se*, but if "attitudinal measures" are considered as such, then the agreement between our results and Goldstein's conclusions are mixed. Partially favorable results were obtained for regard for authority, general activity level, and locus of control, and but not for level of distress.

A primary goal of this study was to identify additional variables (specifically those based on intrinsic attributes) associated with accident involvement, which might have greater utility as predictors from an applied organizational and research perspective. The results of the meta-analysis identified the information-processing variable of selective attention as one such variable. Our results indicate that specific ability measures, such as the

Auditory Selective Attention Test, predicted accident involvement just as well, if not better than general cognitive ability, which is consistent with the findings of Arthur et al. (1990). The meta-analysis results obtained for selective attention (corrected mean correlation of .257) are consistent with other meta-analytic studies. For example, Schmitt et al. (1984) obtained a mean corrected correlation of .162 for special aptitude, .206 for personality, .220 for general mental ability, .317 for biodata, .319 for work samples, and .428 for assessment centers as predictors of job performance ratings. Similar ranges of mean corrected correlations were obtained by Pearlman et al. (1980). and Schmidt et al. (1980) in validity generalization studies of clerical occupations and computer programmers, respectively. Thus, the mean corrected correlation of .257 obtained for selective attention is as good as that obtained for most other occupations.

The applied utility of information-processing variables was demonstrated by Sterns, Barrett, and Alexander (1980), whose results indicate that perceptual information-processing skills of older adults can be improved through training. Furthermore, these effects were not due solely to practice. Sterns et al.'s research demonstrated that training older adults using individual training sessions significantly increased their performance on critical driving tasks. The implication is that information-processing skills such as selective attention not only predict vehicular accident involvement but are also trainable.

Suggestions for Future Research

The results of this meta-analysis highlight several issues that may serve as guidelines for future accident involvement research. These issues can be classified into two general categories: (a) the criterion, and (b) the predictors. The major criterion concern has to do with the operationalization and measurement of vehicular accident involvement. Although a distinction was made between archival and self-report accident measures in this meta-analysis, it is uncertain whether these were defined in the same way across the primary studies. The differences in definitions of vehicular accident involvement was epitomized by McKenna et al. (1986) who included passenger falls and attacks on staff in bus drivers' accident scores. Future research must make distinctions between driving accidents and nondriving accidents, all accidents versus at-fault accidents only, and preventable versus nonpreventable accidents. The importance of these distinctions was demonstrated by Parker (1953) who found a different pattern of tests to be predictive of preventable and nonpreventable accidents for commercial truck drivers. These distinctions are also important from a personnel or organizational perspective in avoiding contaminated criterion measures in decision making.

Criterion time frames (i.e., the time period over which accident data are

collected—past year, past decade, lifetime, etc.) are also of concern. This meta-analysis initially sought to investigate criterion time frames as a potential moderator variable because it was hypothesized that longer time frames would result in an increased base rate. Subsequently, a higher base rate would, in turn, potentially elevate the validity of the predictors. However, due to insufficient data points, we were unable to run these moderator analyses. Also, with longer criterion time frames, the periodic collection of accident data over time would permit an assessment of the reliability of accidents. In addition, Sichel (1965) suggested that the time interval between accidents is a better criterion than the number of accidents.

Associated with the criterion-time-frames issue is the observation that most of the studies analyzed in this meta-analysis used postdictive designs. The popularity of postdictive designs in accident research can be traced to the nature of the criterion, which typically requires relatively long time periods for its manifestation. Consequently, postdictive designs seem most practical and feasible. We must, however, call for more predictive designs, primarily because these designs require researchers to be more involved and proactive in the accident data-collection process. A common problem with postdictive designs, especially those using archival data, is that the data are typically collected for unscientific purposes (e.g., police record keeping). Because the researcher has no control over how data is collected, the quality of the data is often unknown (Maier, 1988).

Given the difficulties associated with “real world” accident data collection, some researchers have turned to performance simulations. For example, Ranney and Pulling (1989, 1990) used a driving performance work sample consisting of 30-min “trips,” with each trip composed of 20 laps on a closed course. The driver was required to respond to a continuous sequence of situations which included responding to traffic signals and signs, route selection, and gap-acceptance tasks. Although they are infrequently used as a criterion in the personnel selection literature, simulations have numerous advantages over the collection of field accident data, including the ability to compress time and permit a finer operationalization and measurement of accidents (Harmon, 1961; Poulton, 1963). We initially sought to investigate criterion type (i.e., performance simulations vs. real world accidents) as a possible moderator, but again there were not enough data points to permit this analysis. Nevertheless, given the refinements in criterion measurement that are possible, the use of performance simulations is encouraged and should be considered whenever feasible.

The second category of future research suggestions has to do with the predictors that have been used in vehicular accident literature. A primary issue concerns the use of an enormous diversity of tests with varying convergent validities and reliabilities to measure ostensibly the same construct. This introduces a source of irrelevant variance. For instance, both the Group Embedded Figures Test and the Portable Rod-and-Frame Test have

been used as measures of perceptual style. There is evidence to suggest, however, that although these tests may be similar, they do not measure exactly the same construct (e.g., Arthur & Day, in press; Witkin & Goode-nough, 1981). The same arguments are pertinent to the use of the complex versus simple reaction times and the Auditory Selective Attention Test and other dichotic listening tests. Future studies should strive to refine and use measures that best reflect the construct of interest, such as the Auditory Se-lective Attention Test for selective attention, the Portable Rod-and-Frame Test for perceptual style, and Rotter's I-E Scale for locus of control.

Predictor categories have too often been conceptualized as though they were mutually exclusive. Previous studies have typically looked at either in-formation-processing, personality, or demographic predictors with little or no attempts made to investigate how combinations of predictors across cat-egories may improve predictability. For example, the theoretical rationale for a relation between locus of control and safe driving is based on the rea-soning that an external locus of control is related to a lack of caution and failure to take precautionary steps to avoid the occurrence of unfavorable outcomes (Hoyt, 1973; Phares, 1976; Strickland, 1977, 1978; Williams, 1972). However, from an information-processing perspective, driving a car is considered to be a task that calls for the perception, identification, proc-essing, and adequate responding to pertinent information (e.g., traffic lights, signs, pedestrians, other vehicles, etc.) in the environment. It seems reasonable to hypothesize that a theoretical and empirical combination of both perspectives may result in incremental predictive validity over and above any single predictor (Hansen, 1989).

As a concluding comment, we reiterate that a conscientious effort must be made to report all pertinent information and data in future studies. Such information should include, but is not limited to, the pertinent test statistic (e.g., r , t , or F); sample sizes, means, and standard deviations; demo-graphic and other characteristics of the sample; operationalization and psy-ychometric properties, such as reliability, validity, and range restriction of predictors; definitions and descriptions, such as criterion type, criterion time frames, and measurement of the criterion along with any relevant psy-ychometric information like reliability and range restriction; and the re-search design. Such information will facilitate the inclusion of more studies in future meta-analyses and permit the investigation of potential moderator variables that we were unfortunately unable to assess.

ACKNOWLEDGMENTS

We thank David Day and two anonymous reviewers for their helpful com-ments on earlier drafts of this article and Juanicki Morrow for her assis-tance on numerous revisions.

REFERENCES

- Alexander, R. A., Scozzaro, M., & Borodkin, L. (1989). Statistical and empirical examination of the chi-square test for homogeneity of correlations in meta-analysis. *Psychological Bulletin*, *106*, 329-331.
- Arthur, W., Jr., Barrett, G. V., & Doverspike, D. (1990). Validation of an information-processing based test battery for the prediction of handling accidents among petroleum product transport drivers. *Journal of Applied Psychology*, *75*, 621-628.
- Arthur, W., Jr., & Day, D. V. (in press). Examination of the construct validity of alternative measures of field dependence/independence. *Perceptual Motor Skills*.
- Avolio, B. J., Kroeck, K. G., & Panek, P. E. (1985). Individual differences in information-processing ability as a predictor of motor vehicle accidents. *Human Factors*, *27*, 577-588.
- Barrett, G. V., & Thornton, C. L. (1968). Relationship between perceptual style and driver reaction to an emergency situation. *Journal of Applied Psychology*, *52*, 169-176.
- Barrett, G. V., Thornton, C. L., & Cabe, P. A. (1969). Relation between Embedded Figures Test performance and simulator behavior. *Journal of Applied Psychology*, *53*, 253-254.
- Brown, P. L., & Berdie, R. G. (1960). Driver behavior and scores on the MMPI. *Journal of Applied Psychology*, *44*, 18-21.
- Clark, A. W. (1976). A social role approach to driver behavior. *Perceptual and Motor Skills*, *42*, 325-326.
- Clement, R., & Jonah, B. A. (1984). Field dependence, sensation seeking, and driving behavior. *Personality and Individual Differences*, *5*, 87-93.
- Day, D. V., & Silverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. *Personnel Psychology*, *42*, 25-36.
- Goldstein, L. G. (1962). *Human variables in traffic accidents: A digest of research and selected bibliography*. Washington, DC: National Academy of Sciences, National Research Council.
- Hansen, C. P. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. *Journal of Applied Psychology*, *74*, 82-90.
- Harano, R. M., Peck, R. C., & McBride, R. S. (1975). The prediction of accident liability through biographical data and psychometric tests. *Journal of Safety Research*, *7*, 16-52.
- Harmon, H. H. (1961). *Simulation: A survey* (Report No. SP-260). Santa Monica, CA: System Development Corporation.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego: Academic.
- Hoyt, M. F. (1973). Internal-external control and beliefs about automobile travel. *Journal of Research in Personality*, *7*, 288-293.
- Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills: Sage.
- Loo, R. (1978). Individual differences and the perception of traffic signs. *Human Factors*, *20*, 65-74.
- Maier, M. H. (1988). On the need for quality control in validation research. *Personnel Psychology*, *41*, 497-502.
- Mayer, R. E., & Treat, J. R. (1977). Psychological, social and cognitive characteristics of high-risk drivers: A pilot study. *Accident Analysis and Prevention*, *9*, 1-8.
- McKenna, F. P., Duncan, J., & Brown, I. D. (1986). Cognitive abilities and safety on the road: A re-examination of individual differences in dichotic listening and search for embedded figures. *Ergonomics*, *29*, 649-663.
- Mihal, W., & Barrett, G. V. (1976). Individual differences in perceptual information processing and their relation to automobile accident involvement. *Journal of Applied Psychology*, *61*, 229-233.

- Orwin, R. G., & Cordray, D. S. (1985). Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. *Psychological Bulletin*, *97*, 134-147.
- Panek, P. E., & Rearden, J. J. (1987). Age and gender effects on accident types for rural drivers. *The Journal of Applied Gerontology*, *6*, 332-346.
- Parker, J. W. (1953). Psychological and personal history data related to accident records of commercial truck drivers. *Journal of Applied Psychology*, *37*, 317-320.
- Pearlman, D., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, *65*, 373-406.
- Pelz, D. C., & Krupat, E. (1974). Caution profile and driving record of undergraduate males. *Accident Analysis and Prevention*, *6*, 45-58.
- Phares, E. J. (1976). *Locus of control in personality*. Morristown, NJ: General Learning Press.
- Poulton, E. C. (1963). On increasing the sensitivity of measures of performance. *Ergonomics*, *8*, 69-76.
- Ranney, T. A., & Pulling, N. H. (1989). Relation of individual differences in information-processing ability to driving performance. *Proceedings of the Human Factors Society 33rd Annual Meeting*, *2*, 965-969.
- Ranney, T. A., & Pulling, N. H. (1990, January). *Performance differences on driving and laboratory tasks between drivers of different ages*. Paper presented at the 69th annual meeting of the Transportation Research Board, Washington, DC.
- Schmidt, F. L., Gast-Rosenberg, I., & Hunter, J. E. (1980). Validity generalization results for computer programmers. *Journal of Applied Psychology*, *65*, 643-661.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, *62*, 529-540.
- Schmidt, F. L., & Hunter, J. E. (1978). Moderator research and the law of small numbers. *Personnel Psychology*, *31*, 215-232.
- Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist*, *36*, 1128-1137.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, *37*, 407-422.
- Schuster, D. H., & Guilford, J. S. (1964). The psychometric prediction of problem drivers. *Human Factors*, *6*, 393-421.
- Selzer, M. L., & Vinokur, A. (1974). Life events, subjective stress and traffic accidents. *American Journal of Psychiatry*, *131*, 903-906.
- Shaw, L., & Sichel, H. S. (1971). *Accident proneness*. New York: Pergamon.
- Sichel, H. S. (1965). The statistical estimation of individual accident liability. *Traffic Safety Research Review*, *9*, 8-15.
- Smith, D. I., & Kirkham, R. W. (1982). Relationship between intelligence and driving record. *Accident Analysis and Prevention*, *14*, 439-442.
- Sterns, H. L., Barrett, G. V., & Alexander, R. A. (1980, November). *Older adult skills related to driving: Individual and group training*. Paper presented at the 33rd annual scientific meeting of the Gerontological Society of America, San Diego.
- Strickland, B. R. (1977). Internal versus external control of reinforcement. In T. Blass (Ed.), *Personality and social behavior* (pp. 219-279). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Strickland, B. R. (1978). Internal-external expectancies and health-related behaviors. *Journal of Consulting and Clinical Psychology*, *46*, 1192-1211.
- Williams, A. F. (1972). Factors associated with seat belt use in families. *Journal of Safety Research*, *4*, 133-138.
- Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive styles: Essence and origins*. New York: International Universities Press.

APPENDIX

Studies in Accident Involvement Meta-Analysis

- Arthur, W., Jr., Barrett, G. V., & Doverspike, D. (in press). Validation of an information-processing based battery for the prediction of handling accidents among petroleum product transport drivers. *Journal of Applied Psychology*.
- Avolio, B. J., Kroeck, K. G., & Panek, P. E. (1985). Individual differences in information-processing ability as a predictor of motor vehicle accidents. *Human Factors*, *27*, 577-588.
- Barrett, G. V., Mihal, W. L., Panek, P. E., Sterns, H. L., & Alexander, R. A. (1977). Information-processing skills predictive of accident involvement for younger and older commercial drivers. *Industrial Gerontology*, *4*, 173-182.
- Bransford, T. (1939). *Relation of performance on drivers' tests to automobile accidents and violations of traffic regulations in the District of Columbia*. Unpublished doctoral dissertation, American University, Washington, DC.
- Brown, C. W., & Ghiselli, E. G. (1947). Factors related to the proficiency of motor coach operators. *Journal of Applied Psychology*, *31*, 447-479.
- Brown, P. L., & Berdie, R. F. (1960). Driver behavior and scores on the MMPI. *Journal of Applied Psychology*, *44*, 18-21.
- Clark, A. W. (1976). A social role approach to driver behavior. *Perceptual and Motor Skills*, *42*, 325-326.
- Clark, A. W., & Prolisko, A. (1979). Social role correlates of driving accidents. *Human Factors*, *21*, 655-659.
- Clement, R., & Jonah, B. A. (1984). Field dependence, sensation seeking and driving behavior. *Personality and Individual Differences*, *5*, 87-93.
- Cobb, P. W. (1939). *Automobile driver tests administered to 3663 persons in Connecticut, 1936-1937, and the relation of tests scores to the accidents sustained*. Unpublished report to Highway Research Board, Washington, DC.
- Evans, G. W., Palsane, M. N., & Carrere, S. (1987). Type A behavior and occupational stress: A cross-cultural study of blue-collar workers. *Journal of Personality and Social Psychology*, *52*, 1002-1007.
- Guastello, S. J., & Guastello, D. D. (1986). The relation between the locus of control construct and involvement in traffic accidents. *The Journal of Applied Psychology*, *120*, 293-297.
- Kahneman, B., Ben-Ishai, R., & Lotan, M. (1973). Relation of a test of attention to road accidents. *Journal of Applied Psychology*, *58*, 113-115.
- Loo, R. (1978). Individual differences and the perception of traffic signs. *Human Factors*, *20*, 65-74.
- Mayer, R. E., & Treat, J. R. (1977). Psychological, social and cognitive characteristics of high-risk drivers: A pilot study. *Accident Analysis and Prevention*, *9*, 1-8.
- McKenna, F. P., Duncan, J., & Brown, I. D. (1986). Cognitive abilities and safety on the road: A re-examination of individual differences in dichotic listening and search for embedded figures. *Ergonomics*, *29*, 649-663.
- Mihal, W., & Barrett, G. V. (1976). Individual differences in perceptual information processing and their relation to automobile accident involvement. *Journal of Applied Psychology*, *61*, 229-233.
- Moffie, D. J., Symmes, A., & Milton, C. R. (1952). *Relation between psychological tests and driver performance* (Bulletin No. 60). Washington, DC: Highway Research Board, National Research Council.
- Montag, I., & Comrey, A. L. (1987). Internality and externality as correlates of involvement in fatal driving accidents. *Journal of Applied Psychology*, *72*, 339-343.
- Pandey, R. P., Kishore, G. S. R., & Jha, S. (1981). Some attitudinal correlates of traffic accidents. *Asian Journal of Psychology and Education*, *7*, 44-48.

- Panek, P. E., Wagner, E. E., Barrett, G. V., & Alexander, R. A. (1978). Selected Hand Test variables related to accidents in female drivers. *Journal of Personality Assessment*, 42, 355-357.
- Parker, J. W. (1953). Psychological and personal history data related to accident record of commercial truck drivers. *Journal of Applied Psychology*, 37, 317-320.
- Peck, R. C., & Kuan, J. (1983). A statistical model of individual accident risk prediction using driver record, territory and other biographical factors. *Accident Analysis and Prevention*, 15, 371-393.
- Pestonjee, D. M., & Singh, U. B. (1980). Neuroticism-extraversion as correlates of accident occurrence. *Accident Analysis and Prevention*, 12, 201-204.
- Roy, G. S., Choundhary, R. K., & Mithila, L. N. (1985). Driver control as a factor in road safety. *Asian Journal of Psychology and Education*, 16, 33-37.
- Selzer, M. L., & Vinokur, A. (1974). Life events, subjective stress and traffic accidents. *American Journal of Psychiatry*, 131, 903-906.
- Schuster, D. H., & Guilford, J. S. (1964). The psychometric prediction of problem drivers. *Human Factors*, 6, 393-421.
- Smith, D. I. (1976). Official driver records and self-reports as sources of accident and conviction data for research purposes. *Accident Analysis and Prevention*, 8, 207-211.
- Smith, D. I., & Kirkham, R. W. (1981). Relationship between some personal characteristics and driving record. *British Journal of Social Psychology*, 20, 229-231.
- Smith, D. I., & Kirkham, R. W. (1982). Relationship between intelligence and driving record. *Accident Analysis and Prevention*, 14, 439-442.
- Williams, J. R. (1971). *Relationships between three-dimensional spatial relations ability and driving performance*. Unpublished doctoral dissertation, New York University, New York.
- Williams, J. R. (1977). Follow-up study of relationships between perceptual style measures and telephone company vehicle accidents. *Journal of Applied Psychology*, 62, 751-754.