

METHODOLOGICAL DEVELOPMENTS

Estimating Individual Influences of Behavioral Intentions: An Application of Random-Effects Modeling to the Theory of Reasoned Action

Donald Hedeker and Brian R. Flay
University of Illinois at Chicago

John Petraitis
University of Alaska, Anchorage

Methods are proposed and described for estimating the degree to which relations among variables vary at the individual level. As an example of the methods, M. Fishbein and I. Ajzen's (1975; I. Ajzen & M. Fishbein, 1980) theory of reasoned action is examined, which posits first that an individual's behavioral intentions are a function of 2 components: the individual's attitudes toward the behavior and the subjective norms as perceived by the individual. A second component of their theory is that individuals may weight these 2 components differently in assessing their behavioral intentions. This article illustrates the use of empirical Bayes methods based on a random-effects regression model to estimate these individual influences, estimating an individual's weighting of both of these components (attitudes toward the behavior and subjective norms) in relation to their behavioral intentions. This method can be used when an individual's behavioral intentions, subjective norms, and attitudes toward the behavior are all repeatedly measured. In this case, the empirical Bayes estimates are derived as a function of the data from the individual, strengthened by the overall sample data.

Psychological theories often posit relations among variables that can vary depending on the individual. Thus, it may be stated that on average, a given variable X influences another variable Y to a certain degree, but that the degree of this relationship is not a constant but varies to some degree in the population of individuals. In other words, a relationship between two variables may be quite strong in some individuals, whereas for others it may be weak or nonexistent. In clinical research, for example, the effect of a given therapeutic agent for treating, say, depression is thought to vary to some degree from individual to individual; in some individuals, the therapeutic agent, given at the same dose and under the same conditions, is effective whereas for others it is not. Most traditional statistical methods, on the other hand, provide estimation of relations among variables on a group and not on an individual level. A

statistical analysis may thus reveal to the clinician that, on average, individuals given a certain amount of drug elicit a given degree of improvement over time. More sophisticated analysis may indicate that, on average, the relationship between drug levels and improvement over time is, say, quite strong for male participants in their 20s and weak for female participants age 30 and over. What is missing from this is the notion that these relations among variables can vary not just for different groups of individuals but for individuals themselves. Furthermore, it is important not just to know whether relations among variables do vary by individuals but to assess the degree to which these relations may vary from individual to individual; that is, how much fluctuation (or variation) is there in a given relationship in a population of individuals.

One such psychological theory that posits individual differences is Fishbein and Ajzen's (1975; Ajzen & Fishbein, 1980) theory of reasoned action (TRA), which has been one of the most elegant and influential theories of behavior. According to TRA, all behaviors are based on behavioral intentions. In fact, TRA asserts that the only immediate cause for any behavior is an individual's intentions to engage in or refrain from that behavior. In turn, TRA asserts that intentions are determined by two components: an individual's attitudes toward the behavior and an individual's perceptions of the social pressures or subjective norms to engage in or refrain from that behavior. When applied to cigarette smoking, for example, TRA predicts that people who have some intention to start smoking should be more likely to smoke in the future than those who have no intentions to smoke. Furthermore, people should have stronger intentions to smoke if they hold relatively positive attitudes toward smoking and feel some social pressure to smoke.

Not only is TRA elegant, its central predictions have been

Donald Hedeker and Brian R. Flay, Prevention Research Center and Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago; John Petraitis, Department of Psychology, University of Alaska, Anchorage.

Preparation of this article was supported by National Institute of Mental Health Grant MH44826-01A2, National Institute of Drug Abuse Grant DA06307, and University of Illinois at Chicago Prevention Research Center Developmental Project—Centers for Disease Control Grant R48/CCR505025.

We thank R. Darrell Bock for many valuable discussions, comments, and suggestions during the preparation of this article. We also thank Ohidul Siddiqui for computer programming assistance.

Correspondence concerning this article should be addressed to Donald Hedeker, Division of Epidemiology and Biostatistics (M/C 922), School of Public Health, University of Illinois at Chicago, 2121 West Taylor Street, Room 510, Chicago, Illinois 60612-7260.

widely supported. In a meta-analytic review of 87 studies, Shepard, Hartwick, and Warshaw (1988) found that the average correlations between intentions and behaviors was above .50 and that the average correlation between intentions and its predictors, attitudes, and subjective norms was above .65. Moreover, TRA has been remarkably versatile, predicting monumental behaviors (e.g., having an abortion, reenlisting in the National Guard, and smoking marijuana), as well as relatively inconsequential behaviors (e.g., watching a rerun of a television program, purchasing a particular brand of grape soda, and making a sandwich).

Although research evidence for TRA's central predictions is impressive, there is an ancillary prediction concerning individual differences that has never been adequately tested. That prediction concerns the relative importance of a person's (a) attitudes toward a behavior and (b) subjective perceptions of norms in determining (c) behavioral intentions. As depicted in Figure 1, the theory has always allowed for the possibility that individuals might weigh their attitudes toward the behavior (ATBs) and subjective norms (SNs) differently when formulating their intentions to engage in a behavior. If this prediction is accurate, we can better model individual behavioral intentions (BI) by assessing the relative weights of ATB and SN at the level of the individual.

Unfortunately, as alluded to earlier, most traditional statistical analyses of TRA would allow for estimation of group-level but not individual-level weights for these two influences of BI. Attempting to get around this difficulty, some researchers have had individual research participants report whether their attitudes or their subjective norms are more important in shaping their intentions. However, the use of self-report measures of weights has proved unsatisfactory when assessing individual-level weights (Ajzen & Fishbein, 1980). More commonly then, group-level weights have been estimated with a linear regression model using ordinary least squares (OLS) methods. However, as mentioned, this method only indicates the relative influence of ATBs and SNs for a group of participants, rather than indicating their influence for individual participants.

Another approach is possible when individuals are repeatedly assessed for their BI, ATB, and SN about the behavior. In this case, individual weights can be obtained using a random-effects regression model and, specifically, empirical Bayes estimation of the individual weights. Empirical Bayes techniques have proved useful in a variety of settings, including estimation of individual trend lines in longitudinal studies (Hui & Berger,

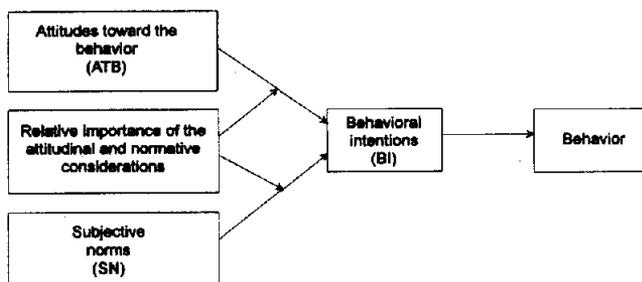


Figure 1. Theory of reasoned action.

1983), estimation of treatment effects for each center in multicenter clinical trials (Louis, 1991), and estimation of the influence of law schools' admissions data on law school success for each school (Rubin, 1980). The present paper will describe how the empirical Bayes methods can be utilized to estimate individual weights for the effects of ATB and SN on BI. By doing so, we will examine the hypothesis that individuals weigh their ATB and SN differently when formulating their intentions to engage in a behavior.

The statistical methods that will be described and used to examine the hypothesis of individual weighing of ATB and SN on BI are not new, but have been developed under a variety of names, including: random-effects models (Laird & Ware, 1982; Ware, 1985), variance component models (Dempster, Rubin, & Tsutakawa, 1981; Harville, 1977); hierarchical linear models (Bryk & Raudenbush, 1987; Raudenbush & Bryk, 1986), multilevel models (Goldstein, 1987), two-stage models (Bock, 1989a; Vacek, Mickey, & Bell, 1989), random coefficient models (DeLeeuw & Kreft, 1986), mixed models (Longford, 1986, 1987), empirical Bayes models (Hui & Berger, 1983; Strenio, Weisberg, & Bryk, 1983), unbalanced repeated measures models (Jennrich & Schluchter, 1986), and random regression models (Bock, 1983a, 1983b; Gibbons, Hedeker, Waternaux, & Davis, 1988; Hedeker, Gibbons, Waternaux, & Davis, 1989). Generally, as we illustrate in this article, these methods involve linear regression models that allow the possibility that some parameters besides the residuals are to be considered random, and not fixed. In addition to the articles cited earlier, a review article (Raudenbush, 1988), three book-length texts (Bryk & Raudenbush, 1992; Goldstein, 1987; Longford, 1993) and two anthologies (Bock, 1989b; Raudenbush & Willms, 1991) further describe and illustrate the use of these statistical models.

Method

Before we introduce the estimation of individual weights, let us describe estimation of weights on a group level. For this, a linear regression model can be used to relate an individual's behavioral intentions in terms of his or her ATBs and SNs as follows:

$$BI_i = \beta_0 + \beta_1 ATB_i + \beta_2 SN_i + \epsilon_i \quad (1)$$

where $i = 1, 2, \dots, N$ individuals. In this model, the regression coefficients β_1 and β_2 represent the weights given to ATB and SN, respectively, in determining BI, whereas ϵ_i represents the model residuals for each individual, and β_0 represents the model intercept. The weights for ATB and SN can be estimated based on a sample of individuals using OLS procedures.

Whereas the weights (the regression coefficients β) associated with the aforementioned regression model reflect the influence of ATB and SN on BI for the sampled population as a whole, they tell us little about how individuals might vary in their weighting of these two components. To modify the regression model to include individual weights, we propose the following linear regression model:

$$BI_i = \beta_{0i} + \beta_{1i} ATB_i + \beta_{2i} SN_i + \epsilon_i \quad (2)$$

In this model, the individual-varying intercept term β_{0i} accommodates scale differences in BI between individuals, whereas β_{1i} represents the influence of ATB on BI for individual i and β_{2i} represents the influence of SN on BI for individual i . In this model, one might view these individual weights (the regression coefficients) as randomly varying over a

population of persons. Knowing the mean of each weight tells us how people weigh attitudes and norms "on average," whereas knowing how they vary tells us about the extent to which individuals differ in weighting attitudes and norms to form intentions. In terms of estimating these individual weights, if individuals are measured on only one occasion, the individual weights cannot be independently determined. However, it is often the case that individuals are repeatedly measured over a series of time points, and in this case, as described later, estimation of the individual weights is possible. The model for repeated assessments of BI, ATB, and SN can be written as

$$BI_{ij} = \beta_{0i} + \beta_{1i}ATB_{ij} + \beta_{2i}SN_{ij} + \epsilon_{ij} \quad (3)$$

where $i = 1 \dots N$ individuals and $j = 1 \dots n_i$ observations for individual i . As these data are repeated measures across time, we might also posit a time effect on BI and allow this term to vary by individuals:

$$BI_{ij} = \beta_{0i} + \beta_{1i}Time_{ij} + \beta_{2i}ATB_{ij} + \beta_{3i}SN_{ij} + \epsilon_{ij} \quad (4)$$

For this model then, the individual effects of ATB and SN are obtained accounting for an individual's intercept and trend across time in terms of BI.

The model can be written in a slightly more general form using matrix notation. For this, consider the following regression model for an individual i ($i = 1, 2, \dots, N$), in which the $n_i \times 1$ response vector y for individual i is modeled in terms of p covariates (including the intercept):

$$y_i = X_i\beta_i + \epsilon_i \quad (5)$$

where y_i is the $n_i \times 1$ vector of responses for individual i , X_i is a known $n_i \times p$ covariate matrix, β_i is the unknown individual subject effects distributed $N(\mu_\beta, \Sigma_\beta)$, and ϵ_i is a $n_i \times 1$ vector of residuals distributed independently as $N(0, \sigma^2 I_{n_i})$. As in the usual linear regression model, the residuals are assumed to be normally distributed in the population with mean 0 and variance σ^2 . However, in the random-effects regression model the regression coefficients are also assumed to have a distribution in the population. In the present case, each coefficient β_{hi} ($h = 0, 1, 2, 3$) is assumed to be normally distributed in the population with a mean of μ_{β_h} and a variance $\sigma_{\beta_h}^2$. Conditional on the other model terms, the residuals are assumed to be independent and uncorrelated. However, the regression coefficients β_i , although independent of the residuals, are allowed to be correlated among themselves; that is, there is no assumption regarding the independence of the coefficient β_{1i} for variable x_1 with the coefficient β_{2i} for variable x_2 . In fact, these correlations can be of great interest and need to be specified and then estimated. As a result, we assume that the vector, in transposed form, of random regression coefficients (with $p = 4$) $\beta'_i = [\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_{3i}]$ are normally distributed with mean vector $\mu_\beta = [\mu_{\beta_0}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\beta_3}]$ and variance covariance matrix Σ_β given by:

$$\Sigma_\beta = \begin{bmatrix} \sigma_{\beta_0}^2 & \sigma_{\beta_0\beta_1} & \sigma_{\beta_0\beta_2} & \sigma_{\beta_0\beta_3} \\ \sigma_{\beta_1\beta_0} & \sigma_{\beta_1}^2 & \sigma_{\beta_1\beta_2} & \sigma_{\beta_1\beta_3} \\ \sigma_{\beta_2\beta_0} & \sigma_{\beta_2\beta_1} & \sigma_{\beta_2}^2 & \sigma_{\beta_2\beta_3} \\ \sigma_{\beta_3\beta_0} & \sigma_{\beta_3\beta_1} & \sigma_{\beta_3\beta_2} & \sigma_{\beta_3}^2 \end{bmatrix}$$

The variance terms represent the degree of variability in the population for these random effects, and the covariance terms reflect the degree of association among these effects. Notice that the usual multiple regression model is essentially a special case of the random-effects regression model with all elements of Σ_β equaling 0.

Because the individual subscript i is present for the y vector and the X matrix, each individual can be observed on a different number of occasions. Furthermore, the actual occasions measured can vary from individual to individual; this is because the time effect, which is in-

cluded as a variable in the X matrix, is treated as a continuous independent variable in the regression model. Thus, the values of the time variable do not have to be assumed to be the same across individuals as is usually done in a repeated measures analysis of variance model, for example. The number of columns in X is p with the intercept as the first parameter of the vector β_i ; The first column of X consists only of ones, whereas the remaining $p - 1$ columns are for the time and covariate values. Under these assumptions, the y_i are distributed as independent normals with mean $X_i\mu_\beta$ and variance-covariance matrix $X_i\Sigma_\beta X_i + \sigma^2 I_{n_i}$.

Estimates of each individual's intercept (β_{0i}), linear trend across time (β_{1i}), and TRA-related effects (β_{2i} and β_{3i}) can be accomplished using empirical Bayes methods. To do this, we make use of the assumption that these individual effects (β_i) have a distribution in the population (a "prior" distribution, namely $\beta_i \sim N[\mu_\beta, \Sigma_\beta]$), and then use the available data to estimate the parameters of this distribution (an "empirical prior" distribution characterized by μ_β and Σ_β) as well as the residual variance (σ^2). Once the residual variance and empirical prior distribution is characterized, estimates of the individual weights can be obtained as a function of the person's data and the empirical prior distribution. As shown in the Appendix, the empirical Bayes estimates are the mean,

$$\bar{\beta}_i = [X_i'X_i + \sigma^2\Sigma_\beta^{-1}]^{-1}X_i'(y_i - X_i\mu_\beta) + \mu_\beta, \quad (6)$$

and covariance matrix,

$$\Sigma_{\beta_i|y_i} = \sigma^2(X_i'X_i + \sigma^2\Sigma_\beta^{-1})^{-1}, \quad (7)$$

of the posterior distribution of the individual effects β_i given the data y_i , where the posterior distribution describes the probability of different values of β given the data y_i . The empirical Bayes estimates are then the expected individual weights for a given individual, and the corresponding estimates of the posterior covariance matrix represent the degree of uncertainty (and covariation) associated with these weights.

For applying Equations 6 and 7, some method of estimating the residual variance σ^2 and the parameters of the prior distribution (μ_β and Σ_β) is necessary. These parameters can be estimated using maximum marginal likelihood (MML) techniques, which are described in the Appendix. Because the empirical Bayes (and associated covariance matrix) estimates depend on the solution of the MML estimates, and vice versa, the estimation proceeds in an iterative manner until convergence. At convergence, the MML procedure provides standard errors for μ_β , Σ_β , and σ^2 that can be used to construct confidence intervals and tests of hypotheses for these parameters (Wald, 1943). Specifically, for a specific parameter, the parameter estimate divided by its standard error is compared to a standard normal frequency table to test the null hypothesis that a given parameter equals 0. While this use of the standard errors to perform hypothesis tests (and construct confidence intervals) for the fixed effects μ_β is generally reasonable, for the variance and covariance components (Σ_β and σ^2) this practice is problematic (see Bryk & Raudenbush, 1992, p. 55).

Instead, to test hypotheses related to the variance and covariance components, as well as the fixed effects, the likelihood-ratio (or difference in log likelihood) chi-square test can be used in certain cases. This test provides a way to test for statistical difference between alternative models when Model A, for example, includes all the parameters of, say Model B, plus some additional terms. The likelihood-ratio test compares the relative fit of the data provided by Models A and B and thus determines the significance of including these additional terms into the statistical model of the data. Evaluating the log-likelihood $\log L$ (given in the Appendix) using the estimated parameters of the two models yields $\log L_A$ and $\log L_B$. The significance of the additional terms in Model A is determined by comparing $-2(\log L_A - \log L_B)$ with a table of the chi-square distribution with degrees of freedom equal to the num-

ber of additional parameters in Model A. If this likelihood-ratio statistic exceeds the critical value of the chi-square distribution, the additional terms significantly improve the model fit of the data. Thus by comparing the log-likelihoods of relative nested models, this test can be used to test hypotheses both for the fixed effects (e.g., specific terms of μ_{β} equal 0) and the variance and covariance terms (e.g., specific terms of Σ_{β} equal 0).

Example

Data Set

The Television School and Family Smoking Prevention and Cessation Project (TVSFP; Flay et al., 1989) was designed to test independent and combined effects of a school-based social resistance curriculum and a television-based program in terms of tobacco use prevention and cessation. The initial study sample consisted of seventh-grade students who were pretested in January 1986 (T1). Students who took the pretest completed an immediate postintervention questionnaire in April 1986 (T2), a 1-year follow-up questionnaire (April 1987; T3), and a second-year follow-up (April 1988; T4). T1 data were collected from 6,695 seventh-grade students in 169 classrooms in Los Angeles County (representing 35 public schools in four school districts) and 67 classrooms in San Diego County (representing 12 public schools in two school districts). Randomization to various design conditions was at the school level, whereas much of the intervention was delivered to students within classrooms. As reported by Flay et al. (1995), there were no significant intervention effects in terms of the behavioral intentions that we examine in this article. Thus, for simplicity, we do not include any intervention effects in the analysis presented later. Of the original participants, 5,475 (81.8%), 4,854 (72.3%), and 3,719 (55.5%) were recontacted at T2, T3, and T4, respectively. Greater attrition was observed in Los Angeles than in San Diego, and African American students and those with lower grades were more likely to drop out (Flay et al., 1995). For this illustration of the random-effects model, a subset of the TVSFP data was used. We concentrated on students who were measured at all four study time points (baseline, postintervention follow-up, and 1- and 2-year follow-ups) and who within a time point had complete data on all measures used in the analyses; in all, there were 1,002 students who were in seventh grade at the first time point satisfying these criteria.

It is important to note that the sample used in the present illustration represents only a fraction of the complete set of students from the study. As such, the analyses presented later may suffer from selection biases to the degree that the sample of students used in the analyses is not representative of the original larger sample of students. This subset of students with complete data across the four study timepoints was chosen primarily to simplify the presentation of the already complex issues examined in this article. For more information regarding the treatment of missing data in longitudinal studies, see Little and Rubin (1987) or Little (1995).

Measures

The behavioral intention examined in this study was the student's intention to smoke cigarettes, which was measured with two items. Participants were asked if they thought they would

ever smoke in the future and if they thought they might ever ask another person to let them try a cigarette (where, for both items, 1 = *definitely would* and 6 = *definitely would not*). Each student's BI score was defined as the sum of these two items. These two items were highly correlated at the four study time points (0.59, 0.65, 0.71, and 0.73, respectively).

Attitudes toward smoking were measured by combining responses from four pairs of Likert-style items about the effects of smoking on lung cancer and heart disease. On the first pair, seventh-grade students indicated how worried they were (where 1 = *extremely worried* and 5 = *not at all worried*) and how likely it is (where 1 = *more than 80% likely* and 5 = *less than 20% likely*) that smokers in general will contract lung or heart disease. On the second pair, students indicated how worried they were and how likely it is that smokers in general might die from lung or heart disease. Whereas the first two pairs focused on adolescents' perceptions of the consequences of smoking among people in general, the second two pairs focused on perceptions of the personal consequences of smoking. On the third pair, students indicated how worried they were (where 1 = *extremely worried* and 5 = *not at all worried*), and how likely it is (where 1 = *more than an 80% chance* and 5 = *less than a 20% chance*) that they personally would contract lung cancer and heart disease if they smoked. Finally, on the fourth pair, students indicated how worried they were and how likely it is that they personally would die from these diseases if they smoked. Adopting a value-expectancy approach to attitudes, responses to the first item in each pair were multiplied by responses to the second item in each pair. Reliability for the resulting four variables was quite high at the four study time points (Cronbach's $\alpha = .84, .88, .89, \text{ and } .88$, respectively). In our analyses, social normative beliefs were represented by a single variable that was the product of two items: how many of an adolescent's 10 closest friends smoke (where 1 = *none* and 6 = *more than seven*), and how much close friends influence what that adolescent does during the week (where 1 = *no influence at all* and 4 = *a great deal of influence*).

Random-Effects Regression Model

The following random-effects regression model was fit to these data:

$$BI_{ij} = \beta_{0i} + \beta_{1i} \text{Month}_j + \beta_{2i} \text{ATB}_{ij} + \beta_{3i} \text{SN}_{ij} + \epsilon_{ij} \quad (8)$$

with $i = 1 \dots 1002$ students and $j = 1 \dots 4$ observations per student, and where BI_{ij} = value of BI for student i at time point j , β_{0i} = baseline BI level for student i , β_{1i} = monthly change in BI for student i , Month_j = value of month (either 0, 2, 14, or 26) at time point j , β_{2i} = change in BI associated with unit change in ATB for student i , ATB_{ij} = value of ATB for student i at time point j , β_{3i} = change in BI associated with unit change in SN for student i , SN_{ij} = value of SN for student i at time point j , and ϵ_{ij} = model residual for student i at time point j .

In the present model, individual effects are being assumed for the intercept (β_{0i}), the linear trend over time (β_{1i}), the ATB effect (β_{2i}), and the SN effect (β_{3i}); that is, an individual will traverse their own trend line across time in terms of BI and additionally will have their own weight for both ATB and SN effects on BI. Because TRA posits that there is significant vari-

ability in terms of the individual weights for ATB and SN on BI, a test of this tenet of TRA amounts to the statistical test of whether the null hypothesis of $\sigma_{\beta_2}^2 = \sigma_{\beta_3}^2 = 0$ (and thus all associated covariate terms also equal zero) can be rejected.

For the ATB and SN measures, *z* scores of these measures were computed (for each of the two measures, this was done across individuals at each of the four time points) to more readily compare the size of the effects of these two components. Using *z* scores also has the advantage that the regression coefficients β_{3i} and β_{4i} reflect change in standard deviation units for ATB and SN. Similarly, to compare the magnitude of the effects that was due to the intercept and linear trend across months, we used orthogonal polynomial transformation for these time-related terms taking into account the nonequally spaced time points (e.g., Months 0, 2, 14, and 26). These orthogonal polynomial transforms were calculated according to the procedure outlined in Bock (1975) for nonequally spaced time points. This model for student *i* can be represented in matrix form as follows:

$$\begin{bmatrix} BI_{i1} \\ BI_{i2} \\ BI_{i3} \\ BI_{i4} \end{bmatrix} = y_i \begin{bmatrix} 0.5 & -0.5034 & z(ATB_1)_i & z(SN_1)_i \\ 0.5 & -0.4075 & z(ATB_2)_i & z(SN_2)_i \\ 0.5 & 0.1678 & z(ATB_3)_i & z(SN_3)_i \\ 0.5 & 0.7432 & z(ATB_4)_i & z(SN_4)_i \end{bmatrix} \begin{bmatrix} \beta_{0i} \\ \beta_{1i} \\ \beta_{2i} \\ \beta_{3i} \end{bmatrix} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \end{bmatrix}$$

X_i β_i ϵ_i

Because β_0 represents the mean of *y* when all *X* variables equal 0, it is important to note that, all other things being equal, using the orthogonal polynomial transformation changes the meaning of β_0 in the model from that representing the mean of *y* at the first time point (i.e., when month = 0) to that representing the mean of *y* when the orthogonal linear transform equals zero (this can be shown to be when month = 10.5). For purposes of interpretation it is often the former representation of β_0 that is of more interest; that is, one is interested in knowing the mean level of students' intention to smoke at the start of the study. As indicated by Bock (1975), one can reexpress the estimated orthogonal polynomial terms back into terms associated with the raw metric of time (i.e., corresponding to month values of 0, 2, 14, and 26), and so in interpreting the results, we shift between the estimates based on the transformed and non-transformed representations of the study time points. For more information regarding how the choice of scale of the independent variables (*X*) influences the interpretation of the model parameters see Bryk and Raudenbush (1992) or Longford (1993).

Results

Means and standard deviations for the BI across the four time points and the correlations of BI with ATB and SN over time

are given in Table 1. These descriptive statistics are given to provide an overview of the data, and to aid in understanding the statistical results of the analyses presented later. As can be seen from Table 1, the intentions to smoke increase slightly over time and the variability in this measure increases over time. Also, the correlations of ATB with BI are generally constant over time, whereas the correlations of SN with BI are generally increasing over time.

To test the TRA assumption (the null hypothesis of whether $\sigma_{\beta_2}^2 = 0$ and $\sigma_{\beta_3}^2 = 0$ can be rejected or not), we first fit a model setting these two variance components and all related covariance terms equal to zero. The results from this analysis are given in Table 2. It is interesting to note that the overall effects of ATB and SN are relatively similar (.49 and .58). Also, although there is a significant linear increase in the behavioral intentions to smoke across time, the magnitude of this increase is small in relation to the magnitude of the constant effect over time. This is also true for the variation in the individual constant and linear effects; there is considerably more individual variation in the constant term (variance = 7.98) than in the linear term (variance = 1.64). In terms of the raw metric of time described earlier (i.e., using values of 1 for the intercept and month values of 0, 2, 14, and 26), reexpressing the estimates of the orthogonally transformed time-related terms yields an intercept of 3.768 (e.g., estimated BI value at Month 0) and a linear increase of 0.012 units per month. The covariance of the constant and linear trend effects expressed as a correlation equals .25, which suggests that there is a moderate positive association between an individual's average BI scores, averaging over time, and their trend in BI across time. Reexpressing this correlation in the raw metric, yields a correlation of -.21, which suggests a moderate negative association between an individual's starting point (Month 0) and the linear trend across time. This negative association between the intercept and the linear trend across time may be due to a ceiling effect, in that individuals with high initial scores cannot increase their scores over time to the same extent as individuals with lower initial scores.

Next, we allowed the effects of ATB and SN to vary with individuals. The results of this analysis are presented in Table 3. To test whether allowing the effects of ATB and SN to vary with individuals significantly improves the fit of the model, we used the likelihood-ratio test. For this, a chi-square value of 194.8 is observed on 7 degrees of freedom, which is significant at the *p*

Table 1
Descriptive Statistics

Measure	Month			
	0	2	14	26
Statistic for BI across time				
<i>M</i>	3.61	3.91	3.98	4.12
<i>SD</i>	1.96	2.22	2.35	2.67
Correlations with ATB and SN				
ATB	.29	.26	.32	.31
SN	.28	.34	.40	.45

Note. *n* = 1,002. BI = behavioral intention; ATB = attitude toward behavior; SN = subjective norm.

Table 2
*Behavioral Intentions to Smoke—Modeling Trend Over Time and Effects of ATB and SN:
 Effects of ATB and SN Do Not Vary With Individuals*

Term	Parameter	Estimate	SE	z	p
Constant	μ_{β_0}	7.784	0.100	78.01	<.0001
Linear	μ_{β_1}	0.246	0.060	4.09	<.0001
ATB	μ_{β_2}	0.490	0.036	13.60	<.0001
SN	μ_{β_3}	0.577	0.033	17.60	<.0001
Constant variance	$\sigma_{\beta_0}^2$	7.986	0.449		
Constant and linear covariance	$\sigma_{\beta_0\beta_1}$	0.916	0.192		
Linear variance	$\sigma_{\beta_1}^2$	1.637	0.173		
Residual variance	σ^2	1.975	0.062		

Note. $\text{Log } L = -8,151.91$. ATB = attitude toward behavior; SN = subjective norm.

< .0001 level. Thus, there is considerable evidence that the ATB and SN weights vary with individuals, and so there is strong empirical support of this tenet of TRA. Interestingly, although as a group the ATB and SN effects are similar (both approximately = .5), the degree of individual variation in these effects is considerably greater for SN (variance = 0.318) than for ATB (variance = 0.152). Also, as indicated by the covariance term $\sigma_{\beta_2\beta_3}$, there is evidence of a positive association between the degree to which an individual weighs ATB and SN (expressed as a correlation this association equals .51). Expressing the full variance-covariance matrix of the individual random effects as a correlation matrix yields:

$$\begin{bmatrix} & \text{Constant} & \text{Linear} & \text{ATB} & \text{SN} \\ \text{Constant} & — & & & \\ \text{Linear} & .198 & — & & \\ \text{ATB} & .401 & .186 & — & \\ \text{SN} & .396 & .534 & .510 & — \end{bmatrix}$$

As this matrix suggests, there is greater association between an individual's weighting of SN and the linear trend in BI across time, than for the weighting of ATB with the BI linear trend. Thus, the greater the weight an individual gives to SN, and to a lesser extent ATB, the more positively increasing is their trend in BI across time. All terms are positively associated with the overall constant term which simply reflects that more positive linear trend and greater weighting of SN and ATB is associated with overall higher BI values across all time points.

To examine the degree to which the effects of ATB and SN vary across time, we included interactions of these terms with the linear time effect into the model. These additional terms were considered to be fixed, and not random, effects; that is, they were treated in the same way as in an ordinary regression analysis. Although one could conceptualize these as random effects (e.g., changes in the ATB effect across time varying by individuals), with data only at four time points and four random effects in the model already, increasing the number of random effects would result in an overparameterized model. The addition of these two interaction terms significantly improved the model fit ($\chi^2 = 19.5$, $p < .0001$); however, only the SN \times Linear term was observed to be statistically significant ($z = 4.22$, $p < .0001$). This suggests that, overall, the influence of SN on BI increases over time, whereas for ATB, there is no evidence

of any significant change over time in its effect. Finally, we included terms into the model to account for the sex of the individual and to examine any potential interactions of sex with ATB and SN. The results of this model are given in Table 4. Adding the main effect of sex and the interactions with ATB and SN significantly improves the model fit ($\chi^2 = 17.8$, $p < .001$). Overall, both the influences of SN and ATB on BI are significantly greater for female students than male students, although this difference between gender groups is more pronounced for SN than ATB. Based on the model estimates given in Table 4, Figure 2 displays the estimated regressions of BI on SN and ATB for male and female students at T1 and T4. Figure 2 clearly illustrates that the effect of SN on BI increases over time, whereas the effect of ATB on BI is consistent across time. Also illustrated in Figure 2 is the increased influence for female students of ATB and especially SN on BI, relative to male students. Note that in the estimation of the regression of BI on SN, ATB was held constant at its mean of 0 and vice versa for the regression of BI on ATB.

Discussion

For our example focusing on the intentions to smoke cigarettes in a longitudinal sample of seventh-grade students, the analyses clearly support the notion that students weigh their ATBs and their SNs differently when assessing their BI to smoke. Additionally, although overall the weights were similar for these two constructs, the amount of individual variation in these weights was quite different. There was considerably more individual variation in the influence of SN than in the influence of ATB. Part of this increased SN individual variation was explained by the observation that the influence of SN increased over time and was more pronounced for female students than for male students. The weight an individual assigned to one construct (ATB or SN) was positively associated with the weight assigned to the other construct, indicating that the influences of these constructs on BIs were not independent of each other.

When individual variation exists in the regression coefficients of the model, as was evidenced for the influences of ATB and SN, one may be interested in modeling this variation in terms of characteristics of the individuals; that is, individual-level covariates. For instance, in the example given earlier, some of the

Table 3
*Behavioral Intentions to Smoke—Modeling Trend Over Time and Effects of ATB and SN:
 Effects of ATB and SN Do Vary With Individuals*

Term	Parameter	Estimate	SE	z	p
Constant	μ_{β_0}	7.700	0.099	78.18	<.0001
Linear	μ_{β_1}	0.197	0.057	3.47	<.0005
ATB	μ_{β_2}	0.467	0.040	11.78	<.0001
SN	μ_{β_3}	0.550	0.041	13.27	<.0001
Constant variance	$\sigma_{\beta_0}^2$	7.196	0.451		
Constant and linear covariance	$\sigma_{\beta_0\beta_1}$	0.605	0.179		
Linear variance	$\sigma_{\beta_1}^2$	1.301	0.157		
Constant and ATB covariance	$\sigma_{\beta_0\beta_2}$	0.419	0.115		
Linear and ATB covariance	$\sigma_{\beta_1\beta_2}$	0.083	0.066		
ATB variance	$\sigma_{\beta_2}^2$	0.152	0.058		
Constant and SN covariance	$\sigma_{\beta_0\beta_3}$	0.599	0.125		
Linear and SN covariance	$\sigma_{\beta_1\beta_3}$	0.343	0.071		
ATB and SN covariance	$\sigma_{\beta_2\beta_3}$	0.111	0.046		
SN variance	$\sigma_{\beta_3}^2$	0.318	0.057		
Residual variance	σ^2	1.765	0.062		

Note. $\log L = -8,054.50$. ATB = attitude toward behavior; SN = subjective norm.

individual variation in the SN effect on BI was explained in terms of the SN by sex interaction; the individual influence of SN on BI was determined to some degree by the gender of the individual. For this reason, the random-effects model is sometimes described as a "slopes as outcomes" model (Burstein,

Linn, & Capell, 1978), a hierarchical linear model (Bryk & Raudenbush, 1992), or a multilevel model (Goldstein, 1987). These representations show that just as time-varying covariates (e.g., ATB and SN) are included in the model to explain variation in the time-varying outcomes (e.g., BI), individual-level

Table 4
*Behavioral Intentions to Smoke—Modeling Trend Over Time and Effects of ATB and SN:
 Effects of ATB and SN Do Vary With Individuals—Interactions With Sex and Time*

Term	Parameter	Estimate	SE	z	p
Constant	μ_0	7.671	0.130	59.13	<.0001
Linear	μ_1	0.219	0.057	3.87	<.0001
ATB	μ_2	0.537	0.055	9.69	<.0001
SN	μ_3	0.674	0.054	12.44	<.0001
Sex (0 = female, 1 = male)	β_4	0.039	0.099	0.39	<.69
ATB \times Sex	β_5	-0.159	0.079	-2.01	<.045
SN \times Sex	β_6	-0.277	0.082	-3.39	<.0007
ATB \times Linear	β_7	0.040	0.060	0.67	<.50
SN \times Linear	β_8	0.269	0.065	4.16	<.0001
Constant variance	$\sigma_{\beta_0}^2$	7.106	0.445		
Constant and linear covariance	$\sigma_{\beta_0\beta_1}$	0.457	0.175		
Linear variance	$\sigma_{\beta_1}^2$	1.236	0.154		
Constant and ATB covariance	$\sigma_{\beta_0\beta_2}$	0.439	0.113		
Linear and ATB covariance	$\sigma_{\beta_1\beta_2}$	0.050	0.065		
ATB variance	$\sigma_{\beta_2}^2$	0.148	0.057		
Constant and SN covariance	$\sigma_{\beta_0\beta_3}$	0.597	0.123		
Linear and SN covariance	$\sigma_{\beta_1\beta_3}$	0.322	0.069		
ATB and SN covariance	$\sigma_{\beta_2\beta_3}$	0.100	0.045		
SN variance	$\sigma_{\beta_3}^2$	0.289	0.055		
Residual variance	σ^2	1.772	0.062		

Note. $\log L = -8,035.81$. ATB = attitude toward behavior; SN = subjective norm.

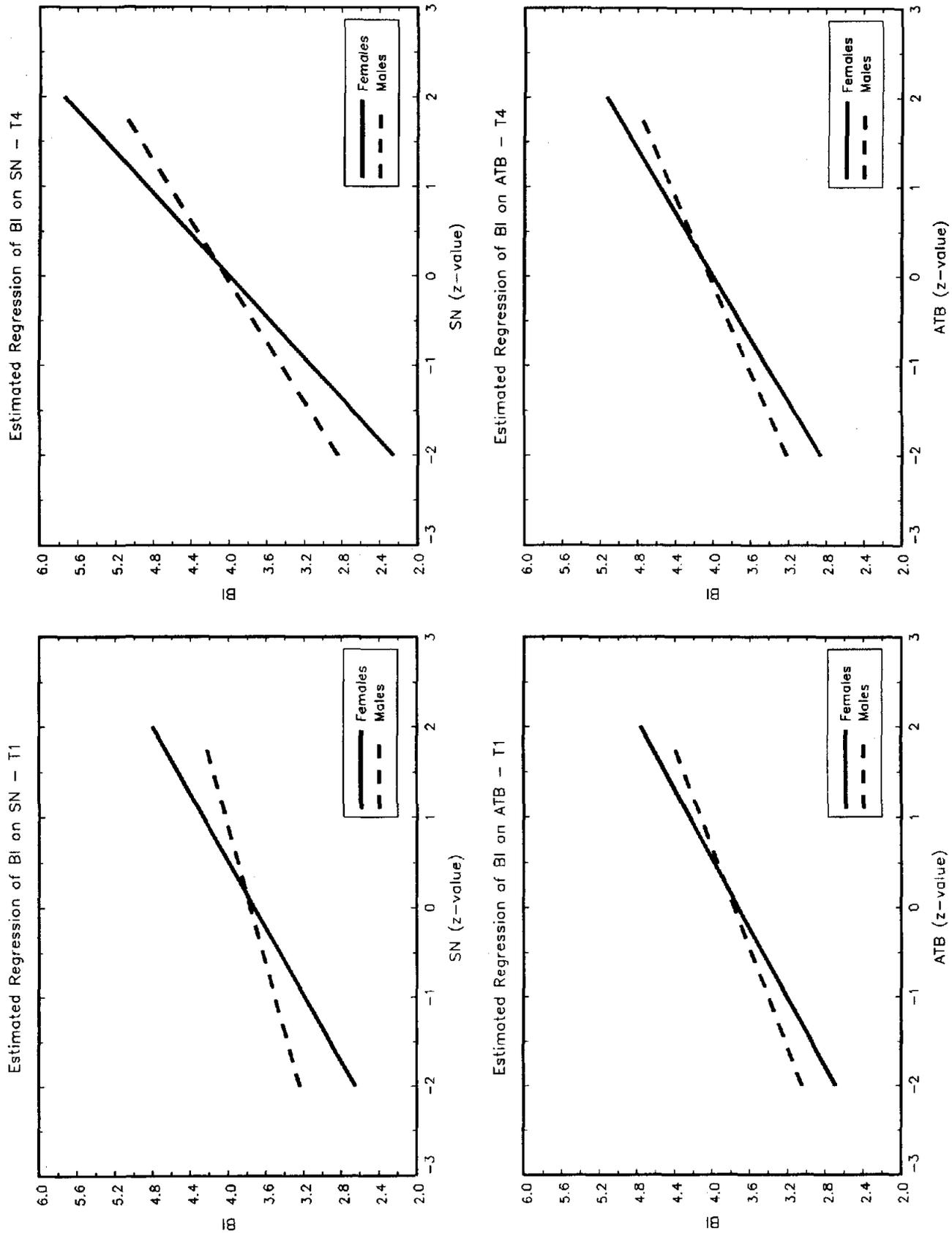


Figure 2. Estimated regressions of behavioral intentions (BI) on subjective norms (SN) and attitudes toward the behavior (ATB). T1 and T4 = Time points 1 and 4, respectively.

covariates (e.g., sex) are included to explain variation in the individual-level regression coefficients of the model (e.g., the individual effects of ATB and SN). Thus, as there is less individual variation in the regression coefficient of a time-varying variable, say X , there is less potential for a significant effect of an individual-level covariate, say Z , on the coefficient for variable X (and thus an $X \times Z$ interaction effect on the outcome variable). Stated in a more obvious way, if the influence of a variable does not vary by individuals, it cannot vary by groups composed of these individuals.

It is important to note that the proposed method of estimating individual weights depends on individuals being repeatedly assessed for their BIs, ATBs, and SNs. Not all individuals need to have the same number of repeated measurements; however, the amount of uncertainty about the individual's weights is a function of the amount of data available for that individual. If there is more data available (the individual being measured at more time points), there will generally be less uncertainty about the individual's weights.

In this example, we have concentrated on the repeated observations that were nested within individuals. In the terminology of multilevel analysis (Goldstein, 1987) and hierarchical linear models (Bryk & Raudenbush, 1992) this is termed a two-level data structure with the individuals representing Level 2 and the nested repeated observations representing Level 1. Thus, the model that we have presented is sometimes referred to as a two-level model. Individuals themselves, however, were observed nested within both classrooms and schools, and so a higher level model could have been used to further analyze these data. A higher level model would reveal the degree of variation that is attributable to the nesting of individuals within classrooms, and classrooms within schools. We chose the two-level model primarily for simplicity; further discussion regarding multiple levels of nesting can be found in Longford's (1993) book.

Although we have focused on a specific example, other applications of the methods presented here are certainly possible in clinical research. One potential application is in examining the association between drug plasma levels and clinical response to mental illness. In these studies, repeated determinations of drug plasma levels and severity of illness are typically made in a sample of psychiatric patients. For example, in the study of Riesby et al., (1977), a sample of depressed inpatients were measured weekly for 4 weeks in terms of their imipramine and desipramine plasma levels, as well as their clinical status, as measured by the Hamilton Rating Scale for Depression (HAM-D). A model of the HAM-D scores over time could then consider the effects of the two drug plasma levels to vary by individual and, thus, treat these as random effects in the model. As another example, in relapse research one is often interested in examining the effects of time-varying influences (e.g., stressful life events or degree of coping) on repeated assessments of whether the person has engaged in the behavior of interest (e.g., smoking, drinking, or taking drugs). Here, one might consider, for example, the possibility that the influence of stressful life events on the behavior varies by individual, and so treat the stress effect as random in the model. The key feature of these potential applications is repeated assessments of both the dependent and independent variables, it is this within-subjects measurement of both which allows one to potentially consider the effect (or

effects) of the independent variable (or variables) to vary by individuals.

In the present example, effects that are due to an individual's overall level and linear trend in BI over time were treated as random, in addition to the influence of time-varying covariates (ATB and SN). Often in longitudinal research, the intercept (or the constant or overall level) and possibly trend effects are treated as random, leaving the effects of both time-invariant and time-varying covariates as fixed. Although, with the exception of the intercept, the effects of time-invariant (or individual-level) covariates cannot vary by individuals (i.e., be treated as random), treating the effects of time-varying covariates as random effects is possible in longitudinal studies. As noted by Longford (1993) the issue of what to consider as random and fixed is a debatable one that depends on both statistical and substantive considerations. In the present example, interest in testing a specific hypothesis of TRA led us to examine the possibility that the effects of ATB and SN on BI varied by individuals. Other times, the data may not support the inclusion of multiple random effects, especially as the number of time points is small. In this case, the iterative estimation procedure may fail because of problems associated with the variance covariance matrix of the random effects (e.g., a variance term becoming zero or non-positive or a correlation between two random effects approaching or exceeding unity). The issue of the number of random effects that are estimable clearly depends on the number of time points measured within individuals as well as the number of individuals; however, general guidelines for determining this are difficult to provide. More work on this issue is clearly needed.

In terms of software to perform the estimation of this model, there are several available programs to perform random-effects regression analysis (e.g., HLM [Bryk, Raudenbush, Seltzer, & Congdon, 1989], ML3 [Prosser, Rasbash, & Goldstein, 1991], VARCL [Longford, 1986], the BMDP 5V procedure; the SAS procedure MIXED), although not all programs can provide all of the results presented in this article. A detailed comparison of some of these programs is included in Kreft, de Leeuw, and van der Leeden (1994). For the results presented in this article, the program MIXREG (Hedeker, 1993b) was used.¹

The model described in this article and the software listed earlier is appropriate when the dependent variable is measured on a continuous scale. Sometimes, however, in psychological research the dependent variable is measured on a dichotomous (e.g., yes or no) or ordinal (e.g., symptom severity of none, mild, or definite) scale. Although not as developed, an increasing amount of work has focused on random-effects models for dichotomous (Anderson & Aitkin, 1985; Gibbons & Bock, 1987; Stiratelli, Laird, & Ware, 1984; Wong & Mason, 1985) and ordinal (Ezzet & Whitehead, 1991; Hedeker & Gibbons, 1994; Jansen, 1990) responses. Also, software programs are available for random-effects models of dichotomous (EGRET; Statistics and Epidemiology Research Corporation, 1991) and ordinal (MIXOR; Hedeker, 1993a) response data, although

¹ This DOS-based program, as well as the MIXOR program, can be obtained from Ann Hohmann, National Institute of Mental Health Services Research Branch, 5800 Fishers Lane, Room 10C-06, Rockville, Maryland 20857, or from Donald Hedeker via Internet at hedeker@uic.edu.

these programs are not as sophisticated as most of the software for continuous response data listed earlier.

Conclusion

Adding to the research evidence for TRA's central predictions, we have provided evidence for an ancillary prediction of TRA. That prediction concerns the relative importance of a person's ATB, SN, and BI. As has been shown, a random-effects regression model can be used to test this prediction and can additionally estimate the weights of these two indicators of BI. The random-effects model then provides behavioral scientists with a useful way of examining relations among variables at the level of the individual, and by doing so, allows a more thorough assessment of the determinants of individual differences.

References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Anderson, D. A., & Aitkin, M. (1985). Variance component models with binary response: Interviewer variability. *Journal of the Royal Statistical Society*, *47*, 203-210.
- Bock, R. D. (1966). Contributions of multivariate experimental designs to educational research. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 820-840). Chicago: Rand-McNally.
- Bock, R. D. (1975). *Multivariate statistical methods in behavioral research*. McGraw-Hill: New York.
- Bock, R. D. (1983a). Within-subject experimentation in psychiatric research. In R. D. Gibbons & M. W. Dysken (Eds.), *Statistical and methodological advances in psychiatric research* (pp. 59-90). New York: Spectrum.
- Bock, R. D. (1983b). The discrete Bayesian. In H. Wainer & S. Messick (Eds.), *Modern advances in psychometric research* (pp. 103-115). Hillsdale, NJ: Erlbaum.
- Bock, R. D. (Ed.). (1989a). Measurement of human variation: A two stage model. *Multilevel analysis of educational data* (pp. 319-342). San Diego, CA: Academic Press.
- Bock, R. D. (Ed.). (1989b). *Multilevel analysis of educational data*. San Diego, CA: Academic Press.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147-158.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Bryk, A. S., Raudenbush, S. W., Seltzer, M., & Congdon, R. T. (1989). *An introduction to HLM: Computer program and users' guide*. Chicago: Scientific Software.
- Burstein, L., Linn, R. L., & Capell, I. (1978). Analyzing multi-level data in the presence of heterogeneous within-class regressions. *Journal of Educational Statistics*, *4*, 347-389.
- DeLeeuw, J., & Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, *11*, 57-85.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, *39*, 1-38.
- Dempster, A. P., Rubin, D. B., & Tsutakawa, R. K. (1981). Estimation in covariance components models. *Journal of the American Statistical Association*, *76*, 341-353.
- Ezzet, F., & Whitehead, J. (1991). A random effects model for ordinal responses from a crossover trial. *Statistics in Medicine*, *10*, 901-907.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior*. Reading, MA: Addison-Wesley.
- Flay, B. R., Brannon, B. R., Johnson, C. A., Hansen, W. B., Ulene, A. L., Whitney-Saltiel, D. A., Gleason, L. R., Sussman, S., Gavin, M., Glowacz, K. M., Sobol, D. F., & Spiegel, D. C. (1989). The television, school and family smoking cessation and prevention project: I. Theoretical basis and program development. *Preventative Medicine*, *17*, 585-607.
- Flay, B. R., Miller, T. Q., Hedeker, D., Siddiqui, O., Britton, C. F., Brannon, B. R., Johnson, C. A., Hansen, W. B., Sussman, S., & Dent, C. (1995). The television, school and family smoking cessation and prevention project: VIII. Student outcomes and mediating variables. *Preventative Medicine*, *24*, 29-40.
- Gibbons, R. D., & Bock, R. D. (1987). Trend in correlated proportions. *Psychometrika*, *52*, 113-124.
- Gibbons, R. D., Hedeker, D., Waternaux, C., & Davis, J. M. (1988). Random regression models: A comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacology Bulletin*, *24*, 438-443.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. New York: Oxford University Press.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems (with discussion). *Journal of the American Statistical Association*, *72*, 320-385.
- Hedeker, D. (1993a). *MIXOR: A FORTRAN program for mixed-effects linear ordinal probit and logistic regression* (Tech. Rep.). Chicago: Prevention Research Center, School of Public Health, University of Illinois.
- Hedeker, D. (1993b). *MIXREG: A FORTRAN program for mixed-effects linear regression with autocorrelated errors* (Tech. Rep.). Chicago: Prevention Research Center, School of Public Health, University of Illinois.
- Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, *50*, 933-944.
- Hedeker, D., Gibbons, R. D., & Flay, B. R. (1994). Random regression models for clustered data: With an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*, *62*, 757-765.
- Hedeker, D., Gibbons, R. D., Waternaux, C., & Davis, J. M. (1989). Investigating drug plasma levels and clinical response using random regression models. *Psychopharmacology Bulletin*, *25*, 227-231.
- Hui, S. L., & Berger, J. O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association*, *78*, 753-759.
- Jansen, J. (1990). On the statistical analysis of ordinal data when extravariation is present. *Applied Statistics*, *39*, 75-84.
- Jennrich, R. I., & Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, *42*, 805-820.
- Kreft, I. G., de Leeuw, J., & van der Leeden, R. (1994). Comparing five different statistical packages for hierarchical linear regression: BMDP-5V, GENMOD, HLM, ML3, and VARCL. *American Statistician*, *48*, 324-335.
- Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics*, *40*, 961-971.
- Little, R. J. A. (1995). Modeling the drop-out mechanism in longitudinal studies. *Journal of the American Statistical Association*, *90*, 1112-1121.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Longford, N. T. (1986). VARCL—Interactive software for variance component analysis: Applications for survey data. *Professional Statistician*, *5*, 28-32.
- Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, *74*, 817-827.
- Longford, N. T. (1993). *Random coefficient models*. New York: Oxford University Press.

- Louis, T. A. (1991). Using empirical Bayes methods in biopharmaceutical research. *Statistics in Medicine*, 10, 811–829.
- Prosser, R., Rasbash, J., & Goldstein, H. (1991). *ML3 software for three-level analysis, users' guide for Version 2*. London: Institute of Education, University of London.
- Raudenbush, S. W. (1988). Educational applications of hierarchical linear models: a review. *Journal of Educational Statistics*, 13, 85–116.
- Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociological Education*, 59, 1–17.
- Raudenbush, S. W., & Willms, J. D. (Eds.). (1991). *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective*. San Diego, CA: Academic Press.
- Riesby, N., Gram, L. F., Bech, P., Nagy, A., Petersen, G. O., Ortmann, J., Ibsen, I., Dencker, S. J., Jacobsen, O., Krautwald, O., Sondergaard, I., & Christiansen, J. (1977). Imipramine: Clinical effects and pharmacokinetic variability. *Psychopharmacology*, 54, 263–272.
- Rubin, D. B. (1980). Using empirical Bayes techniques in the law school validity studies. *Journal of the American Statistical Association*, 75, 801–827.
- Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications for future research. *Journal of Consumer Research*, 15, 325–343.
- Silvey, S. D. (1975). *Statistical inference*. New York: Chapman and Hall.
- Statistics and Epidemiology Research Corporation. (1991). EGRET [Computer software]. Reference manual. Seattle, WA: Author.
- Stiratelli, R., Laird, N. M., & Ware, J. H. (1984). Random-effects models for serial observations with binary response. *Biometrics*, 40, 961–971.
- Strenio, J. F., Weisberg, H. I., & Bryk, A. S. (1983). Empirical Bayes estimation of individual growth curve parameters and their relationship to covariates. *Biometrics*, 39, 71–86.
- Vacek, P. M., Mickey, R. M., & Bell, D. Y. (1989). Application of a two-stage random effects model to longitudinal pulmonary function data from sarcoidosis patients. *Statistics in Medicine*, 8, 189–200.
- Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. *The American Statistician*, 39, 95–101.
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54, 426–482.
- Wong, G. Y., & Mason, W. M. (1985). The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association*, 80, 513–524.

Appendix

Estimation

A combination of two complementary methods have been proposed for the estimation of the random-effects regression model parameters (Bock, 1989a; Laird & Ware, 1982). For estimation of the random effects β_i , empirical Bayes methods have been recommended, whereas MML methods are recommended for estimation of variance parameters, σ^2 and Σ_β , and mean vector μ_β . A thorough treatment of parameter estimation is included both in Bryk and Raudenbush (1992) and Longford (1993). In what follows, we describe the general approach to parameter estimation for the model with multiple random effects. Hedeker, Gibbons, and Flay (1994) provide a less technical description of estimation for the special case of a random-intercepts regression model.

Empirical Bayes estimates are sometimes termed EAP ("Expected a posteriori") estimates, because they are derived as the mean of the posterior distribution of β_i , given y_i . From a Bayesian viewpoint, all of the information about β_i that is available from the data y_i is expressed in the "posterior" probability or probability density of β_i , given y_i , which is given in Bayes theorem as follows:

$$p(\beta_i | y_i) = \frac{f(y_i | \beta; \sigma^2)g(\beta | \mu_\beta, \Sigma_\beta)}{h(y_i | \sigma^2, \mu_\beta, \Sigma_\beta)}, \quad (\text{A1})$$

where $f(\cdot)$ is called the "likelihood function,"

$$f(y_i | \beta; \sigma^2) = (2\pi\sigma^2)^{-n_i/2} \exp\left[-\frac{1}{2} \frac{(y_i - X_i\beta)'(y_i - X_i\beta)}{\sigma^2}\right], \quad (\text{A2})$$

$g(\cdot)$ is the "prior" probability density,

$$g(\beta | \mu_\beta, \Sigma_\beta) = (2\pi)^{r/2} |\Sigma_\beta|^{-1/2} \exp\left[-\frac{1}{2} (\beta - \mu_\beta)' \Sigma_\beta^{-1} (\beta - \mu_\beta)\right], \quad (\text{A3})$$

and $h(y_i | \sigma^2, \mu_\beta, \Sigma_\beta)$ is the marginal probability of the observation,

$$h(y_i | \sigma^2, \mu_\beta, \Sigma_\beta) = \int_\beta f(y_i | \beta; \sigma^2)g(\beta | \mu_\beta, \Sigma_\beta)d\beta. \quad (\text{A4})$$

The posterior distribution then describes the probability of different values of β given the data y_i . The mean of this distribution is the expected value of β given the data, whereas the variance of this distribution represents the dispersion of the distribution about these random effects. Because both the likelihood and the prior are assumed to be normal, the posterior distribution is also normal, with mean given by Equation 6 and covariance matrix given by Equation 7. These are derived as the mean and variance-covariance matrix of the conditional distribution of β_i given y_i (see Bock, 1983a, 1983b), from which it can be shown that the mean is equal to

$$\bar{\beta}_i = R_i \hat{\beta}_i + (I_r - R_i) \mu_\beta, \quad (\text{A5})$$

and that the covariance matrix is

$$\Sigma_{\beta_i | y_i} = (I_r - R_i) \Sigma_\beta, \quad (\text{A6})$$

where $R_i = \Sigma_\beta [\Sigma_\beta + (X_i'(\sigma^2 I_{n_i})^{-1} X_i)^{-1}]^{-1}$ is the "multivariate analog of reliability" (Bock, 1966) and $\hat{\beta}_i$ is the least squares estimator for individual i . These forms make clear the fundamental property of empirical Bayes estimation; namely, that $\bar{\beta}_i$ is a function of both the actual data and the empirical prior distribution specified for β . As information about an individual decreases (the reliability decreases toward 0), the estimate for the individual approaches the posited mean of the empirical prior distribution of β_i , namely μ_β . Alternatively, as information about an individual increases (e.g., the reliability increases toward 1), the estimate of the individual approaches the least squares estimator $\hat{\beta}_i$ based only on that individual's data. The degree to which an individual's estimate is weighted by his or her data, then, is a function of the reliability of those data. Similarly, the form of the variance of the posterior

distribution reveals the nature of this empirical Bayes estimator of the posterior variance: As information about the an individual's data increases, the posterior variance becomes a fraction of the empirical prior variance (Σ_{β}); as information about an individual decreases, this variance approaches the empirical prior variance.

To estimate the mean vector μ_{β} and variance parameters Σ_{β} and σ^2 , MML estimation proceeds by maximizing the log-marginal likelihood of the data from N individuals with respect to these parameters; $\log L = \sum_{i=1}^N \log[h(y_i)]$. Setting the first derivatives of the parameters to zero and solving, the MML solutions are:

$$\hat{\mu}_{\beta} = \frac{1}{N} \sum_i^N \bar{\beta}_i \quad (\text{A7})$$

$$\hat{\Sigma}_{\beta} = \frac{1}{N} \sum_i^N (\bar{\beta}_i \bar{\beta}_i' + \Sigma_{\beta|y_i}) - \hat{\mu}_{\beta} \hat{\mu}_{\beta}' \quad (\text{A8})$$

$$\hat{\sigma}^2 = \left(\sum_{i=1}^N n_i \right)^{-1} \sum_{i=1}^N \text{tr} [(y_i - X_i \bar{\beta}_i)(y_i - X_i \bar{\beta}_i)' + X_i \Sigma_{\beta|y_i} X_i'], \quad (\text{A9})$$

where "tr" represents the trace of a matrix (e.g., the sum of the elements on the main diagonal of a matrix).

The EM solution proceeds by iterating between EB Equations 6 and 7 (or Equations A5 and A6) and MML Equations A7-A9 until convergence. As the EM iterative process uses only information from first derivatives, it is termed a first-order solution and, as a result, is somewhat slow to converge under certain circumstances (see Discussion of Dempster, Laird, & Rubin, 1977; Longford, 1993). Thus, at some point in the iterative procedure, it is desirable to switch to a second-order solution (e.g., using second derivatives in addition to first derivatives), for example, the Fisher scoring solution (Silvey, 1975). The Fisher scoring solution is an iterative process that uses first derivatives and expectations of the second derivatives (the negative of the information matrix) of the

likelihood of the data with respect to the estimated parameters. Specifically, multiplying the vector of first derivatives by the inverse of the information matrix provides the vector of corrections which, added to parameter values, yield improved estimates. From the improved estimates, values for the first derivatives and information matrix are reobtained, yielding further improved estimates. This process is repeated until convergence. At convergence, the inverse of the information matrix, denoted I^{-1} , provides the large-sample variances and covariances of the MML estimates, which can be used to construct confidence intervals and tests of hypotheses for the parameters (Wald, 1943). Specifically, the square root of each diagonal element of I^{-1} provides standard errors for the MML estimates, which are used to determine asymptotically normal test statistics (estimate divided by its standard error) for each parameter. More details regarding the Fisher scoring solution for the random-effects model can be found in Bock (1989a) or Longford (1987).

To test for statistical difference between alternative models using the likelihood-ratio chi-square test, Dempster et al. (1981) show that the log likelihood is given as follows:

$$\begin{aligned} \log L = & -\frac{1}{2} \sum_{i=1}^N n_i \log(2\pi\sigma^2) - \frac{N}{2} \log |\Sigma_{\beta}| \\ & + \frac{1}{2} \sum_{i=1}^N \log |\Sigma_{\beta|y_i}| - \frac{1}{2\sigma^2} \left[\sum_{i=1}^N (y_i - X_i \mu_{\beta})'(y_i - X_i \bar{\beta}_i) \right], \quad (\text{A10}) \end{aligned}$$

where $|\cdot|$ denotes the determinant of a matrix. Comparing this likelihood using the estimated parameters of nested models (say, Model B is nested within A) is then done by calculating $-2(\log L_A - \log L_B)$, which follows the chi-square distribution with degrees of freedom equal to the number of additional parameters in Model A .

Received February 22, 1993

Revision received February 6, 1995

Accepted March 8, 1995 ■