Journal of Gerontology: PSYCHOLOGICAL SCIENCES
1994, Vol. 49, No. 6, P251–P264

Age Differences in the Symptoms of Depression: A Latent Trait Analysis

Joseph J. Gallo,[1] James C. Anthony,[1] and Bengt O. Muthén[2]

[1]The Johns Hopkins University School of Hygiene and Public Health, Baltimore.
[2]University of California, Los Angeles Graduate School of Education.

Our hypothesis was that older adults are less likely than younger adults to acknowledge dysphoria or anhedonia even at the same level of depression. Study subjects were 3,141 participants in Baltimore, Maryland, and 3,469 participants in the Durham-Piedmont region of North Carolina who had complete data on symptoms of depression active in the one month prior to interview, as well as several covariates thought to be related to depression. The effect of age on the endorsement of the dysphoria/anhedonia stem question from the section on Major Depression in the Diagnostic Interview Schedule was estimated in the two independently gathered samples employing structural equations with a measurement model. The results indicate that, even accounting for differences due to overall level of depressive symptoms, as well as gender, minority status, educational attainment, marital status, employment status, and cognitive impairment, dysphoria was less likely to be endorsed by persons 65 years of age and older. This bias against older adults may account in part for the low rates of Major Depression reported for older persons from epidemiologic studies employing the standard diagnostic criteria.

EPIDEMIOLOGIC studies that use the standard diagnostic criteria show a decrease in the prevalence of depression with age, yet clinical studies, and epidemiologic studies using symptom scales, show symptoms of mood disturbance to be more common for older adults (see Newmann, 1989, for a review). Specifically, reports from the Epidemiologic Catchment Area (ECA) Program have suggested that the prevalence of depressive disorders *declines* in late life (Blazer, Hughes, & George, 1987; Eaton et al., 1989; Kramer, German, Anthony, Von Korff, & Skinner, 1985; Myers et al., 1984; Weissman et al., 1988), while studies using symptom scales (e.g., Kessler, Foster, Webster, & House, 1992, and other studies cited by Newmann, 1989) suggest a pattern of *increasing* depression in late life. Newmann (1989) implored researchers in late-life depression to carefully scrutinize measurement of depression at the level of *symptoms* in order to delineate the features of different strategies of measurement which lead to these disparate conclusions. Although age-related differences in measurement might explain the discrepancy between prevalence studies using standard diagnostic criteria and those using symptom scales, few epidemiologic studies have addressed the issue of whether there are age-specific differences in endorsement of the symptoms of depression.

Our analysis was motivated by the realization that the requirement of the dysphoria symptom for the diagnosis of Major Depression in the *Diagnostic and Statistical Manual, 3rd ed.* (DSM-III; American Psychiatric Association, 1980), as reflected in the Diagnostic Interview Schedule (DIS) employed in the ECA, might underestimate prevalence of disorder among older adults if older adults were less likely than the young to endorse dysphoria. This bias might explain the low rates of Major Depression and other affective disorders among older adults previously reported from the ECA. Prior work from the Baltimore ECA, in which every subject who fulfilled DSM-III criteria for Major Depression or Dysthymic Disorder from the DIS was examined by a psychiatrist, offered preliminary evidence that underestimation occurs (Romanoski et al., 1992). As a result of standardized clinical examinations, Romanoski and colleagues concluded that, while prevalence of DSM-III Major Depressive Disorder declined with age, prevalence of all other DSM-III depressive disorders *increased* with age, even after adjustment for sociodemographic variables and recent life events.

The occurrence of anxious, somatic, or hypochondriacal complaints without dysphoria in older patients felt to be clinically depressed has been observed by experienced clinicians (e.g., Fogel & Fretwell, 1985; Salzman & Shader, 1978). Cognitive theories of depression, combined with an appreciation of the diminution of physical abilities with advanced age, suggest that feelings of helplessness (Depue & Monroe, 1978) and hopelessness (Abramson, Metalsky, & Alloy, 1989) may play a more central role than for younger persons. Other aspects of depression that may be characteristic for older adults, such as perceived cognitive deficit (Weiss, Nagel, & Aronson, 1986) or irritability (Rohrbaugh, Siegal, & Giller, 1988), are not included in the standard criteria. Anxiety symptoms may accompany depression, persisting even after the depressive symptoms have improved (Alexopoulos, 1990; Blazer, Hughes, & Fowler, 1989). Fogel and Fretwell (1985) observed that since many depressed older adults do not complain of "depression," a diagnosis of depression emphasizes a symptom which does not speak to the illness experience of the elder. They suggested that reluctance on the part of older adults to accept the term "depression" may be specific to the current cohort of older people, and instead offered an ostensibly more acceptable term to describe this state, the "depletion syndrome."

Clinical lore suggesting that older adults are less likely than the young to assent to dysphoria has been codified in recent documents dealing with depression produced by several

#52

prominent organizations. The National Institutes of Health Consensus Statement on Late-life Depression asserted that dysphoria, the "typical signature of depression," may be less prominent for older adults in comparison to other symptoms such as sleeplessness (NIH Consensus Development Panel on Depression in Late Life, 1992). The "Practice Guideline for Major Depressive Disorder in Adults" of the American Psychiatric Association states that, compared to younger persons with depression, older persons tend to complain less of "subjective dysphoria" (American Psychiatric Association, 1993). Finally, guidelines for the detection of depression in primary care settings emphasize the dysphoria and anhedonia criterion by recommending that sadness and loss of interest or pleasure be the first symptoms elicited in history taking (Depression Guideline Panel, 1993).

To return to the measurement theme, Newmann (1989) clearly sketched the philosophical differences of the psychometric tradition, which uses symptom scales and assumes that depression varies on a continuum, and the clinical tradition, which emphasizes "caseness" and assumes that depression is a disease that is either "present" or "absent" (see Klerman, 1989; Mirowsky & Ross, 1989; Swartz, Carroll, & Blazer, 1989; Tweed & George, 1989, for spirited exchange regarding these contrasting perspectives). From the perspective of public mental health, the case can be made that the psychometric approach has distinct advantages with regard to understanding psychologic distress in populations and to developing models of prevention (Anderson, Huppert, & Rose, 1993; Goldberg & Huxley, 1992; Rose, 1993). Mirowsky and Ross (1989) draw an analogy with the measurement of weight:

> Although the bathroom scale rarely tells a person's true weight, usually it gives correctly the five-pound or 10-pound range in which the person's weight belongs. The broader the ranges, the more likely it is that a scale will give the correct range. The broadest range would sort everybody into two categories: heavy and not heavy. Almost everyone would be classified correctly. The current bathroom scale, which almost never shows a person's true weight, can be improved so that it shows only whether or not a person is heavy. Now suppose you are trying to lose weight. Which scale would you choose? The old one, which is almost never correct, or the new and improved one, which is almost always correct? (p. 18)

This dimensional perspective is made particularly poignant by recent studies suggesting that symptoms of depression are enough to signal poor functional outcomes even in the absence of disorder; indeed, these studies suggest that the majority of disability in the population attributable to depression occurs among persons who do not meet standard criteria for depression (Horwath, Johnson, Klerman, & Weissman, 1992; Wells, Burnam, Rogers, Hays, & Camp, 1992).

Our perspective lines up most closely with the psychometric tradition and is guided by the model presented by Goldberg and Huxley (1992) and Duncan-Jones, Grayson, and Moran (1986) and applied to public mental health by Anderson et al. (1993). For this study, however, we constrain the "depression as continuum" model within a framework that is grounded in DSM-III, since the DIS items employed in the measurement of depression in the ECA are so closely linked to the DSM-III criteria of Major Depression. We apply a continuous rather than the usual categorical measurement model to the DIS, asking the question posed by Newmann (1989): ". . . are there features of the measurement process . . . that introduce bias in estimates of true levels or rates of depression among contrasting age cohorts?" We are interested in examining possible bias with regard to age in the dysphoria/anhedonia item of the DIS. In doing so, while we hope to shed some light on how depressive symptoms differ across age groups, we recognize that there is the larger issue of the structure of depression. Our focused aim in this analysis was to examine the measurement characteristics of the depression items of the DIS, short of clear evidence for more than one factor, within the context of the measurement of a single continuous dimension.

Several alternative strategies might have been used to compare the symptoms of depression in different age groups, some of which were outlined by Weiss and colleagues (1992). First, a simple comparison of the prevalence of symptoms across age groups is not helpful because age-related changes in the level of specific symptoms may be unrelated to depression (e.g., Berry, Storandt, & Coyne, 1984). For example, somatic symptoms may simply reflect the prevalence of physical illness, not depression. Also, if a symptom decreases with age in general but increases with depression in the aged, the overall trend might be no average change with age. Second, if correlations of individual symptom items with an item-total score are examined to determine differences by age group, this has the effect of weighing each item equally, when in fact, particular items, such as thoughts of death or suicide, should be given more weight. A similar approach involves stratification on the total score and examining prevalence of the item of interest among different groups at levels of the summary score (e.g., Arean & Miranda, 1992; Zwick & Ercikan, 1989). Third, logistic regression, with the item response as the dependent variable, can be used to study the relationship of covariates to individual items (e.g., Anthony & Aboraya, 1992), but cannot simultaneously adjust for the level of latent variables. Fourth, latent class analysis, which assumes the latent construct is a category, rather than a dimensional "trait," can examine the probabilities of response to a given symptom in different age groups (e.g., Heithoff, 1992); however, it may be difficult to incorporate more than one covariate into the model (because to do so it is necessary to subdivide the sample for every covariate). Finally, comparing the endorsement of dysphoria or anhedonia by older "cases" of depression with younger patients (e.g., Museti et al., 1989; Oxman, Barrett, Barrett, & Gerber, 1990) begs the question of differential reporting of dysphoria and anhedonia by age if criteria for "caseness" *require* the presence of dysphoria or anhedonia.

The approach we have taken to address these difficulties, the MIMIC (multiple indicators, multiple causes) model, allows simultaneous factor analysis and regression of factor scores on covariates for the comparison of item functioning across groups, while accounting for differences in several personal characteristics such as gender (Muthén, 1989a). Unlike multigroup factor analysis (e.g., Horn & McArdle, 1992), several covariates can be incorporated into the MIMIC model without subdividing the sample. In an educa-

tional context the MIMIC model has been used to detect differential item functioning in mathematics test questions, with implications for the construction of tests evaluating mathematics competency. For example, girls answer certain questions less well than boys, even at the same level of algebra skill, and accounting for the influence of other characteristics which might affect mathematics ability, such as opportunity to learn the material (Muthén, 1988a; Muthén, Kao, & Burstein, 1991). Algebra questions that require translation of verbal problems into algebraic expressions seemed to be particularly biased against girls; in other words, such items test an additional skill unrelated to the "algebra achievement" construct. In an illustration of the MIMIC model employing data on symptoms of anxiety and depression, Muthén showed that despite a higher anxiety factor mean for women than men, the endorsement of fear of heights was less for women than would be expected (Muthén, 1989b). Further analyses illustrate the utility of a multi-factor MIMIC model in delineating the correlates of the factors that underlie the criteria for alcohol abuse and dependence (Muthén, Grant, & Hasin, 1993).

Working within the framework of DSM-III, our hypothesis was that older adults are less likely than younger adults to acknowledge dysphoria or anhedonia even at the same level of depression, in much the same way that cultural groups or persons experiencing chronic pain differ from others in reporting these symptoms. Gender, minority status, educational attainment, recent unemployment, marital status, and cognitive status influence the tendency to assent to dysphoria and might produce differences ascribed to age. In this study, our model accounts for these covariates, as well as for the level of depression, without requiring that predetermined diagnostic criteria be met.

Our investigation differs in several ways from prior work on depression in older adults. First, this study is based on the ECA Program, a multisite collaborative epidemiologic survey using a standardized interview keyed to diagnostic criteria of the DSM-III (American Psychiatric Association, 1980). In consequence, the response characteristics of the DSM-III criterion for dysphoria/anhedonia can be examined directly and within the context of the set of standard criteria measuring a single continuous dimension. Second, the symptom data were gathered in a large epidemiologic study of adults over the age of 18 years, and are not restricted to persons in treatment, to specific sociodemographic groups, or to persons who meet diagnostic criteria for Major Depression (e.g., Ben-Arie, Swartz, & Dickman, 1987; Berry et al., 1984; Good, Vlachoniklolis, Griffiths, & Griffiths, 1987; Koenig, Cohen, Blazer, Ranga Rama Krishnan, & Sibert, 1993; Newmann, Engel, & Jensen, 1990; Oxman et al., 1990; Zemore & Eames, 1979). Finally, the analysis makes use of the MIMIC model, a special application of the latent trait model with distinct advantages over other strategies used to study and conceptualize psychologic distress (Duncan-Jones et al., 1986; Goldberg & Huxley, 1992) or used to study heterogeneity in measurement (Muthén, 1989a).

METHODS

The Epidemiologic Catchment Area Program was a series of epidemiologic surveys conducted by collaborators between 1980 and 1984 at five sites in the United States: New Haven, CT (Yale University), Baltimore, MD (Johns Hopkins University), St. Louis, MO (Washington University), Durham-Piedmont, NC (Duke University), and Los Angeles, CA (University of California, Los Angeles). At each site, ECA collaborators used multistage probability sampling to select 3000 to 5000 adult respondents, and then applied the standardized Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981). The standardized questions of the DIS were keyed to individual criteria of the DSM-III case definition for Major Depression.

Positive responses to the questions were followed by further questioning to determine whether a threshold for severity had been met and whether the symptom was plausible as a psychiatric symptom; that is, the symptom could not be explained by physical illness, medications, alcohol, or drug use (Robins et al., 1981; Robins & Regier, 1991). Finally, respondents rated the onset and recency of *syndromes*, that is, clusters of symptoms, so that the timing of disorders could be estimated. To meet the DSM-III diagnostic criteria for Major Depression, an individual must have depressed mood (dysphoria) or loss of interest in things normally enjoyed (anhedonia) for two weeks or more, and at least four of eight additional criteria; however, a diagnosis of Major Depression was not required for inclusion in our analysis.

In Baltimore and in Durham-Piedmont, information on onset and recency was also gathered at the *symptom level*, so that data on recency of individual symptoms are available for the criteria of DSM-III Major Depression (Von Korff & Anthony, 1982). If the respondent "ever" had the symptom, the respondent was then asked when the last time the symptom was present. Detailed descriptions of the sampling design, the diagnostic assessments, and other aspects of the ECA have been published elsewhere, including several studies of DIS reliability and validity (Anthony et al., 1985; Eaton & Kessler, 1985; Helzer et al., 1985; Robins et al., 1981). Only symptoms present in the one month prior to the interview were rated as "present" for the current analysis.

Analytic Strategy

The analysis was restricted to household data from the Baltimore and Durham-Piedmont ECA sites. The availability of two sets of data allowed us to check our findings in two independently gathered samples. At the Baltimore and Durham-Piedmont sites and no others, procedures for oversampling the elderly were developed in order to ensure adequate numbers of older people (Helzer et al., 1985). Only persons with complete data on symptoms along with the covariates of age, gender, minority status, educational status, employment status, marital status, and mental status score were included in this analysis (Table 1). The implication of selecting only individuals with complete data is considered in our discussion of study results.

The MIMIC model, described below, includes dichotomous variables for age (age less than 65 years = 0; age 65 years and older = 1), sex (male = 0; female = 1), minority status (White = 0; African American or other minority = 1), educational status (12 or more years of schooling = 0; fewer than 12 years = 1), employment status (worked for

Table 1. Totals and Percent Available for Analysis From Baltimore and Durham-Piedmont ECA Study Samples

	Baltimore		Durham-Piedmont	
	Total from ECA Study Sample	% Available for the Present Analysis	Total from ECA Study Sample	% Available for the Present Analysis
Age ≥ 65 years	923	78.3	1241	80.6
Females	2159	88.9	2371	87.1
Minority	1288	92.2	1439	86.9
Education < H.S.	1886	86.2	1874	84.3
MMSE < 24	368	65.2	806	71.2
Unemployed	1831	93.7	1777	91.2
Not married	1908	89.4	1927	84.9
Total	3141	90.3	3469	88.5

pay within the 6 months prior to interview = 0; last worked for pay more than 6 months prior to interview = 1), and marital status (married or living with someone as though married = 0; not currently married = 1). We coded the employment variable with a 6-month threshold in order to exclude persons who were interviewed during a period of temporary unemployment. The Mini-Mental State Examination (MMSE) score was entered into the MIMIC model as a continuous variable (range, 0–30) (Folstein, Folstein, & McHugh, 1975). The decisions on how to code the variables were reached after considering what variables might modify the tendency to endorse sadness or loss of interest or pleasure, but before any data analysis was completed.

Latent Trait Analysis

Lazarsfeld and Henry (1968) and, later, Bartholomew (1987) defined latent variables as unobserved variables that explain the association among a set of observed variables. This definition is consistent with the notion of a "factor" in factor analysis: the factor "explains" the observed correlations in the observed variables. The approach in standard factor analysis assumes that variables are normally distributed and related to each other in a linear and additive fashion. However, new tools allow a more complete use of symptom data, which are often dichotomous.

The essence of the latent trait model is the estimation of the level of a continuous latent variable (the latent "trait") from dichotomous data (that is, data expressed as symptom present or absent). The term "trait" as used here simply denotes a *continuous* latent variable in contrast to a *categorical* one, and does not indicate a fixed quality of the individual (e.g., a "personality trait").

We use the formulation of the latent trait model developed by Muthén (1983, 1984, 1988b). Under this model, the underlying latent trait, say, depression, is denoted by η. We do not observe an individual's level on η directly, but instead observe dichotomous indicator variables, y, which provide information about η. Grayson, Henderson, and Kay (1987) liken this to ascertaining height and weight, continuous variables, by asking a series of yes/no questions such as, "Are you tall?", "Are you heavy?", and "Are you large?" The first question ascertains information specific to height, the second to weight, and the third, while not specific to height or weight, is descriptive of height and weight nonetheless.

The relationship of the observed dichotomous variables to the unobserved latent variables is given by a system of equations called the *measurement model:*

$$\mathbf{y} = \nu + \Lambda\eta + \epsilon \quad (1)$$

where ν are intercepts, Λ relates the observed variables to the latent traits, η, and ϵ contains the error terms. Relationships among the latent variables as well as observed covariates, $\mathbf{x}$, on which the latent variables are regressed, are expressed in a second, related system of equations called the *structural equation model:*

$$\eta = \alpha + \beta\eta + \Gamma\mathbf{x} + \zeta \quad (2)$$

where α are intercepts, β represent the effects of latent traits on one another, Γ contains the coefficients for the regression of η on $\mathbf{x}$, and ζ are error terms. The error terms, ϵ and ζ, are assumed to be uncorrelated with each other and with η. The variances and covariances of the latent variables are contained in an additional matrix, Ψ.

The MIMIC Model

The MIMIC model is a special application of the latent trait model which enables item bias to be detected across age groups (Muthén, 1989a). A straightforward application of the latent trait framework permits inference about the personal characteristics that influence reporting of depressive symptoms. Since the MIMIC model has not previously been applied in aging research for the evaluation of differential item functioning, we discuss it in further detail in the Appendix.

The latent trait in the MIMIC model, η_0, is defined for this analysis by the symptoms which form the diagnostic criteria for depression from the DSM-III (Figure 1). The observed variables, y, are the symptom items from the DIS. For the purposes of the MIMIC model, the DIS items are considered to be indicators of a set of latent variables η_1 through η_{15}, representing the symptoms from the diagnostic criteria of the DSM-III. (One item was dropped from the MIMIC model, leaving 15 measurement items, as discussed in detail in Results.)

The MIMIC model assumes that the DIS/DSM-III criteria for depression constitute an adequate model for a unidimensional construct of depression across all age groups. The latent variables η_1 through η_{15} serve as indicators of the single latent dimension of depression, η_0, so that this MIMIC model is consistent with a unidimensional model of depression. (The assumption of a unidimensional model for these items will be evaluated through an examination of the principal components of the covariance matrix.)

We introduce covariates to the factor model in order to examine differential item functioning by age, adjusting for differences in the level of the latent trait and for the effect of other covariates such as gender. Differential item functioning or item bias is present if individuals in different groups at the same level of a trait or ability do not have the same probability of answering the item in the affirmative (Duncan-

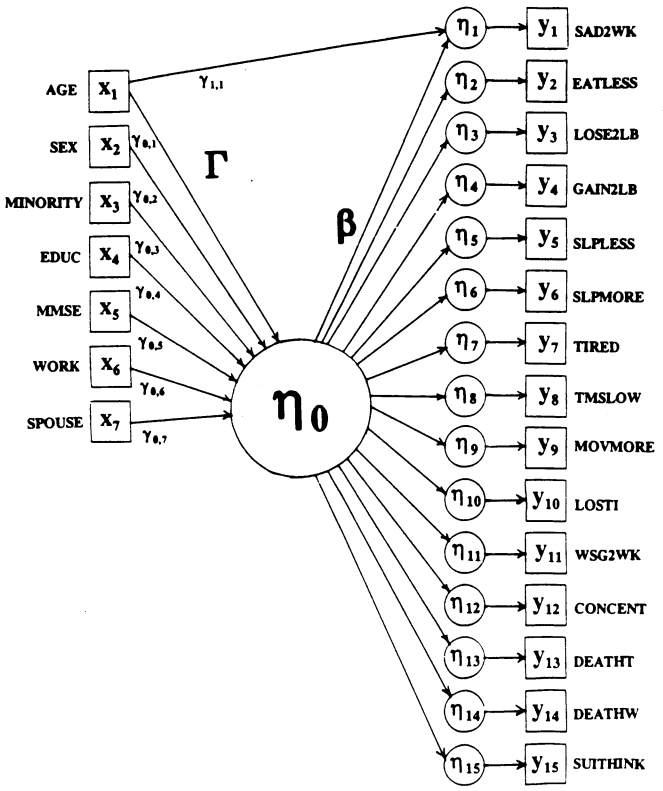

Figure 1. Path diagram for the MIMIC model. The y's represent the diagnostic items from the DIS for Major Depression, each assumed to measure an associated η. Coefficients for the regression of the latent constructs on the x covariates are contained in Γ (Table 4). The factor loadings for the underlying latent variable are contained in β (Table 5).

Jones et al., 1986). In addition to age, six covariates are introduced to adjust for differences in level of depression due to gender (x_2), minority status (x_3), educational status (x_4), MMSE score (x_5), employment status (x_6), and marital status (x_7), as defined earlier. In summary, the MIMIC model consists of three parts:

(1) A *measurement model* (a continuous latent variable underlies the dichotomous symptom responses; on the right side of Figure 1 the measurement model relates the observed y's to the unobserved latent trait, η_0),
(2) A *regression model* (analogous to multiple regression of the continuous outcome variable onto several covariates; on the left side of Figure 1 the regression model relates the observed outcome variable, η_0, to the covariate x's), and finally,
(3) A *"direct effect" estimate* (to detect any residual variance in item response associated with membership in a particular group; the path at the top of Figure 1 associated with the parameter $\gamma_{1,1}$ relates the covariate of interest to the item of interest).

Assessing Model Fit

The usual tests, such as χ^2, for assessing fit of a model to empirical data pose some difficulty in large data sets because even small departures can result in rejection of an adequately fitting model. A large sample will have a great deal of power to reject the hypothesized model. For this reason, we use several approaches to assessing model fit. Appropriate tests for the adequacy of structural equation models are still under development (Bollen & Long, 1993).

One simple alternative approach to using the solitary χ^2, the Descriptive Fit Value (DFV, Muthén, 1989b), uses the χ^2, but takes sample size and degrees of freedom into account. The DFV is calculated by dividing the χ^2 by the product of the degrees of freedom and sample size. The DFV does not provide a probability statement for testing of a hypothesis, but is descriptive of the model fit, with values above 1.5 indicating that the model can be substantially improved or that the variables are not suited to factor analysis.

The overall fit of a model can be assessed by comparing the estimated sample covariance matrix, **S**, with the predicted covariance matrix, Σ, derived from the model and its parameter estimates. The closer Σ is to **S**, the better the fit of the model. The Goodness-of-Fit Index (GFI) makes use of this idea (Alwin, 1988). If the data fit the model perfectly, the GFI equals 1. The GFI will approach 0 as the model fit deteriorates. The Adjusted GFI (AGFI) accounts for the number of variables in the covariance matrix and the degrees of freedom in the model, thus "rewarding" parsimony in model specification (Alwin, 1988). Finally, the Critical Number, CN, differs from other measures of model fit in using sample size as a measure of fit (Hoelter, 1983; Wheaton, 1988). The CN is interpreted as the sample size above which a particular model is likely to be rejected under the null for the χ^2 test, despite a reasonable fit. A model with an acceptable fit should have a CN of at least 200 (Wheaton, 1988).

Model Estimation Procedure

Models were estimated using the LISCOMP program's limited-information generalized least squares estimator for dichotomous response (Muthén, 1988b, 1989b). The MIMIC analysis takes advantage of a feature of LISCOMP which allows regression of latent variables onto observed x covariates; at the same time, the latent factor underlying the observed dichotomous y's is included in the model. This structural analysis has the advantage of avoiding normality assumptions regarding tetrachoric correlations, employing instead the conditional normality assumption of regression analysis (Muthén, 1989b, 1993). If the x covariates are not normally distributed (which they are not if the x covariates are categorical), the distribution of the latent variable they predict also will be nonnormal, following the nonnormal distribution of the x's (Muthén, 1989a).

RESULTS

A total of 3,481 subjects were interviewed in the household survey in Baltimore, while 3,921 subjects compose the household sample in Durham-Piedmont. After excluding persons without complete information on DIS symptom data and all covariates, 3,141 subjects remained for analysis from the Baltimore site and 3,469 subjects remained from the Durham-Piedmont site (because of limitations on computer memory, a single random sample of 3,400 subjects (i.e., a 98% random sample) of the 3,469 subjects at the Durham-

Table 2. Characteristics of the Samples. Percents Represent Site- and Age-Specific Proportions of Given Characteristics

	Baltimore		Durham-Piedmont		Baltimore		Durham-Piedmont	
	Age < 65 years		Age < 65 years		Age ≥ 65 years		Age ≥ 65 years	
Total household sample:	2558		2680		923		1241	
Total available for analysis:	2418		2469		723		1000	
	n	%	*n*	%	*n*	%	*n*	%
Female	1491	61.7	1408	57.0	428	59.2	656	65.6
Minority	1023	42.3	909	36.8	165	22.8	342	34.2
Less than high school education	1060	43.8	859	34.8	565	78.2	721	72.1
Not working for pay	1030	42.6	706	28.6	685	94.7	914	91.4
Not living with spouse	1310	54.2	1019	41.3	395	54.6	617	61.7
MMSE score less than 24	101	4.2	230	9.3	139	19.3	344	34.6

Note. Data gathered from the Baltimore, Maryland and Durham-Piedmont, North Carolina Epidemiologic Catchment Area Program sites, 1980–1984.

Piedmont site was analyzed). Table 2 enumerates the samples by site, age, and personal characteristics. The sites were comparable with respect to gender, minority status, educational attainment, and marital status. Proportionately more younger persons in Baltimore reported unemployment, and a higher proportion in both young and old groups in Durham-Piedmont scored 24 points or less on the MMSE.

Table 3 contains the wording of the DIS questions used to make operational the DSM-III criteria for depression as well as the one-month prevalence of individual items. Inspection of the prevalences shown in the table reveals that the one-month prevalence for any individual DIS symptom was rare, about 8% or less. Sleep disturbance was the most common symptom at both sites.

Inclusion of two items, concentration difficulty (CONCENT) and thinking more slowly than usual (THINKSLO), was found to result in a model which could not be estimated due to system limitations. The item eliciting the symptom "thinking more slowly than usual" was dropped from the MIMIC model. A separate analysis showed that results were substantially unchanged if THINKSLO was substituted for CONCENT in the analysis.

Latent roots for the sample correlation matrix of the y's (that is, the items of the DIS) were obtained from exploratory factor analysis in LISCOMP, equivalent to principal components analysis for dichotomous data. Parsimony in model specification would prescribe the smallest number of latent variables possible. The scree plot of the latent roots derived from the covariance matrix of the 15 symptoms of depression (Figure 2) was consistent with a unidimensional factor structure at both ECA sites (Cattell, 1965; Edelbrock, 1987; Gibbons, Clark, Cavanaugh, & Davis, 1985). The first eigenvalue accounted for 54.7% of the variance in the Baltimore data and for 51.5% of the variance in the Durham-Piedmont data.

The parameter estimates for the MIMIC models using DIS data are presented in Table 4. Examining first the parameter estimates from the Baltimore site, there was evidence for differential item response to the dysphoria/anhedonia item by age group. After adjustment for other covariates which may mediate differential item functioning by age, the estimated coefficient for a direct effect of age on differential item functioning for dysphoria/anhedonia was significantly less than the null value (estimated difference, $\gamma_{1,1} = -.169$, 95% confidence interval (CI) [−.332, −.006]). The estimated value of this coefficient from the analysis of data from the Durham-Piedmont site affirms the Baltimore finding (estimated difference, $\gamma_{1,1} = -.294$, CI [−.439, −.149]), despite different factor levels across the two sites (see below). Thus, our initial hypothesis of differential item functioning by age in the dysphoria/anhedonia item of the DIS was supported.

The factor level for the elderly compared to younger persons is captured by the term $\gamma_{0,1}$. The estimated factor level was lower among older adults than among younger adults in the Baltimore sample (estimated age difference in factor level, $\gamma_{0,1} = -.346$, CI [−.460, −.232]). In the Durham-Piedmont sample, the estimated factor level did not significantly differ for persons aged 65 and older compared to persons less than 65 years old (estimated age difference in factor level, $\gamma_{0,1} = -.045$, CI [−.143, +.005]). Thus, older adults at each site were less likely to endorse sadness or loss of interest or pleasure, even accounting for differences due to the other x covariates in the MIMIC model, and despite different factor mean levels at the two sites. While factor levels were lower for older adults at the Baltimore site, the significant diminished estimated direct effect of age group on the dysphoria/anhedonia DIS item means that the endorsement is even less than expected given the decreased level of the latent trait of depression among persons aged 65 years and older. At Durham-Piedmont, the estimated mean level of depression from young to old did not significantly differ, yet, as in Baltimore, older adults were less likely to endorse sadness or loss of interest or pleasure.

The β coefficients of the MIMIC model, similar to the factor loadings in factor analysis, are also of interest. The β coefficients in Table 5 may be thought of as "weights" representing the "strength" of the relationship of an item to the underlying construct, in this case, DIS depression (represented in Figure 1 by η_0). The β coefficient for the "tired" symptom was fixed to 1.0 in order to identify the MIMIC model; that is, to provide a "metric" for the latent variable η_0. The results are substantially unchanged when the β corresponding to the symptom for "thinking about suicide" was fixed to 1.0 with β otherwise free. As might be ex-

Table 3. One-month Prevalence of DIS Symptoms

Diagnostic Interview Schedule Items based on *Diagnostic and Statistical Manual* (3rd ed.)	One-month Prevalence (per 100 individuals)			
	Baltimore		Durham-Piedmont	
	Age < 65 years	Age ≥ 65 years	Age < 65 years	Age ≥ 65 years
Dysphoria/anhedonia				
72. Have you ever had two weeks or more during which you felt sad, blue, depressed, or when you lost all interest and pleasure in things that you usually cared about or enjoyed? (SAD2WK)	4.4	2.2	3.6	2.9
Appetite disturbance				
74. Has there ever been a period of two weeks or longer when you lost your appetite? (EATLESS)	1.9	1.2	1.1	2.4
75. Have you ever lost weight without trying to — as much as two pounds a week for several weeks (or as much as 10 pounds altogether)? (LOSE2LB)	1.6	0.4	0.9	0.6
76. Have you ever had a period when your eating increased so much that you gained as much as two pounds a week for several weeks (or 10 pounds altogether)? (GAIN2LB)	4.2	1.7	2.2	0.7
Sleep disturbance				
77. Have you ever had a period of two weeks when you had trouble falling asleep, staying asleep, or waking up too early? (SLPLESS)	8.8	8.7	7.0	12.2
78. Have you ever had a period of two weeks or longer when you were sleeping too much? (SLPMORE)	2.8	1.2	1.7	1.1
Fatigue				
79. Has there ever been a period lasting two weeks or more when you felt tired out all the time? (TIRED)	7.4	5.5	6.1	7.1
Psychomotor retardation or agitation				
80. Has there ever been a period of two weeks or more when you talked or moved more slowly than is normal for you? (TMSLOW)	3.1	2.2	1.6	3.0
81. Has there ever been a period of two weeks or more when you had to be moving all the time — that is, you couldn't sit still and paced up and down? (MOVMORE)	3.8	1.7	2.1	2.7
Loss of libido				
82. Was there ever a period of several weeks when your interest in sex was a lot less than usual? (LOSTI)	2.7	1.0	1.5	1.0
Feelings of worthlessness, sinfulness, or guilt				
83. Has there ever been a period of two weeks or more when you felt worthless, sinful, or guilty? (WSG2WK)	3.0	1.9	2.2	1.4
Trouble thinking				
84. Has there ever been a period of two weeks or more when you had a lot more trouble concentrating than is normal for you? (CONCENT)	3.8	1.9	1.9	2.6
85. Have you ever had a period of two weeks or more when your thoughts came much slower than usual or seemed mixed up? (THINKSLO)	3.0	3.0	1.7	3.0
Thoughts of death or suicide				
86. Has there ever been a period of two weeks or more when you thought a lot about death — either your own, someone else's, or death in general? (DEATHT)	8.1	7.0	5.2	8.3
87. Has there ever been a period of two weeks or more when you felt like you wanted to die? (DEATHW)	2.0	1.7	0.9	1.4
88. Have you ever felt so low you thought of committing suicide? (SUITHINK)	0.8	0.1	0.7	0.5

Note. If the respondent "ever" had the symptom, the respondent was then asked when the last time the symptom was present. Only symptoms present in the one month prior to the interview were rated as "present" for the current analysis. Data gathered from the Baltimore, Maryland and Durham-Piedmont, North Carolina Epidemiologic Catchment Area Program sites, 1980–1984.

pected, the items referring to death wishes or to thoughts of suicide are given relatively more weight than other items in determining the level of depression.

The estimated direct effect of age on the dysphoria/anhedonia DIS item found in Baltimore was replicated in the independently surveyed Durham-Piedmont sample. In addition to replication of results in independent samples, fit of the model was evaluated through the use of the DFV, GFI, AGFI, and CN indices described earlier (see Table 4). Although simple application of the χ^2 value would result in rejection of the null hypothesis, the DFV for the models were much less than 1.5, suggesting that the χ^2 value for the model relative to the number of degrees of freedom would not result in rejection of the model when the large sample

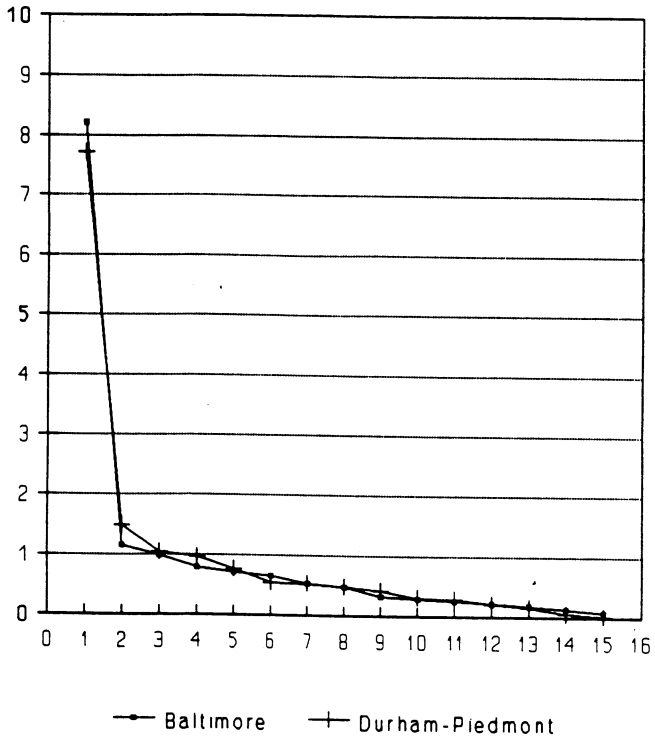

Figure 2. Scree plot of eigenvalues for 15 items of DIS Major Depression. Data gathered from the Baltimore and Durham-Piedmont ECA program sites, 1980–1984.

size is taken into account. The sample sizes at both ECA sites are well above the CNs for the models, which also suggests good fit. The GFI calculated for the Baltimore MIMIC model was .93 (AGFI = .96), indicating good reproduction of the estimated sample covariance matrix from the model-predicted covariance matrix. The GFI for the Durham-Piedmont MIMIC model of .64 (AGFI = .77) indicated that the predicted covariance matrix did a less adequate job of reproducing the estimated sample covariance matrix in the Durham-Piedmont data than in the Baltimore data. We did not attempt to improve model fit by introducing post hoc correlated measurement errors in the y's, which would be equivalent to positing more than one factor. There were no substantial changes in the significance of the parameter estimates obtained with separate MIMIC models that included age as a continuous variable, or with models that excluded the MMSE (the only quasi-continuous variable in the model).

Finally, using separate MIMIC models, we estimated the direct effect of membership in age group 65 years and older on the remaining items in the DIS/DSM-III criteria for Major Depression. In Figure 1, "sliding" the pointer from x_1 to each individual DIS item (first to η_2, for EATLESS, then to η_3, for LOSE2LB, and so on), estimating the direct effect of age group on item endorsement for each item of the DIS at each step, results in the estimates for direct effects of age on each DIS/DSM-III criterion for Major Depression found in Table 6.

Table 4. Parameter Estimates and Indices of Model Fit for MIMIC Models

	Baltimore	Durham-Piedmont
$\gamma_{1,1}$ Estimated direct effect of age group 65 years and older (reference group: age less than 65 years)	−0.169* [−0.332, −0.006]	−0.294* [−0.439, −0.149]
Estimated differences in latent trait of depression according to potentially influential covariates under study		
$\gamma_{0,1}$ Age 65 years and older (reference group: age less than 65 years)	−0.346* [−0.460, −0.232]	−0.045 [−0.143, 0.005]
$\gamma_{0,2}$ Female (reference group: males)	+0.220* [0.130, 0.310]	+0.184* [0.262, 0.106]
$\gamma_{0,3}$ African American or other minority (reference group: White)	+0.114* [0.030, 0.199]	+0.035 [−0.005, 0.115]
$\gamma_{0,4}$ Education less than high school (reference group: high school education and beyond)	+0.148* [0.062, 0.234]	+0.234* [0.150, 0.318]
$\gamma_{0,5}$ MMSE score (continuous score, range 0 to 30)	−0.012* [−0.024, −0.0002]	−0.016* [−0.026, −0.006]
$\gamma_{0,6}$ Unemployed more than 6 months (reference group: employed)	+0.269* [0.175, 0.363]	+0.117* [0.029, 0.205]
$\gamma_{0,7}$ Not married (reference group: married)	+0.126* [0.042, 0.210]	+0.180* [0.104, 0.256]
Degrees of freedom	187	187
Measures of Model Fit		
χ^2	506.6	696.7
DFV	0.001	0.001
CN	1402	1126
GFI	0.93	0.64
AGFI	0.96	0.77

Note. Data gathered from the Baltimore, Maryland and Durham-Piedmont, North Carolina Epidemiologic Catchment Area Program sites, 1980–1984. Measures of model fit are defined in the text. The 95% confidence intervals for the estimated MIMIC model coefficients are given in brackets.
*$p < .05$.

Table 5. Coefficients of β, the Factor Loadings for the Dichotomous Factor Analysis of DIS Items

	Baltimore	Durham-Piedmont
SAD2WK	1.019	1.004
EATLESS	0.889	1.093
LOSE2LB	0.805	0.831
GAIN2LB	0.766	0.413
SLPLESS	0.828	0.783
SLPMORE	0.776	0.915
TIRED	1.000[a]	1.000[a]
TMSLOW	0.965	1.096
MOVMORE	0.876	0.999
LOSTI	0.766	0.914
WSG2WK	1.124	1.036
CONCENT	1.021	1.080
DEATHT	0.910	0.989
DEATHW	1.180	1.257
SUITHINK	1.172	1.368

Note. Data gathered from the Baltimore, Maryland and Durham-Piedmont, North Carolina Epidemiologic Catchment Area Program sites, 1980–1984. The variable names refer to the DIS items listed in Table 3.

[a]The value for this coefficient was fixed to 1.0 in order to identify the MIMIC model.

Table 6. Estimates for Direct Effect of Age Group 65 Years and Older on the Items of DIS/DISM-III Major Depression, Adjusted for Covariates (Gender, Minority Status, Education, MMSE, Unemployment, and Marital Status)

DIS/DSM-III item	Baltimore	Durham-Piedmont
SAD2WK	−0.169*	−0.294*
EATLESS	−0.256*	−0.032
LOSE2LB	−0.495*	−0.019
GAIN2LB	−0.319*	−0.296*
SLPLESS	+0.142*	+0.327*
SLPMORE	−0.206	+0.062
TIRED	+0.142*	−0.116
TMSLOW	+0.034	+0.297*
MOVMORE	−0.219*	+0.037
LOSTI	−0.015	−0.182
WSG2WK	−0.016	−0.423*
CONCENT	−0.132	−0.182*
DEATHT	+0.162*	+0.140*
DEATHW	+0.257*	+0.258*
SUITHINK	−0.202	−0.289*

Note. Variable names refer to DIS items listed in Table 3. Data gathered from the Baltimore, Maryland and Durham-Piedmont, North Carolina Epidemiologic Catchment Area Program sites, 1980–1984.

*$p < .05$.

The estimated parameters revealed differential item functioning by age for several items in addition to dysphoria/anhedonia. For example, adjusting for gender, minority status, educational status, MMSE score, employment status, marital status, and the level of depression, older adults in Baltimore were significantly *more* likely than younger persons to assent to difficulty with sleep, feeling tired, thinking about or even wishing for death, and, in addition to dysphoria/anhedonia, were *less* likely to assent to anorexia, weight change, or agitation (Table 6). In Durham-Piedmont, older adults were *more* likely than younger persons to assent to difficulty with sleep, talking or moving more slowly than usual, thinking about or wishing for death, and, in addition to dysphoria/anhedonia, were *less* likely to assent to gaining weight, feeling worthless, sinful, or guilty, having trouble concentrating, or having thoughts of suicide (Table 6).

Discussion

In the community samples studied, the dysphoria/anhedonia criterion for DIS Major Depression was less likely to be endorsed by persons 65 years of age and older, even adjusting for differences due to level of a latent trait of depression and for covariates associated with depression. To derive this result, we used a variation of factor analysis for dichotomous variables that included regression of potentially influential covariates on factor scores, for the comparison of item functioning across groups. As an analytic approach, the MIMIC technique is generally applicable to situations in which it is necessary to adjust for levels of depression, as well as for confounding covariates, in examining item-level differences among groups (be the group defined by age, culture, physician recognition of psychological distress, and so on). Few prior studies have focused on age differences in the symptoms of depression, particularly with regard to a key symptom such as dysphoria, as we have done here. In accord with Costello's (1993a, 1993b) and Newmann's (1989) assertion that an understanding of depression is best approached through study of the origins and correlates of its component symptoms, we focused on a key symptom in the diagnostic criteria, namely, dysphoria.

Before discussing the implications of this finding, the limitations of the study require comment. First, the DIS items are based on psychiatric symptoms reported to a lay interviewer, and are subject to imperfect recall, socially desirable responding, and other sources of error in retrospective interview data (Anthony et al., 1985; Eaton & Kessler, 1985; Helzer et al., 1985; Robins et al., 1981). To some degree, we limit error due to these factors by restricting the analysis to symptoms active in the one month prior to the interview. Second, the wording of the DIS stem question for Major Depression precludes separate analysis of dysphoria and anhedonia as symptoms of depression. Third, the elderly are heterogeneous, varying widely in functional capacity, social resources, and cognitive status. For the purposes of this study, we have divided the samples into two groups based on an arbitrary cutpoint of 65 years of age, although in future analyses it might be desirable to consider persons 85 years and older separately. Even large samples such as these include relatively few of the "oldest old." Fourth, measures of physical illness and disability were not available for inclusion in the MIMIC model although these aspects of health and function might mediate the relationship between age and the tendency to endorse depressive symptoms. In addition, by limiting our analysis to household respondents who reported complete information on the covariates in the model, we have effectively excluded the most impaired respondents. Fifth, cross-sectional data on age-related phenomena cannot disentangle the extent to which age-related

differences in item performance reflect cohort and period effects rather than "aging" per se. Future analyses of multiwave data from the 13-year follow-up of the Baltimore ECA subjects will help to clarify the relationship of aging and the tendency to endorse symptoms.

The limitations of the study specific to the MIMIC model require additional comment. The first set of comments considers whether the relative size of measurement error is different for the young and old. While the model compares item endorsement at the same level of the latent trait as estimated by the model, the "true" level of depression may be either over- or underestimated for the elderly, relative to the difference in the estimated and "true" level of depression for younger persons. If the estimated level of the latent trait is *higher* than the "true" level of depression for older adults relative to younger adults, the model then *overestimates* the level of depression for the old. Older persons might be *more* likely to endorse symptoms that are etiologically unrelated to Major Depression but are related to physical illness (e.g., fatigue) or to age-related psychological changes (e.g., thoughts of death), thus "inflating" the estimated level of the latent trait in the aged. Inspection of the raw prevalences of the DIS items, however, shows that the elderly do not universally endorse somatic symptoms at a rate higher than younger persons. For example, in Baltimore, compared to younger persons, older persons report a similar or lower one-month prevalence of loss of appetite, sleep disturbance, feeling tired, and moving more slowly than usual. On the other hand, if the estimated level of the latent trait is *lower* than the "true" level of depression for older persons relative to younger persons, the model then *underestimates* the level of depression for the old. Older adults might be *less* likely to endorse symptoms in general, especially those of a "psychological" nature. In this instance, our estimate of the extent to which older adults underreport sadness or loss of interest will be conservative; that is, the "true" direct effect of age on the dysphoria/anhedonia item of the DIS will be even more negative than the estimated direct effect reported here.

The second set of comments on limitations specific to the MIMIC model considers other explanations for the findings of the study. First, dysphoric mood may simply be a less salient feature of depression among older adults; that is, the form of depression is different in the old than in the young (Newmann, Engel, & Jensen, 1991). Specifically, our findings are consistent with the notion that the dysphoria criterion as expressed in the DIS may not be a prominent feature of depression in older people. Dimensional models like the MIMIC will be a fruitful area for future research in delineating symptom profiles specific to the elderly because the MIMIC draws attention to age-related differences in item response while simultaneously adjusting for other important characteristics and the level of the latent trait. Second, the measurement component of the MIMIC model results in a single set of parameter estimates for the factor loadings at each site (the β's in Table 5) irrespective of age. Theoretically, if there are group differences in the factor loading for the dysphoria/anhedonia item, this could account for the direct effect observed. However, we point out that the factor loadings are estimated simultaneously with the regression of the latent trait on the covariates. For this reason, we can include several covariates thought to be influential, and avoid ascribing to age any differences in factor loadings which are actually due to personal characteristics such as gender and ethnicity. In addition, despite the skewed nature of the items on the measurement side of the model, the latent trait in the model is normally distributed *conditional* on the covariates. As a result, transformations of the variables to satisfy assumptions of normality for the factor analysis are not necessary (cf., Newmann et al., 1991). This is particularly important in analysis of epidemiologic data such as the ECA, in which the symptom data are dichotomous, skewed, and of low prevalence.

Notwithstanding the limitations of the study, the finding of differential item functioning by age in the symptom of dysphoria deserves attention. In two independently gathered samples, in Baltimore and in Durham-Piedmont, adults aged 65 years and older were found to be less likely to endorse the dysphoria/anhedonia item of the DIS, even accounting for the level of depression and other personal characteristics, such as gender and ethnicity, which might influence the reporting of symptoms. Gaitz and Scott (1972) pointed out the pitfalls involved in ascribing to age, differences in reporting of mental symptoms due to gender and ethnicity. The MIMIC model has not been previously applied to a psychometric study of age differences in item response. The MIMIC model sharpens the focus on the relationship of a given characteristic, such as age, to a particular measurement item, such as the dysphoria/anhedonia criterion of DIS Major Depression. This focus is achieved by removing the effects of other characteristics that might otherwise modify the prevalence of the item through their influence on the measurement trait of interest. This strength of the analysis encouraged us to examine the influence of age on the reporting of other symptoms which constitute the DIS section on Major Depression, separate from our primary hypothesis.

The finding that the elderly are less likely to endorse the dysphoria/anhedonia item of the DIS might explain, at least in part, the low prevalence of Major Depression found in the older age groups of the ECA (Blazer, Hughes, & George, 1987; Eaton et al., 1989; Kramer et al., 1985; Myers et al., 1984; Weissman et al., 1988). It would be important to confirm this finding using a latent trait MIMIC model formulated on the basis of symptom scales such as the General Health Questionnaire (Goldberg & Blackwell, 1970). Other investigators have suggested that dysphoric mood may be less salient for older persons with depression. Blazer and colleagues (1988), analyzing the Durham-Piedmont ECA data, found a symptom profile tending to occur in older persons in which cognitive impairment was prominent and accompanied by sleep and appetite disturbance, hopelessness, and thoughts of death. Newmann and coworkers (1990, 1991), studying women aged 51 to 92 years, used multigroup latent variable models to compare the structure of symptoms of depression measured by a symptom scale in two age groups. Their data were consistent with the "depletion syndrome" drawn from Fogel and Fretwell's (1985) observations coexisting with more "delimited" and transient symptom groupings, such as sleep disturbance. Arean and Miranda (1992) found that, after stratifying on total

score to achieve some adjustment for severity of depressive symptoms, older medical patients were less likely to endorse the item "I felt depressed" from the Center for Epidemiologic Studies Depression Scale (CES-D, Radloff, 1977). Adjusting for gender, educational status, and marital status, Craig and Van Natta (1979) also found that older adults were less likely than young adults to endorse depression in the CES-D. In contrast, Ben-Arie et al. (1987) in a South African community survey, Oxman and colleagues (1990) in primary care, and Koenig and coworkers (1993) in a study of aged male veterans on admission to the hospital, found that virtually all older persons with a diagnosis of depression exhibited dysphoric mood. Reported age differences in dysphoria and anhedonia symptoms are difficult to evaluate since dysphoria and anhedonia are so intertwined with the clinical criteria for depression. In any case, patients in clinical samples are probably more severely depressed than persons in samples drawn from the community, so that dysphoric mood is present regardless of age.

While our primary finding in Baltimore, and replicated in Durham-Piedmont, was that the dysphoria/anhedonia item of the DIS was *less likely* to be endorsed by older adults, it is important to take note of symptoms that were *more likely* to be endorsed, accounting for the level of depression and for covariates such as gender. In the one month prior to interview, older persons at both sites studied were more likely than younger persons at the same level of the latent trait to report sleep disturbance, to have thought a lot about death, and to have felt like they wanted to die, but were less likely to endorse weight gain as well as sadness or loss of interest or pleasure (Table 6). Differences by site in the estimated direct effects in Table 6 probably reflect site differences not included in the MIMIC model, such as the contrast between rural and urban communities. To illustrate that a direct age comparison of the raw prevalences can be misleading, consider the symptom, "thought a lot about death" in the Baltimore sample. In Baltimore, despite a *lower* one-month prevalence of "thoughts of death" in older persons compared to younger persons, when the level of depression and personal characteristics such as gender were taken into account, older individuals were found to be *more* likely to acknowledge thoughts of death.

Our findings with regard to age variation in the symptoms of depression in the ECA were similar though not wholly consistent with previous work. For example, Museti and coworkers (1989) found that weight loss, psychomotor retardation, and fatigue were more common in older persons presenting to a university center in Italy, while guilt and suicidal ideation were more common in the young. Blazer, Bachar, and Hughes (1987) suggested that, compared to younger inpatients, older inpatients with Major Depression exhibited weight loss more commonly and suicidal thoughts less commonly. Berry, Storandt, and Coyne (1984) found that sleep and appetite disturbance were more salient symptoms among older adults. Zemore and Eames (1979) reported that, compared to young adults enrolled in psychology courses, older adults more commonly assented to insomnia and fatigue and less commonly to self-blame. Adjusting for gender, education, and marital status, Craig and Van Natta (1979) reported that older adults in a community survey were less likely to endorse most symptoms from the CES-D, including trouble concentrating, poor appetite, and feeling depressed. Unlike the current investigation, these studies do not adjust for level of depressive symptoms, are limited in their accounting for other characteristics which might mediate the relationship of age and the reporting of symptoms, and are drawn for the most part from clinical or convenience samples.

The greater tendency of older persons than younger persons to endorse sleep disturbance and thoughts of death at the same level of depression is an important lead in efforts to improve screening for depression in older adults, particularly in general medical settings. Sleep disturbance has already been noted to be a clue to the presence of psychiatric disorder (Ford & Kamerow, 1989). Kales and Kales (1974) observed that elderly patients with insomnia thought to be secondary to depression frequently denied dysphoric mood. In addition, while thoughts of death might be a "normal" psychological development in old age, the high rate of suicide among older adults, in the light of a reluctance to endorse dysphoria or anhedonia, raises the question of whether thoughts of death and attitudes toward death should be routinely discussed with older persons. The design of screening instruments and criteria for depression in older adults should account for this age-related heterogeneity in item response along the continuum of depression.

A discussion of differences in the level of depression with regard to personal characteristics other than age is beyond the scope of this study; however, while levels of latent trait depression according to gender, educational attainment, employment status, and marital status were generally in the expected direction, it is of interest to note that the MIMIC model implied a significantly increased level of depression among minorities at the Baltimore site (but not Durham-Piedmont), in contrast to prior studies from the ECA employing prevalence data (e.g., Somervell, Leaf, Weissman, Blazer, & Bruce, 1989) and incidence data (Anthony & Petronis, 1991; Gallo, Royall, & Anthony, 1993). This observation serves to highlight differences when a dimensional approach is employed rather than conventional categorical criteria, and suggests an application of the MIMIC model to identify bias in measurement among cultural or ethnic groups.

Dimensional models make fewer assumptions than categorical models, are generally more parsimonious, do not impose the structure of categorical entities on the symptom pattern, and are consistent with a public health paradigm (Anderson et al., 1993; Blashfield, 1990; Duncan-Jones et al., 1986; Goldberg & Huxley, 1992; Mirowsky & Ross, 1989). For older adults, "depression" may encompass symptom markers that are more prominent than dysphoria, but not part of standard criteria, such as helplessness, hopelessness, real or perceived cognitive deficit, and anxiety. Latent trait models can help elucidate these patterns, but the MIMIC model highlights age differences in measurement by accounting for other characteristics such as gender, that may modify reporting of symptoms. Employing dimensional models to improve the measurement of depression among older adults will require precise measures of the *symptoms* that form the dimensions of depression, recovery of the

interrelationships of the dimensions in empirical data through the use of multivariate statistical techniques, and the association of the dimensions with pathology and prognostic power. The result will be a clearer definition and understanding of depression and its correlates in late life.

In summary, persons aged 65 years and older were less likely to endorse dysphoria or anhedonia compared to younger persons (but more likely to endorse sleep disturbance and thoughts of death), even at the same level of depression, and accounting for differences due to a number of personal characteristics, such as gender, thought to influence the reporting of depressive symptoms. Future studies using the latent trait model should be able to determine whether the model-derived linkages found in this study lead to useful improvements in screening and conceptualization of depression in the elderly.

ACKNOWLEDGEMENTS

This work was supported by National Institute of Drug Abuse grant DA03992, National Institute of Mental Health grant MH47447, National Institute of Mental Health training grant MH14592, and National Institute on Aging training grant AG00161. Data gathering was supported by the Epidemiologic Catchment Area (ECA) Program of the National Institute of Mental Health Division of Biometry and Epidemiology. The principal investigators and grant award numbers during data gathering were Jerome K. Myers at Yale University (MH34224), Morton Kramer at The Johns Hopkins University (MH33870), Lee N. Robins at Washington University (MH33883), Dan Blazer and Linda George at Duke University (MH35386), Richard Hough and Marvin Karno at the University of California, Los Angeles (MH35865). At the National Institute of Mental Health, principal collaborators during data gathering were Darrel A. Regier, Ben Z. Locke, William W. Eaton, Carl A. Taube, and Jack D. Burke, Jr. The authors gratefully acknowledge the assistance of Scott Hubbard in data analysis.

Address correspondence to Dr. Joseph J. Gallo, The Johns Hopkins University School of Hygiene and Public Health, 624 North Broadway, Baltimore, MD 21205. Internet: GALLO@JHUHYG.SPH.JHU.EDU

REFERENCES

Abramson, L.Y., Metalsky, G.I., & Alloy, L.B. (1989). Hopelessness depression: A theory based subtype of depression. *Psychological Review, 96,* 358–372.

Alexopoulos, G.S. (1990). Anxiety-depression syndromes in old age, editorial. *International Journal of Geriatric Psychiatry, 5,* 351–353.

Alwin, D.F. (1988). Structural equation models in research on human development and aging. In K.W. Schaie, R.T. Campbell, W. Meredith, & S.C. Rawlings (Eds.), *Methodological issues in aging research* (pp. 71–170). New York: Springer.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1993). *Practice guideline for major depressive disorder in adults.* Washington, DC: Author.

Anderson, J., Huppert, F., & Rose, G. (1993). Normality, deviance and minor psychiatric morbidity in the community. *Psychological Medicine, 23,* 475–485.

Anthony, J.C., & Aboraya, A. (1992). The epidemiology of selected mental disorders in later life. In J.E. Birren, R.B. Sloane, & G.D. Cohen (Eds.), *Handbook of aging and mental health* (pp. 27–73). New York: Academic Press.

Anthony, J.C., Folstein, M., Romanoski, A.J., Von Korff, M.R., Nestadt, G.R., Chahal, R., Merchant, A., Brown, H., Shapiro, S., Kramer, M., & Gruenberg, E.M. (1985). Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: Experience in eastern Baltimore. *Archives of General Psychiatry, 42,* 667–675.

Anthony, J.C., & Petronis, K.R. (1991). Suspected risk factors for depression among adults 18–44 years old. *Epidemiology, 2,* 123–132.

Arean, P.A., & Miranda, J. (1992, October). *Differential item functioning of CES-D items in older versus younger medical patients.* Paper presented at the Sixth Annual NIMH International Research Conference on Primary Care Mental Health Research: Concepts, Methods, and Obstacles. Bethesda, Maryland.

Bartholomew, D.J. (1987). *Latent variable models and factor analysis.* New York: Oxford University Press.

Ben-Arie, O., Swartz, L., & Dickman, B.J. (1987). Depression in the elderly living in the community: Its presentation and features. *British Journal of Psychiatry, 150,* 169–174.

Berry, J.M., Storandt, M., & Coyne, A. (1984). Age and sex differences in somatic complaints associated with depression. *Journal of Gerontology, 39,* 465–467.

Blashfield, R.K. (1990). Comorbidity and classification. In J.D. Maser & C.R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 61–82). Washington, DC: American Psychiatric Press.

Blazer, D., Bachar, J.R., & Hughes, D.C. (1987). Major depression with melancholia: A comparison of middle-aged and elderly adults. *Journal of the American Geriatrics Society, 35,* 927–932.

Blazer, D., Hughes, D.C., & Fowler, N. (1989). Anxiety as an outcome symptom of depression in elderly and middle-aged adults. *International Journal of Geriatric Psychiatry, 4,* 273–278.

Blazer, D.G., Hughes, D.C., & George, L.K. (1987). The epidemiology of depression in an elderly community population. *The Gerontologist, 27,* 281–287.

Blazer, D., Swartz, M., Woodbury, M., Manton, K.G., Hughes, D., & George, L.K. (1988). Depressive symptoms and depressive diagnoses in a community population: Use of a new procedure for analysis of psychiatric classification. *Archives of General Psychiatry, 45,* 1078–1084.

Bollen, K.A., & Long, J.S. (Eds.). (1993). *Testing structural equation models.* Newbury Park, CA: Sage Publications.

Cattell, R.B. (1965). Factor analysis: An introduction to essentials. *Biometrics, 21,* 190–215.

Costello, C.G. (1993a). The advantages of the symptoms approach to depression. In C.G. Costello (Ed.), *Symptoms of depression* (pp. 1–21). New York: John Wiley & Sons.

Costello, C.G. (1993b). Research on symptoms versus research on syndromes: Arguments in favour of allocating more research time to the study of symptoms. *British Journal of Psychiatry, 160,* 304–308.

Craig, T.J., & Van Natta, P.A. (1979). Influence of demographic characteristics on two measures of depressive symptoms: The relation of prevalence and persistence of symptoms with sex, age, education, and marital status. *Archives of General Psychiatry, 36,* 149–154.

Depression Guideline Panel. (1993). *Depression in primary care: Volume 1. Detection and diagnosis. Clinical practice guideline, Number 5.* Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, AHCPR Pub. No. 93-0550.

Depue, R.A., & Monroe, S.M. (1978). Learned helplessness in the perspective of the depressive disorders: Conceptual and definitional issues. *Abnormal Psychology, 87,* 3–20.

Duncan-Jones, P., Grayson, D.A., & Moran, P.A.P. (1986). The utility of latent trait models for psychiatric epidemiology. *Psychological Medicine, 16,* 391–405.

Eaton, W., & Kessler, L.G. (1985). *Epidemiologic field methods in psychiatry: The NIMH Epidemiologic Catchment Area program.* New York: Academic Press.

Eaton, W.W., Kramer, M., Anthony, J.C., Dryman, A., Shapiro, S., & Locke, B.Z. (1989). The incidence of specific DIS/DSM-III mental disorders: Data from the NIMH Epidemiologic Catchment Area Program. *Acta Psychiatrica Scandinavica, 79,* 163–178.

Edelbrock, C. (1987). Psychometric research on children and adolescents. In C.G. Last & M. Hersen (Eds.), *Issues in diagnostic research* (pp. 219–240). New York: Plenum Press.

Fogel, B.S., & Fretwell, M. (1985). Reclassification of depression in the medically ill elderly. *Journal of the American Geriatrics Society, 33,* 446–448.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Ford, D.E., & Kamerow, D.B. (1989). Epidemiologic study of sleep disturbances and psychiatric disorders: An opportunity for prevention? *Journal of the American Medical Association, 262,* 1479–1484.

Gaitz, C.M., & Scott, J. (1972). Age and the measurement of mental health. *Journal of Health and Social Behavior, 13,* 55–67.

Gallo, J.J., Royall, D.R., & Anthony, J.C. (1993). Risk factors for the

onset of depression in middle age and late life. *Social Psychiatry and Psychiatric Epidemiology, 28,* 101–108.

Gibbons, R.D., Clark, D.C., Cavanaugh, S.V., & Davis, J.M. (1985). Application of modern psychometric theory in psychiatric research. *Journal of Psychiatric Research, 19,* 43–55.

Goldberg, D.P., & Blackwell, B. (1970). Psychiatric illness in general practice: A detailed study using a new method of case identification. *British Medical Journal, 2,* 439–443.

Goldberg, D., & Huxley, P. (1992). *Common mental disorders: A biosocial model.* New York: Tavistock/Routledge.

Good, W.R., Vlachoniklolis, I., Griffiths, P., & Griffiths, R.A. (1987). The structure of depressive symptoms in the elderly. *British Journal of Psychiatry, 150,* 463–470.

Grayson, D.A., Henderson, A.S., & Kay, D.W.K. (1987). Diagnoses of dementia and depression: A latent trait analysis of their performance. *Psychological Medicine, 17,* 667–675.

Heithoff, K.A. (1992). *Depression in late-life: Analysis of data from the NIMH epidemiologic catchment area (ECA) program.* Unpublished doctoral dissertation, The Johns Hopkins University, Baltimore, Maryland.

Helzer, J.E., Robins, L.N., McEvoy, L.T., Spitznagel, E.L., Stolzman, R.K., Farmer, A., & Brockington, I.F. (1985). A comparison of clinical and diagnostic interview schedule diagnoses: Physician reexamination of lay-interviewed cases in the general population. *Archives of General Psychiatry, 42,* 657–666.

Hoelter, J.W. (1983). The analysis of covariance structures: Goodness of fit indices. *Sociological Methods and Research, 11,* 325–344.

Horn, J.L., & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18,* 117–144.

Horwath, E., Johnson, J., Klerman, G.L., & Weissman, M.M. (1992). Depressive symptoms as relative and attributable risk factors for first-onset major depression. *Archives of General Psychiatry, 49,* 817–823.

Kales, A., & Kales, J.D. (1974). Sleep disorders. *New England Journal of Medicine, 200,* 487–499.

Kessler, R.C., Foster, C., Webster, P.S., & House, J.S. (1992). The relationship between age and depressive symptoms in two national surveys. *Psychology and Aging, 7,* 119–126.

Klerman, G.L. (1989). Psychiatric diagnostic categories: Issues of validity and measurement. *Journal of Health and Social Behavior, 30,* 26–32.

Koenig, H.G., Cohen, H.J., Blazer, D.G., Ranga Rama Krishnan, K., & Sibert, T.E. (1993). Profile of depressive symptoms in younger and older medical inpatients with Major Depression. *Journal of the American Geriatrics Society, 41,* 1169–1176.

Kramer, M., German, P.S., Anthony, J.C., Von Korff, M., & Skinner, E.A. (1985). Patterns of mental disorders among the elderly residents of Eastern Baltimore. *Journal of the American Geriatrics Society, 33,* 236–245.

Lazarsfeld, P.F., & Henry, N.W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin.

Mirowsky, J., & Ross, C.E. (1989). Psychiatric diagnosis as reified measurement. *Journal of Health and Social Behavior, 30,* 11–25.

Museti, L., Perugi, G., Soriani, A., Rossi, V.M., Cassano, G.B., & Akiskal, H.S. (1989). Depression before and after age 65: A re-examination. *British Journal of Psychiatry, 155,* 330–336.

Muthén, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics, 22,* 43–65.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49,* 115–132.

Muthén, B.O. (1988a). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer & H. Braun (Eds.), *Test validity.* Hillsdale, NJ: Lawrence Erlbaum.

Muthén, B.O. (1988b). *LISCOMP: Analysis of linear structural equations with a comprehensive measurement model* (2nd ed.). User's guide. Mooresville, IN: Scientific Software.

Muthén, B.O. (1989a). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557–585.

Muthén, B.O. (1989b). Dichotomous factor analysis of symptom data. *Sociological Methods and Research, 18,* 19–65.

Muthén, B.O. (1993). Goodness of fit with categorical and other nonnormal variables. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 205–234). Newbury Park, CA: Sage Publications.

Muthén, B.O., Grant, B., & Hasin, D. (1993). The dimensionality of alcohol abuse and dependence: Factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. *Addiction, 88,* 1079–1090.

Muthén, B.O., Kao, C.F., & Burstein, L. (1991). Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *Journal of Educational Measurement, 28,* 1–22.

Myers, J.K., Weissman, M.M., Tischler, G.L., Holzer, C.E., Leaf, P.J., Orvaschel, H., Anthony, J.C., Boyd, J.H., Burke, J.D., Kramer, K., & Stoltzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities. *Archives of General Psychiatry, 41,* 959–967.

NIH Consensus Development Panel on Depression in Late Life. (1992). Diagnosis and treatment of depression in late life. *Journal of the American Medical Association, 268,* 1018–1024.

Newmann, J.P. (1989). Aging and depression. *Psychology and Aging, 4,* 150–165.

Newmann, J.P., Engel, R.J., & Jensen, J. (1990). Depressive symptom patterns among older women. *Psychology and Aging, 5,* 101–118.

Newmann, J.P., Engel, R.J., & Jensen, J. (1991). Age differences in depressive symptom experiences. *Journal of Gerontology: Psychological Sciences, 46,* P224–P235.

Oxman, T.E., Barrett, J.E., Barrett, J., & Gerber, P. (1990). Symptomatology of late-life minor depression among primary care patients. *Psychosomatics, 31,* 174–180.

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401.

Robins, L.N., Helzer, J.E., Croughan, J., & Ratcliff, K.S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Robins, L.N., & Regier, D.A. (1991). *Psychiatric disorders in America: The Epidemiologic Catchment Area study.* New York: Free Press.

Rohrbaugh, R.M., Siegal, A.P., & Giller, E.L. (1988). Irritability as a symptom of depression in the elderly. *Journal of the American Geriatrics Society, 36,* 736–738.

Romanoski, A.J., Nestadt, G., Folstein, M.F., Chahal, R., Merchant, A., Brown, C.H., Gruenberg, E.M., & McHugh, P.R. (1992). The epidemiology of psychiatrist-ascertained depression and DSM-III depressive disorders. *Psychological Medicine, 22,* 629–655.

Rose, G. (1993). Mental disorder and the strategies of prevention. *Psychological Medicine, 23,* 553–555.

Salzman, C., & Shader, R.I. (1978). Depression in the elderly: I. Relationship between depression, psychologic defense mechanisms and physical illness. *Journal of the American Geriatrics Society, 26,* 253–260.

Somervell, P.D., Leaf, P.J., Weissman, M.M., Blazer, D.G., & Bruce, M.L. (1989). The prevalence of Major Depression in black and white adults in five United States communities. *American Journal of Epidemiology, 130,* 725–735.

Swartz, M., Carroll, B., & Blazer, D. (1989). In response to "Psychiatric diagnosis as reified measurement." *Journal of Health and Social Behavior, 30,* 33–34.

Tweed, D.L., & George, L.K. (1989). A more balanced perspective on "Psychiatric diagnosis as reified measurement." *Journal of Health and Social Behavior, 30,* 35–37.

Von Korff, M., & Anthony, J.C. (1982). The NIMH Diagnostic Interview Schedule modified to record current mental status. *Journal of Affective Disorders, 4,* 365–371.

Weiss, B., Weisz, J.R., Politano, M., Carey, M., Nelson, W.M., & Finch, A.J. (1992). Relations among self-reported depressive symptoms in clinic-referred children versus adolescents. *Journal of Abnormal Psychology, 101,* 391–397.

Weiss, I.K., Nagel, C.L., & Aronson, M.K. (1986). Applicability of depression scales to the old old person. *Journal of the American Geriatrics Society, 34,* 215–218.

Weissman, M.M., Leaf, P.J., Tischler, G.L., Blazer, D.G., Karno, M., Bruce, M.L., & Florio, L.P. (1988). Affective disorders in five United States communities. *Psychological Medicine, 18,* 141–153.

Wells, K.B., Burnam, M.A., Rogers, W., Hays, R., & Camp, P. (1992). The course of depression in adult outpatients: Results from the Medical Outcomes Study. *Archives of General Psychiatry, 49,* 788–794.

Wheaton, B. (1988). Assessment of fit in overidentified models with latent

variables. In J.S. Long (Ed.), *Common problems / proper solutions: Avoiding error in quantitative research* (pp. 202–203). Newbury Park, CA: Sage Publications.

Zemore, R., & Eames, N. (1979). Psychic and somatic symptoms of depression among young adults, institutionalized aged and noninstitutionalized aged. *Journal of Gerontology, 34,* 716–722.

Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement, 26,* 55–66.

Received September 30, 1993
Accepted February 22, 1994

Appendix

Parameters of the MIMIC Model

As described in the text, the first covariate, x_1, takes the value of 1 for individuals age 65 years and older and 0 for individuals less than 65 years of age. In the MIMIC model, paths from an x covariate to an η capture group differences in the factor mean for η. Consider x_1 in Figure 1, which represents group membership according to age. From equation 2 (see page P254), the factor level for $x_1 = 1$ (age 65 years and older) is:

$$\eta_0 = \alpha_0 + \gamma_{0.1}x_1 + \zeta_0 \quad (3)$$

We set the value of $\beta_{7.}$ for TIRED, to 1.0 so that the model is identified. Again, applying equation 2, the factor level for $x_1 = 0$ (age less than 65 years) is:

$$\eta_0 = \alpha_0 + \zeta_0 \quad (4)$$

The difference of equations 3 and 4 then represents the difference in factor levels of η_0 for persons age 65 years and older compared to persons less than 65 years of age:

$$[\alpha_0 + \gamma_{0.1}x_1 + \zeta_0] - [\alpha_0 + \zeta_0] = \gamma_{0.1}x_1 = \gamma_{0.1} \quad (5)$$

This demonstrates that the coefficient for the path from x_1 to η_0, $\gamma_{0.1}$, represents the difference in the factor level of the latent trait, η_0, for persons aged 65 years and older compared to persons younger than age 65 years. This is analogous to the regression of a continuous outcome variable onto several dichotomous x covariates in multiple regression. The coefficients of the x's in multiple regression *also* represent the difference in the value of the outcome variable for one group ($x = 1$) compared to a reference group ($x = 0$).

The MIMIC model allows for estimation of differential item response through the use of "direct effects" from the x covariate to the item of interest, and was the focus of our hypothesis regarding the dysphoria / anhedonia item of the DIS and age. The coefficient for a direct effect accounts for variation in item endorsement due directly to membership in a given group, over and above that accounted for by differences in the level of the latent trait. The factor level for η_0, dysphoria or anhedonia, can be estimated from equation 2 for the case when $x_1 = 1$ (age 65 years and older):

$$\eta_1 = \alpha_1 + \beta_{1.0}\eta_0 + \gamma_{1.1}x_1 + \zeta_1 \quad (6)$$

From equation 2, the factor level for η_1 when $x_1 = 0$ (age less than 65 years) is:

$$\eta_1 = \alpha_1 + \beta_{1.0}\eta_0 + \zeta_1 \quad (7)$$

The difference of equations 6 and 7 then represents the difference in factor level of the item, 1, over and above variation due to differences in the factor level of the latent trait, 0, for persons age 65 years and older compared to persons less than 65 years of age:

$$[\alpha_1 + \beta_{1.0}\eta_0 + \gamma_{1.1}x_1 + \zeta_1] - [\alpha_1 + \beta_{1.0}\eta_0 + \zeta_1] = \gamma_{1.1}x_1 = \gamma_{1.1} \quad (8)$$

This demonstrates that the coefficient for the path from x_1 to η_1 (i.e., $\gamma_{1.1}$) represents the difference in endorsement of the item (i.e., "prevalence") for persons aged 65 years and older compared to persons younger than age 65 years, adjusted for age differences in the level of the latent trait η_0. Put another way, the direct effect in a MIMIC model captures group differences in item prevalence when individuals in different groups are at the same level of the latent trait. In this instance, the model relates to differences in endorsement for η_1, which represents dysphoria or anhedonia. We can examine differences in endorsement for other items in the criteria for Major Depression, such as sleep disturbance, in the same way.