An Analysis of Meehl's MAXCOV-HITMAX Procedure for the Case of Dichotomous Indicators

Michael D. Maraun, Kathleen Slaney, and Louis Goddyn
Simon Fraser University

MAXCOV-HITMAX was invented by Paul Meehl for the detection of latent taxonic structures (i.e., structures in which the latent variable, θ, is not continuously, but rather Bernoulli, distributed). It involves the examination of the shape of a certain conditional covariance function, and is based on Meehl's claims that: (R1) Given a latent taxonic structure, this conditional covariance function is single peaked; and that (R2), continuous latent structures produce a flat, rather than single-peaked, curve. While Meehl has recommended that continuous indicators be used as input into MAXCOV-HITMAX, the use of dichotomous indicators has become popular. The current work investigates whether (R1) and (R2) are true for the case of dichotomous indicators. The conclusions will be that, for dichotomous indicators: (a) (R1) is not true; (b) (R1) is made true given that there are a large number of indicators; and (c) (R2) is not true, certain unexceptional Rasch structures, for example, producing single-peaked curves. Implications are briefly discussed of these results for the case of MAXCOV-HITMAX with continuous indicators.

In a series of articles (Meehl, 1965, 1973, 1992; Meehl & Golden, 1982; Meehl & Yonce, 1996; Waller & Meehl, 1998), noted theoretician Paul Meehl developed what he calls *taxometrics*, a set of procedures which, he claims, can be used to detect latent taxa (i.e., discrete *types* which underlie, perhaps causally, responding to a set of indicator variables) when, in fact, they do exist. The most widely employed of these procedures is MAXCOV-HITMAX, which involves the examination of the shape of a certain conditional covariance function. In particular, MAXCOV-HITMAX is based on Meehl's claims that: (R1) Given a latent taxonic structure, this conditional covariance function is single peaked; and that (R2), continuous latent structures produce a flat, rather than single-peaked, curve. If this were the case, MAXCOV-HITMAX would be a very useful criterion of latent taxonicity, for it then could not only be used to judge when data are in keeping with the hypothesis of latent taxonicity, but also to rule out as a possibility continuous latent structures.

However, there has recently arisen controversy with regard MAXCOV-HITMAX. Miller (1996) provides a counterexample that appears to contradict (R2), suggesting that MAXCOV-HITMAX may signal taxonicity when the latent structure is, in fact, continuous. Questions have also been raised regarding the appropriateness of the practice of employing dichotomous indicators as input into MAXCOV-HITMAX. This controversy has arisen, to a certain extent, because, while he has offered up an impressive number of examples and Monte Carlo studies in support of MAXCOV-HITMAX, for neither of the cases of continuous, nor dichotomous, indicators has Meehl provided formal proofs of (R1) and (R2).

According to Meehl (Meehl & Yonce, 1996), the use of dichotomous indicators as input into MAXCOV was first suggested in Meehl (1965). As indicated by Meehl and Yonce (1996, p. 1112), their use has been quite popular in applied research. Meehl, however, has been careful to admit his uncertainty regarding the practice: "Despite the impressive results that have been obtained by investigators using dichotomous outputs, we retain a strong preference for quantitative output indicators until more adequate Monte Carlo tests have been done" (Meehl & Yonce, 1996, p. 1114); "The limitations of using dichotomous output indicators remain to be investigated" (1995, p. 272). Varied accounts have been given as to the perceived nature of the danger in using dichotomous indicators. Meehl and Yonce (1996, p. 1113) explain that the concern is "...about the danger of spurious results...", and speculate that "If we are dealing with items which all have the same difficulty level, very steep discrimination ogives are also needed to give spurious results." They do not provide further detail regarding the latent structures to which they refer, but conclude that such structures will, in practice, rarely arise. It appears then that the issue is that (R2) may, under certain rare scenarios, not be true.

Given its popularity in mainstream application, the status of the dichotomous indicator paradigm is in need of resolution, this being the aim of the work herein described. In particular, it is investigated whether, for the case of dichotomous indicators, (R1) and (R2) are true. The case of continuous indicators is briefly discussed in light of the findings. It will turn out that, for the case of dichotomous indicators:

1. A latent taxonic structure does not necessarily produce a single-peaked conditional covariance function. That is, (R1) is not true.

2. (R1) is true given that there are a large number of indicators. Hence, there exists a non-prohibitive, but less general, condition, under which (R1) is true.

3. It is not the case that continuous latent structures do not produce single-peaked conditional covariance functions. Certain unexceptional

Rasch structures, for example, can produce such curves. That is, for the case of dichotomous indicators, (R2) is not true.

The Logic of MAXCOV-HITMAX

Meehl derived MAXCOV-HITMAX on the basis of a characterization of latent taxonicity. Based on his many discussions of MAXCOV-HITMAX (e.g., Meehl, 1973, 1992), it may be deduced that this characterization involves three elements, here called M1, M2, and M3.

M1: Taxon and Complement Class

There exist two (latent) classes of individuals, one class called the taxon (T) and the other, the complement class (T'). This situation may be represented by defining θ, the latent variable, to be a random variate with Bernoulli distribution, such that

$$(1) \qquad 0 < P(\theta = T) = \pi_T < 1, \text{ and } P(\theta = T') = (1 - \pi_T),$$

a property that will, henceforth, be referred to as M1.

M2: Indicators

Define an indicator to be a dichotomous random variate (1 = endorsement; 0 = lack of endorsement), X_i, with the property that, after recoding,[1]

$$(2) \qquad P(X_i = 1|\theta = T) > P(X_i = 1|\theta = T'),$$

a property known as positive regression dependence (Lehmann, 1966; Tukey, 1958) and, within the domain of latent variable modelling (e.g., Holland & Rosenbaum, 1986), "latent monotonicity." Property Equation 2 is a defining feature of all commonly encountered latent variable models, and is, of course, equivalent to

$$(3) \qquad E(X_i|\theta = T) > E(X_i|\theta = T').$$

---

[1] Holland and Rosenbaum (1986, p. 1540) describe how the recoding should be done: One simply codes the X_i so that

$$C(X_i, \sum_{i=1} X_i)$$

is positive for all i. Such a recoding is not necessarily achievable, and, if it is not, **X** does not conform to *any* latent variable model in which θ only assumes two values.

Condition M2 then states that there exists a set of p indicators of the taxon, stored in the random vector $\mathbf{X}$.

M3: Conditional Independence

One interpretation given by Meehl to latent taxa is that they are causes of the responding of individuals to the indicators. He paraphrases this notion in the usual way: The association that exists amongst the indicators is completely "explained" by the existence of the latent taxon and complement class. This condition (henceforth, M3) is stated as

$$(4) \qquad P(\mathbf{X}_1 = x_1, \mathbf{X}_2 = x_2, ..., \mathbf{X}_p = x_p | \theta = t) = \prod_{i=1}^{p} P(\mathbf{X}_i = x_i \mid \theta = t)$$

that is, the joint discrete mass function of $\mathbf{X}$ given $\theta = t$ is equal to the product of the individual conditional discrete mass functions. It follows from Equation 4 that the two $p \times p$ conditional covariance matrices, $C(\mathbf{X}|\theta = T) = \Phi_T$ and $C(\mathbf{X}|\theta = T') = \Phi_{T'}$, are diagonal matrices. This diagonality condition, Meehl acknowledges, "is an idealization that will rarely be satisfied in MAXCOV-HITMAX applications" (Waller & Meehl, 1998, p. 17), but whose failure to obtain, he claims, "...only rarely vitiates MAXCOV-HITMAX parameter estimates" (Waller & Meehl, 1998, p. 17). In what follows, the case of a latent taxonic structure with dichotomous indicators will be symbolized as [M1∩M2∩M3].

Meehl has billed MAXCOV-HITMAX as a criterion of latent taxonicity. It is, according to Meehl, a quantitative basis for deciding when latent taxonicity underlies responding to a set of items. In a series of articles (e.g., Meehl, 1965, 1973, 1992; Meehl & Golden, 1982; Waller & Meehl, 1998), he has described the ingenious reasoning that led to MAXCOV-HITMAX, reasoning that may be paraphrased as follows:

1. Partition $\mathbf{X}$ as $[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}, \mathbf{X}^*]$, in which $\mathbf{X}_{1(i)}$ and $\mathbf{X}_{2(j)}$ are any two choices, $i \neq j$, from $(\mathbf{X}_1, ..., \mathbf{X}_p)$, and $\mathbf{X}^*$ contains the $(p - 2)$ remaining indicators;

2. Define the random variate $H = \mathbf{1}'\mathbf{X}^*$, that is, as the sum of the $(p - 2)$ indicators in $\mathbf{X}^*$;

3. Since H is the sum of $(p - 2)$ indicators, it too should be an indicator of T;

4. Define: $\boldsymbol{\pi}_{Th} = P(\theta = T|H = h)$; $\Phi_{Th} = C\{[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}]|H = h \cap \theta = T\}$ and $\Phi_{Th} = C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}]|H = h \cap \theta = T'\}$, each a 2×2 conditional covariance matrix of $\mathbf{X}_{1(i)}$ and $\mathbf{X}_{2(j)}$; $\boldsymbol{\mu}_{Th}$ a 2×1 vector with elements $E[\mathbf{X}_{1(i)}|H = h \cap \theta = T]$ and $E[\mathbf{X}_{2(j)}|H = h \cap \theta = T]$; and $\boldsymbol{\mu}_{Th}$ a 2×1 vector with elements $E[\mathbf{X}_{1(i)}|H = h \cap \theta = T']$ and $E[\mathbf{X}_{2(j)}|H = h \cap \theta = T']$. MAXCOV-HITMAX rests on a

consideration of the shape of the conditional covariance function, $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$. As is well known, the 2×2 covariance matrix of $\mathbf{X}_{1(i)}$ and $\mathbf{X}_{2(j)}$ conditional on $H = h$, can be expressed as:

(5) $\quad C\{[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}]|H = h\}$
$\quad\quad = \boldsymbol{\pi}_{Th} * \boldsymbol{\Phi}_{Th} + (1 - \boldsymbol{\pi}_{Th}) * \boldsymbol{\Phi}_{Th} + \boldsymbol{\pi}_{Th} * (1 - \boldsymbol{\pi}_{Th}) * (\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})'$

5. Meehl reasons that $\mathbf{X}_{1(i)}$ and $\mathbf{X}_{2(j)}$ should be statistically independent when conditioned on both $H = h$ and $\theta = t$, and, hence, that $\boldsymbol{\Phi}_{Th}$ and $\boldsymbol{\Phi}_{Th}$ should be diagonal;

6. He likewise reasons that $(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})$ should be constant over the range of H;

7. If his reasoning is correct, $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$, the off-diagonal element of $C\{[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}]|H = h\}$, will then vary with h only through $\boldsymbol{\pi}_{Th}(1 - \boldsymbol{\pi}_{Th})$;

8. Since $0 < \boldsymbol{\pi}_{Th} < 1$, $\boldsymbol{\pi}_{Th}(1 - \boldsymbol{\pi}_{Th})$ has a maximum at $\boldsymbol{\pi}_{Th} = .5$, and, if $\boldsymbol{\pi}_{Th}$ is increasing, will be a single peaked function of h. $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ will then also be a single peaked function of h. This is the fundamental claim on which MAXCOV-HITMAX is based: If $[M1 \cap M2 \cap M3]$, then $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ is a single peaked function of h.

9. Meehl also claims that $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ will, conversely, be flat over the range of H if θ is distributed continuously. In his words: "If the latent structure is not taxonic, the curve will be flat" (Meehl, 1992, p. 134); "In MAXCOV-HITMAX the factorial situation does not give a dish...but a flat graph" (Meehl, 1995, p. 272). Hence, according to Meehl, a single peaked conditional covariance function distinguishes the continuous from the taxonic latent structure.

If this were true, an investigator could make an inference about the shape of the curve $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$, and, if this inference suggested single peakedness, justifiably, but provisionally, conclude that the data arose from latent taxonicity. She could then estimate $\boldsymbol{\pi}_T$, and classify individuals into T and T'. What is of interest in the current treatment is whether MAXCOV-HITMAX is, in fact, a criterion of latent taxonicity. Meehl, it should be emphasized, downplays the importance of results based on any one of his procedures, instead laying "heavy emphasis upon approximate empirical agreement between numerical inferences made from different procedures and indicators..." (Meehl, 1992, p. 132). As he argues, "...the most persuasive evidence of theories concerning theoretical entities, or events and processes observable in principle but not observed because of spatial or temporal remoteness from the scientist, lay in finding *consistent numerical results via quantitatively disparate epistemic paths*" (Meehl, 1992, p. 33; emphasis in original). Hence, the researcher must look for consistency in the

results that accrue from the employment of distinct criteria derived on the basis of similar assumptions.

Clearly then, it would be a gross misportrayal of Meehl to have him suggesting that definitive conclusions can be reached regarding the existence of latent taxa on the basis of a single application of MAXCOV-HITMAX. But it is also true that convergence in the results of multiple individual procedures will be of little significance unless these individual procedures are, indeed, criteria of latent taxonicity. And since, for neither of the cases of continuous, nor dichotomous, indicators, have formal proofs, to date, been given of (R1) and (R2), the status of MAXCOV-HITMAX is, to a certain extent, open to question. The task, then, is to ascertain, for the case of dichotomous indicators, whether MAXCOV is a criterion of latent taxonicity. Before turning to this task, it is worth reviewing what is meant by the claim that a manifest property is a criterion for a particular latent structure.

Criteria of Latent Structures

The standard, and primary, justification for the claim that a manifest property, C, is a criterion for a particular latent structure, LS, is that it is true that LS $\Rightarrow C$ (i.e., it is true that "if LS, then C";C is a necessary condition for LS), or, equivalently, that it is true that $\sim C \Rightarrow \sim$LS. Such a property will, herein, be called a sense 1 criterion of LS. If C is a sense 1 criterion of LS, then, if the researcher finds that, for a particular set of data, $\sim C$ is the case, he is justified in taking this fact as evidence against the hypothesis that the data was generated by LS. For example, it is true that LS = (2-dimensional linear factor structure) $\Rightarrow C = (\Sigma_X = \Lambda_2 \Lambda_2' + \Psi)$, in which $\Sigma_X = C(X)$, Λ_2 is a $p \times 2$ matrix of "factor loadings", and Ψ a $p \times p$, diagonal, positive definite matrix. Hence, $\Sigma_X = \Lambda_2 \Lambda_2' + \Psi$ is a sense 1 criterion for the 2-dimensional linear factor structure. Generally speaking, if it were *not* the case that $\Sigma_X = \Lambda_t \Lambda_t' + \Psi$ is a sense 1 criterion for t-dimensional linear factor structures, it would be a pointless exercise to employ judgments of the conformity of data to such covariance structures to decide whether data might have arisen from such a structure.

Different latent structures may, of course, imply at least some of the same manifest properties. If a property is an implication of just one, or a narrow class, of latent structures, that is, it is true that $\sim$LS$\Rightarrow \sim C$ (or, equivalently, that it is true that $C \Rightarrow$ LS; C is sufficient for LS), it will, herein, be called a sense 2 criterion for LS. Note that $C = (\Sigma_X = \Lambda_2 \Lambda_2' + \Psi)$ is not a sense 2 criterion for LS = (2-dimensional linear factor structure) since it is not true that $\sim$LS = (2-dimensional linear factor) $\Rightarrow \sim C = (\Sigma_X = \Lambda_2 \Lambda_2' + \Psi)$.

The unidimensional, quadratic factor structures, for example, also imply that $\Sigma_{\mathbf{X}} = \mathbf{\Lambda}_2 \mathbf{\Lambda}_2' + \mathbf{\Psi}$ (McDonald, 1967). Hence, the manifest property $\Sigma_{\mathbf{X}} = \mathbf{\Lambda}_2 \mathbf{\Lambda}_2'$ $+ \mathbf{\Psi}$ does not constitute a basis for distinguishing between the 2-dimensional linear, and 1-dimensional quadratic, factor structures. The shape of the distribution of factor scores *does* constitute such a basis (McDonald, 1967; Maraun & Rossi, 2001), and, in particular, given that $\Sigma_{\mathbf{X}} = \mathbf{\Lambda}_2 \mathbf{\Lambda}_2' + \mathbf{\Psi}$, a quadratic distribution of factor score estimates eliminates as a candidate the 2-dimensional linear case. This is so because it is true that $C =$ (quadratic distribution of factor score estimates) $\Rightarrow$ ~LS = (2-dimensional linear factor). It should be acknowledged, however, that even if a property C is not a criterion (in either of senses 1 or 2) for a given LS, it may be so given a particular set of side-conditions (i.e., under special circumstances). The usefulness of a C that is a criterion for an LS, *only given a particular set of side-conditions*, will have to be judged on a case-by-case basis, the judgment dependent upon, among other factors, the likelihood that the side-conditions are realizable in practice.

Whether a manifest property is a criterion, in either of senses 1 or 2, for a particular latent structure is established via mathematical proof. There are no degrees of criterion-hood: A property either is or is not a criterion. This degree of certainty is, of course, not achievable when the researcher employs a particular criterion, C, in the making of decisions as to whether data square with the hypothesis of a particular latent structure. In the first place, he must judge whether, for a set of data, C does in fact obtain. Since the putative latent structure will never hold exactly, and, hence, C will never obtain exactly, the task becomes one of judging whether the data are approximately in keeping with C. Secondly, such decision making is usually inferential in nature. Hence, the researcher must decide, on the basis of a sample, whether C holds (approximately) in the population. Both approximation and sampling error render decisions made about the existence of latent structures via the employment of criteria, inherently tentative. Now, these issues are certainly important issues. However, their consideration is contingent upon a clear understanding of whether a property C is, in fact, a criterion of a latent structure LS, this being the issue of interest in the current work.

Criteria of Latent Taxonic Structures

The condition [M1∩M3] implies that

$$(6) \qquad P(\mathbf{X}=x) = \sum_{t=1}^{2} \boldsymbol{\pi}_t \left\{ \prod_{i=1}^{P} P(\mathbf{X}_i = 1 | \theta = t)^{x_i} \left[1 - P(\mathbf{X}_i = 1 | \theta = t) \right]^{1-x_i} \right\}.$$

Since the discrete mass function of Equation 6 is that of a latent class model (see, e.g., Bartholomew & Knott, 1999; Heinen, 1996), [M1∩M2∩M3] is a latent class structure. Consider, in the first place, several sense 1 criteria of [M1∩M2∩M3]. The condition [M1∩M3] implies that the covariance structure of the indicators has the form

$$(7) \quad C(\mathbf{X}) = \boldsymbol{\pi}_T * \boldsymbol{\Phi}_T + (1 - \boldsymbol{\pi}_T) * \boldsymbol{\Phi}_{T'} + \boldsymbol{\pi}_T * (1 - \boldsymbol{\pi}_T) * (\boldsymbol{\mu}_T - \boldsymbol{\mu}_{T'})(\boldsymbol{\mu}_T - \boldsymbol{\mu}_{T'})'$$

in which $\boldsymbol{\mu}_T = E(\mathbf{X}|\theta = T)$ and $\boldsymbol{\mu}_{T'} = E(\mathbf{X}|\theta = T')$. It thus follows that

$$(8) \quad C(\mathbf{X}) = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi},$$

in which $\boldsymbol{\Lambda} = \sqrt{\boldsymbol{\pi}_T (1 - \boldsymbol{\pi}_T)} * (\boldsymbol{\mu}_{T'} - \boldsymbol{\mu}_T)$ and $\boldsymbol{\Psi} = \boldsymbol{\pi}_T * \boldsymbol{\Phi}_T + (1 - \boldsymbol{\pi}_T) * \boldsymbol{\Phi}_{T'}$ is a diagonal matrix (McDonald, 1967; Bartholomew & Knott, 1999; Molenaar & von Eye, 1994). Since, from Equation 7, it follows that

$$(9) \quad \sigma_{ij} = \boldsymbol{\pi}_T(1 - \boldsymbol{\pi}_T)(\boldsymbol{\mu}_{Ti} - \boldsymbol{\mu}_{T'i})(\boldsymbol{\mu}_{Tj} - \boldsymbol{\mu}_{T'j}),$$

[M1∩ M2∩M3] implies that

$$(10) \quad \sigma_{ij} > 0,$$

or, equivalently, that the elements of $\boldsymbol{\Lambda}$ in the decomposition (Equation 8) are of the same sign.

The structure [M1∩M2∩M3] is, additionally, a member of the class of unidimensional, monotone, latent variable (UMLV) structures (Holland & Rosenbaum, 1986, p. 1529).[2] The indicators of a UMLV structure exhibit a range of strong, testable manifest dependencies, many of which have been described by Holland and Rosenbaum (1986). These authors proved that, for dichotomous indicators, UMLV structures imply that the indicators are associated (A), conditionally strongly positively orthant dependent (CSPOD), conditionally associated (CA), and conditionally multivariate positive of order 2 (CMTP$_2$). For continuous indicators, these are progressively stronger forms of multivariate dependency, while, for dichotomous indicators, they are equivalent forms (in the sense of entailment). Hence, A, CSPOD, CA, and CMTP$_2$ are sense 1 criteria of [M1∩M2∩M3]. Property Equation 10 is, in fact, a special case of the associatedness that is a property of the indicators of any UMLV structure. Holland and Rosenbaum (1986) and Rosenbaum (1984) have employed

---

[2] In the terminology of Stout (1990) and Junker (1993), the class of strictly unidimensional models.

these sense 1 criteria to derive tests of whether a set of dichotomous items conforms to *any* UMLV structure.

None of the sense 1 criteria listed above are, additionally, sense 2 criteria for [M1∩M2∩M3]. The covariance structure (Equation 8), for example, is also implied by the unidimensional, linear factor structure, while A, CA, CSPOD, and CMTP$_2$ are implied by UMLV structures in general.[3] These criteria cannot then be used to distinguish between the case in which θ has a continuous distribution and that in which it has a Bernoulli distribution (i.e., is taxonic).[4] McDonald (1967, pp. 58-64) provided a comprehensive discussion of the issue of criteria for latent class structures, and derived on the basis of nonlinear factor theory what, under certain special conditions, is a sense 2 criterion. His reasoning may be paraphrased as follows. From Equation 8, it follows that **X** may be given a linear factor representation,

$$(11) \qquad\qquad \mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\theta + \mathbf{e},$$

in which $E(\theta) = 0$, $E(\theta^2) = 1$, $E(\mathbf{X}|\theta = t) = \boldsymbol{\mu} + \boldsymbol{\Lambda}t$, $E(\mathbf{e}) = \mathbf{0}$, and $C(\mathbf{e}) = \boldsymbol{\Psi}$, a diagonal $p \times p$ matrix. The latent variate, θ, has, of course, a two-point distribution, and, since $E(\theta) = 0$ and $E(\theta^2) = 1$, takes the values

$$T = -\sqrt{\frac{(1-\boldsymbol{\pi})}{\boldsymbol{\pi}}}$$

and

$$T' = -\sqrt{\frac{\boldsymbol{\pi}}{(1-\boldsymbol{\pi})}}$$

with probabilities $\boldsymbol{\pi}$ and $(1 - \boldsymbol{\pi})$, respectively (McDonald, 1967). The matrix $C(\mathbf{X}) - \boldsymbol{\Psi}$ is of rank one, and may be represented as

$$(12) \qquad\qquad C(\mathbf{X}) - \boldsymbol{\Psi} = \mathbf{mm}'\omega = \boldsymbol{\Lambda}\boldsymbol{\Lambda}',$$

---

[3] Molennar and von Eye (1994) establish, additionally, the indistinguishability of the *k*-class latent profile model for continuous indicators, and the *k*-factor linear factor model, at both the levels of covariance and second-order moment structure.

[4] In fact, in discussing covariance structure based methods, McDonald (1967, p. 58) notes that "Existing methods of latent class analysis do not test whether the latent variates have a discrete distribution..."

in which $\mathbf{m}'\mathbf{m} = 1$, $\mathbf{m}$ and ω are, respectively, the single eigenvector and eigenvalue of $C(\mathbf{X}) - \boldsymbol{\Psi}$, and in which $\boldsymbol{\Lambda} = \sqrt{\omega}\,\mathbf{m}$. From Equations 11 and 12,

$$(13) \quad v = \frac{1}{\sqrt{\omega}}\mathbf{m}'(\mathbf{X} - \boldsymbol{\mu}) = \frac{1}{\sqrt{\omega}}\mathbf{m}'\boldsymbol{\Lambda}\theta + \frac{1}{\sqrt{\omega}}\mathbf{m}'\mathbf{e} = \frac{1}{\omega}\boldsymbol{\Lambda}'\boldsymbol{\Lambda}\theta + \frac{1}{\omega}\boldsymbol{\Lambda}'\mathbf{e} = w + d,$$

in which v is a "component variate" (McDonald, 1967, pp. 31-32). Note that w has the same shape of distribution as θ, and, if σ_d^2 is small, will have roughly the same distribution as v. Hence, if σ_d^2 is small, the taxonic and continuous latent scenarios will be distinguishable on the basis of the shape of the distribution of v. Now, since $\omega = \boldsymbol{\Lambda}'\boldsymbol{\Lambda}$,

$$(14) \qquad\qquad \sigma_d^2 = \frac{1}{\omega^2}\boldsymbol{\Lambda}'\boldsymbol{\Psi}\boldsymbol{\Lambda} = \frac{\boldsymbol{\Lambda}'\boldsymbol{\Psi}\boldsymbol{\Lambda}}{[\boldsymbol{\Lambda}'\boldsymbol{\Lambda}]^2}.$$

Let $(\boldsymbol{\mu}_{T'i} - \boldsymbol{\mu}_{Ti}) = \delta_i^2$, and recall that $\boldsymbol{\Lambda} = \sqrt{\boldsymbol{\pi}_T(1 - \boldsymbol{\pi}_T)} * (\boldsymbol{\mu}_T - \boldsymbol{\mu}_{T'})$. Since $\sigma_{Xi}^2 = \boldsymbol{\mu}_i(1 - \boldsymbol{\mu}_i)$, the ith element of $\boldsymbol{\Psi}$ is equal to

$$(15) \qquad\qquad \boldsymbol{\mu}_i(1 - \boldsymbol{\mu}_i) - \boldsymbol{\pi}(1 - \boldsymbol{\pi})\delta_i^2.$$

On the basis of Equations 14 and 15, McDonald established that

$$(16) \qquad \sigma_d^2 = \frac{\displaystyle\sum_{i=1}^{p}\delta_i^2\left[\boldsymbol{\mu}_i(1 - \boldsymbol{\mu}_i) - \boldsymbol{\pi}(1 - \boldsymbol{\pi})\delta_i^2\right]}{\boldsymbol{\pi}(1 - \boldsymbol{\pi})\left(\displaystyle\sum_{i=1}^{p}\delta_i^2\right)^2}$$

and that, for the special case in which, for all i, $\boldsymbol{\mu}_i = \boldsymbol{\pi}$ and $\delta_i^2 = \delta^2$,

$$(17) \qquad\qquad \sigma_d^2 = \frac{1 - \delta^2}{p\delta^2}.$$

For this special case, as $\delta^2 \to \infty$, $\sigma_d^2 \to 1/p$, and as $p \to \infty$, $\sigma_d^2 \to 0$. That is, for a set of very good indicators (those for which δ^2 is large), or, alternatively, a large number of indicators, the shapes of the distributions of v and θ will be very similar, and latent taxonicity will be detectable in the shape of the distribution of the component variate v.

MAXCOV-HITMAX, was developed by Meehl, along with a number of other procedures, explicitly for the purpose of allowing the researcher to distinguish between taxonic and continuous latent structures. MAXCOV-HITMAX is based on (R1) and (R2), which may now be stated formally as:

R1) It is true that: $[M1 \cap M2 \cap M3] \Rightarrow$ For any partition $[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}, \mathbf{X}^*]$ of $\mathbf{X}$, $\{C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ is a single peaked function of $h\}$;

R2) It is true that: (θ is distributed continuously) $\Rightarrow$ For any partition $[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}, \mathbf{X}^*]$ of $\mathbf{X}$, $\{C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ is not a single peaked function of $h\}$.

The claim (R1) is equivalent to asserting that {single-peakedness of $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]\}$ is a sense 1 criterion of latent taxonicity, while (R2), since it is equivalent to the claim that $\{C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ is a single peaked function of $h\} \Rightarrow \sim(\theta$ is distributed continuously), is *nearly* equivalent to the claim that the single-peakedness of $C[\mathbf{X}_{1(i)}, \mathbf{X}_{2(j)}|H = h]$ is a sense 2 criterion of latent taxonicity. It is not quite so because $\sim(\theta$ is distributed continuously) is not synonymous with $[M1 \cap M2 \cap M3]$. Nevertheless, if (R1) and (R2) were true, MAXCOV-HITMAX would be a very useful procedure, since it would constitute not only a logical basis for deciding upon whether a set of data was in keeping with the hypothesis of latent taxonicity, but also for deciding upon when continuous latent structures should be ruled out as candidates. This is clearly what Meehl has in mind, for he states that "For a taxometric procedure to "work" it must detect real taxa and not concoct nonexistent ones..." (1992, p. 138); "...by its overall shape, a MAXCOV-HITMAX plot indicates whether the data measure a latent taxon (type, class) or a latent dimension (continuum, factor)" (Waller & Meehl, 1998). Meehl was, of course, speaking with respect to the case of continuous indicators. Attention will now be turned to the question of whether (R1) and (R2) are true for the case of dichotomous indicators.[5]

Is R1 True?

From Equation 5, it is clear that, for (R1) to be true, it must be true that, for any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, $[M1 \cap M2 \cap M3]$ implies that:

(18) $(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})$ is constant over the range of H;

(19) $\boldsymbol{\Phi}_{Th}$ and $\boldsymbol{\Phi}_{T'h}$ are diagonal;

---

[5] In the proofs that follow, the more precise notations $\mathbf{X}_{1(i)}$ and $\mathbf{X}_{2(j)}$ are abandoned in favour of the more compact $\mathbf{X}_1$ and $\mathbf{X}_2$.

M. Maraun, K. Slaney, and L. Goddyn

(20) $\boldsymbol{\pi}_{Th} = P(\theta = T|H = h)$ is an increasing function of h;

(21) $\boldsymbol{\pi}_{Th}$ crosses .5.

In addition, a trivial requirement for the single-peakedness of $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is that the range of H contains at least three values. This requirement is met by ensuring that $\mathbf{X}^*$ contains at least two indicators, a condition assumed throughout. For any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, observe that $H = 1'\mathbf{X}^*$ has a discrete distribution on $[0, (p - 2)]$.

Theorem 1

For any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, [M1∩M2∩M3] implies Property 18, that is, that $(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})$ is constant over h.

Proof

The vth element of the 2×1 vector $(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})$ is equal to $E(\mathbf{X}_v|H = h \cap \theta = T) - E(\mathbf{X}_v|H = h \cap \theta = T')$. In the case of dichotomous indicators, this is equivalent to

(22) $$\frac{P(\mathbf{X}_v = 1 \cap H = h|\theta = T)\boldsymbol{\pi}_T}{P(H = h \cap \theta = T)} - \frac{P(\mathbf{X}_v = 1 \cap H = h|\theta = T')(1 - \boldsymbol{\pi}_T)}{P(H = h \cap \theta = T')}$$

which, from M3, is equivalent to

(23) $$\frac{P(\mathbf{X}_v = 1|\theta = T)P(H = h|\theta = T)\boldsymbol{\pi}_T}{P(H = h \cap \theta = T)} - \frac{P(\mathbf{X}_v = 1|\theta = T')P(H = h|\theta = T')(1 - \boldsymbol{\pi}_T)}{P(H = h \cap \theta = T')}$$

which, in turn, is equivalent to $P(\mathbf{X}_v = 1|\theta = T) - P(\mathbf{X}_v = 1|\theta = T')$. Hence, $(\boldsymbol{\mu}_{Th} - \boldsymbol{\mu}_{T'h})$ is constant over h. ∎

Theorem 2

For any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, [M1∩M2∩M3] implies Property 19, that is, that $\boldsymbol{\Phi}_{Th}$ and $\boldsymbol{\Phi}_{T'h}$, $h = 0..(p - 2)$, are diagonal.

Proof

The off-diagonal element of $\mathbf{\Phi}_{Th}$ is $C(\mathbf{X}_1, \mathbf{X}_2|H = h \cap \theta = t)$ which, for the case of dichotomous indicators, is equal to

(24) $P(\mathbf{X}_1 = 1 \cap \mathbf{X}_2 = 1|H = h \cap \theta = t) - P(\mathbf{X}_1 = 1|H = h \cap \theta = t)P(\mathbf{X}_2 = 1|H = h \cap \theta = t)$.

It was already established that the right member of Equation 24 is equal to $P(\mathbf{X}_1 = 1|\theta = t)P(\mathbf{X}_2 = 1|\theta = t)$. Now, the left member, $P(\mathbf{X}_1 = 1 \cap \mathbf{X}_2 = 1|H = h \cap \theta = t)$ is equal to

$$(25)\quad \frac{P\left(\mathbf{X}_1=1\cap \mathbf{X}_2=1\cap H=h\cap\theta=t\right)}{P\left(H=h\cap\theta=t\right)}=\frac{P\left(\mathbf{X}_1=1\cap \mathbf{X}_2=1\cap H=h|\theta=t\right)P\left(\theta=t\right)}{P\left(H=h\cap\theta=t\right)}$$

which, from M3, is equal to

$$(26)\quad \frac{P\left(\mathbf{X}_1=1|\theta=t\right)P\left(\mathbf{X}_2=1|\theta=t\right)P\left(H=h|\theta=t\right)P\left(\theta=t\right)}{P\left(H=h\cap\theta=t\right)}$$

which, finally, is equal to $P(\mathbf{X}_1 = 1|\theta = t)P(\mathbf{X}_2 = 1|\theta = t)$. Hence, $C(\mathbf{X}_1, \mathbf{X}_2|H = h \cap \theta = t) = 0$, for $h = 0..(p - 2)$, and the theorem is proven. ∎

To establish Property 20 it must be shown that $P(\theta = T|H = k) > P(\theta = T|H = m)$ for $0 \leq m < k \leq (p - 2)$. Note that

$$(27)\quad P\left(\theta=T|H=h\right)=\frac{P\left(H=h|\theta=T\right)\mathbf{\pi}_T}{P\left(H=h\right)},$$

and, hence, what must be shown is that

$$(28)\quad \frac{P\left(H=k|\theta=T\right)\mathbf{\pi}_T}{P\left(H=k\right)}>\frac{P\left(H=m|\theta=T\right)\mathbf{\pi}_T}{P\left(H=m\right)}.$$

However, $P(H = h) = \mathbf{\pi}_T * P(H = h|\theta = T) + (1 - \mathbf{\pi}_T) * P(H = h|\theta = T')$, hence, Inequality 28 is equivalent to

$$(29)\quad P\left(H=k|\theta=T\right)P\left(H=m|\theta=T'\right)>P\left(H=m|\theta=T\right)P\left(H=k|\theta=T'\right).$$

This is the requirement that H and θ be monotone likelihood ratio dependent (mlrd).

Monotone likelihood ratio dependence is discussed by Lehmann (1966), Karlin and Rinott (1980), and, within the context of item response models, Holland and Rosenbaum (1986) (under the heading of TP_2 distribution), Grayson (1988), Huynh (1994), and Hemker, Sijtsma, Molenaar, and Junker (1997). In the course of their investigations of the 2-point classification problem in item response theory, Grayson (1988) and Huynh (1994) proved, for the case of dichotomous indicators and continuous θ, that the sum of the indicators (total score) and θ are mlrd, Grayson's proof , unlike Huynh's, requiring that the derivatives of all item characteristic curves be non-negative over the range, and positive for at least one value, of θ. These proofs, however, do not establish that H and θ are mlrd for the case of dichotomous indicators and θ with a 2-point distribution. Hence, they do not establish Inequality 29. This special case is proven in Theorem 3.

Theorem 3

For any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, [M1∩M2∩M3] implies Property 20, that is, that $P(\theta = T|H = h)$ is an increasing function of h.

Proof

The proof is given in the Appendix.

What then remains is the question as to whether $P(\theta = T|H = h)$ necessarily crosses .5. If it does not, $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ will not be a single-peaked function of h. In fact, it will be shown that $P(\theta = T|H = h)$ does not necessarily cross .5, and, hence, $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is not necessarily a single-peaked function of h. In the following theorem, conditions are given under which $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is, and is not, a single-peaked function of h. Define

$$k_1 = \left[\prod_{i=1}^{(p-2)} \frac{P\left(\mathbf{X}_i^* = 1|\theta = T\right)}{P\left(\mathbf{X}_i^* = 1|\theta = T'\right)} \right]^{\frac{1}{(p-2)}}$$

the geometric mean of the ratios

$$\frac{P\left(\mathbf{X}_i^* = 1|\theta = T\right)}{P\left(\mathbf{X}_i^* = 1|\theta = T'\right)}$$

and

$$k_2 = \left\{ \prod_{i=1}^{(p-2)} \frac{\left[1 - P\left(\mathbf{X}_i^* = 1 | \theta = T\right)\right]}{\left[1 - P\left(\mathbf{X}_i^* = 1 | \theta = T'\right)\right]} \right\}^{\frac{1}{(p-2)}}$$

the geometric mean of the ratios

$$\frac{\left[1 - P\left(\mathbf{X}_i^* = 1 | \theta = T\right)\right]}{\left[1 - P\left(\mathbf{X}_i^* = 1 | \theta = T'\right)\right]}.$$

Theorem 4

For any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, if $[\text{M1} \cap \text{M2} \cap \text{M3}]$ then:

1. $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is an increasing function of h if

$$k_1 \leq \left[\frac{\left(1 - \pi_T\right)}{\pi_T} \right]^{\frac{1}{(p-2)}}.$$

2. $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is a decreasing function of h if

$$k_2 \geq \left[\frac{\left(1 - \pi_T\right)}{\pi_T} \right]^{\frac{1}{(p-2)}}.$$

3. $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is a single-peaked function of h if

$$k_2 < \left[\frac{\left(1 - \pi_T\right)}{\pi_T} \right]^{\frac{1}{(p-2)}} < k_1.$$

4. $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is a single-peaked function of h if $\pi_T = .5$.

Proof

$P(\theta = T | H = h)$ is an increasing function of h. If its range is $(0, .5)$, $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is increasing in h. If its range is $(.5, 1)$, $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is decreasing in h.

M. Maraun, K. Slaney, and L. Goddyn

1. Because $P(\theta = T|H = h)$ is increasing in h, its range is $(0, .5)$ if and only if

(30) $$P\big[\theta = T|H = (p-2)\big] = \frac{P\big[H = (p-2)|\theta = T\big]\pi_T}{P\big[H = (p-2)\big]} \leq .5,$$

Since $P[H = (p - 2)] = \pi_T * P[H = (p - 2)|\theta = T] + (1 - \pi_T) * P[H = (p - 2)|\theta = T']$, Equation 30 is equivalent to

(31) $$\frac{P\big[H = (p-2)|\theta = T\big]}{P\big[H = (p-2)|\theta = T'\big]} \leq \frac{(1-\pi_T)}{\pi_T}$$

But the left member of the inequality is equal to

$$\prod_{i=1}^{(p-2)} \frac{P\big(\mathbf{X}_i^* = 1|\theta = T\big)}{P\big(\mathbf{X}_i^* = 1|\theta = T'\big)} = k_1^{(p-2)}$$
∎

2. Because $P(\theta = T|H = h)$ is increasing in h, its range is $(.5,1)$ if and only if

(32) $$P\big(\theta = T|H = 0\big) = \frac{P\big(H = 0|\theta = T\big)\pi_T}{P\big(H = 0\big)} \geq .5$$

Following along the same path as in the proof of (1), Equation 32 is equivalent to

(33) $$\frac{P\big(H = 0|\theta = T\big)}{P\big(H = 0|\theta = T'\big)} \geq \frac{(1-\pi_T)}{\pi_T}$$

The left member of the inequality is equal to

$$\prod_{i=1}^{(p-2)} \frac{\big[1 - P\big(\mathbf{X}_i^* = 1|\theta = T\big)\big]}{\big[1 - P\big(\mathbf{X}_i^* = 1|\theta = T'\big)\big]} = k_2^{(p-2)}$$
∎

3. Point 3 is a simple consequence of points 1 and 2. If the range of $P(\theta = T|H = h)$, an increasing function, is not $(.5, 1)$, and also is not $(.5, 1)$,

then $P(\theta = T | H = h)$ crosses .5, and $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is a single-peaked function of h. ∎

4. From M2, $k_1 > 1$ and $k_2 < 1$. If $\pi_T = .5$,

$$\left[\frac{(1-\pi_T)}{\pi_T} \right]^{\frac{1}{(p-2)}} = 1.$$

Hence, the inequality of point 3 is satisfied. ∎

Theorem 4 establishes that (R1) is not true. Single-peakedness arises when the inequalities

$$k_2 < \left[\frac{(1-\pi_T)}{\pi_T} \right]^{\frac{1}{(p-2)}} < k_1$$

are satisfied, as when, for example, $\pi_T = .5$. Whether or not these inequalities are satisfied is a function of π_T, $(p-2)$, and the quality of the indicators, that is, the magnitudes of the ratios

$$\alpha_i = \frac{P\left(\mathbf{X}_i^* = 1 | \theta = T\right)}{P\left(\mathbf{X}_i^* = 1 | \theta = T'\right)}.$$

In particular, points 1 and 2 establish that, if T has a low (high) base rate, that is, π_T is small (large), $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ will be an increasing (decreasing) function of h unless the indicators in $\mathbf{X}^*$ are "very good." Hence, theorem 4 proves, for the case of dichotomous indicators, an observation of Meehl's: "Note that for extreme base rates- say, when the base rate is smaller than .10- the conditional slope (derivative) of a (possibly smoothed) MAXCOV function can fail to change sign" (Waller & Meehl, 1998, p. 22). It is also worth noting that the indicators the quality of which are important to the success of MAXCOV-HITMAX are those in the set $\mathbf{X}^*$, the "conditioning set", and not the indicators $\mathbf{X}_1$ and $\mathbf{X}_2$. Let, for all i, $\alpha_i = \alpha$, in which case $k_1 = \alpha$. Figure 1 then depicts the behaviour of

$$\left[\frac{(1-\pi_T)}{\pi_T} \right]^{\frac{1}{z}},$$

$z = (p - 2) > 2$, in relation to k_1 and k_2, given three different values of α (1.1, 1.5, and 2.0) and four different values of π_T (.05, .2, .8, and .95).

Side-Conditions Under Which R1 is True

It has been shown that, for the case of dichotomous indicators, (R1) is not true. That is, it is not true that [M1∩M2∩M3] → [$C(X_1, X_2|H = h)$ is a single peaked function of h]. However, under conditions of moderate base rate and/or excellent indicators in the conditioning set, (R1) is true. Meehl (1992; Meehl & Yonce, 1996) has, in fact, argued that the base rates of certain of the putative taxa that he has investigated may not be particularly extreme. Meehl's opinion notwithstanding, since the researcher cannot know in advance what the base rates are of the taxa he is to encounter, nor, for that matter, whether his indicators are of good quality, it would be of use to find other reasonable side-conditions (special circumstances) under which (R1) is true. The following theorem provides one such a side-condition.

Theorem 5

If [M1∩M2∩M3] then, for any partition [$\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}^*$] of $\mathbf{X}$, as $(p - 2)$ becomes large, $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ converges to a single-peaked function of h.

Proof

k_1 and k_2, being, as they are, geometric means of the ratios of conditional probabilities, are not functions of $(p - 2)$. From M2, $k_1 > 1$ and $k_2 < 1$. The function

$$\left[\frac{(1 - \pi_T)}{\pi_T} \right]^{\frac{1}{(p-2)}}$$

converges to unity as $(p - 2) \to \infty$. Hence, as $(p - 2)$ becomes large, the pair of inequalities

$$k_2 < \left[\frac{(1 - \pi_T)}{\pi_T} \right]^{\frac{1}{(p-2)}} < k_1$$
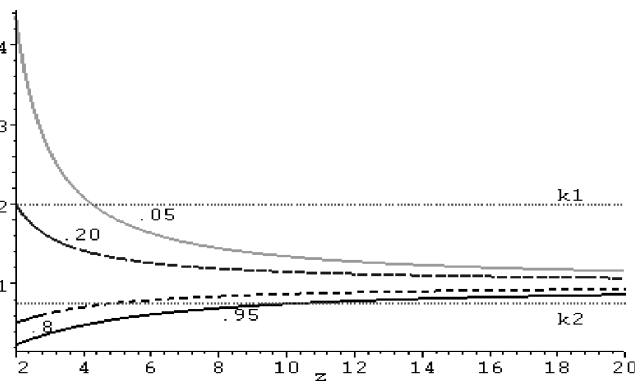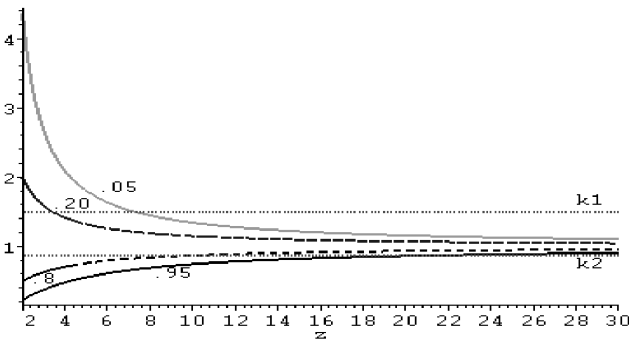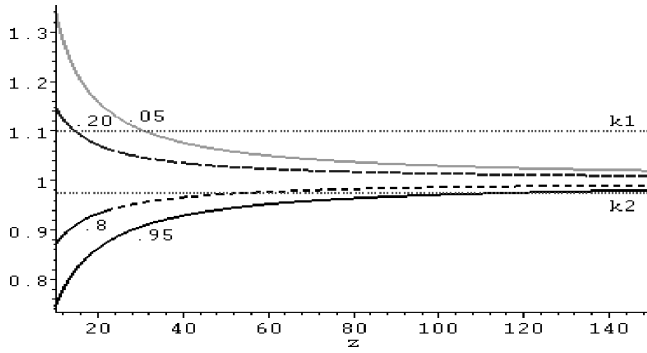
will eventually be satisfied. ∎

Figure 1

Behavior of $\left[\dfrac{\left(1-\pi_T\right)}{\pi_T}\right]^{\frac{1}{z}}$ for Selected Values of α and π_T

Theorem 5 suggests that, regardless of the quality of the indicators in $\mathbf{X}^*$, as $(p - 2)$, the number of indicators in $\mathbf{X}^*$, becomes large, $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ becomes single-peaked. The rate of convergence is a function of both π_T and the quality of the indicators in $\mathbf{X}^*$. Generally speaking, the better are the indicators and the less extreme is π_T, the faster the rate of convergence. Table 1 then gives, for selected values of α and π_T, the value of $(p - 2)$ at which $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ becomes single-peaked, that is, at which

$$ k_2 < \left[\frac{\left(1 - \pi_T\right)}{\pi_T} \right]^{\frac{1}{(p-2)}} < k_1 $$

is satisfied. Also included in Table 1 are "effect sizes",

$$ \frac{(\alpha - 1) P\left(\mathbf{X}_i | \theta = T'\right)}{\sqrt{\pi_T V\left(\mathbf{X}_i | \theta = T\right) + \left(1 - \pi_T\right) V\left(\mathbf{X}_i | \theta = T'\right)}} , $$

even though, as shown in Theorem 4, the quantities that matter are the α_i. Table 1 shows that, when dichotomous indicators are employed, even when the researcher is confronted with a latent taxonic structure with extreme base rate (e.g., $\pi_T = .005$ or $.999$), four very good (say, e.g., $\alpha = 4$), or a larger number of mediocre, indicators in the $\mathbf{X}^*$ set will make (R1) true. When the indicators are of poor quality (say, e.g., $\alpha = 1.1$), a high base rate scenario (e.g., $\pi_T = .999$) will require many more indicators than a low base rate scenario (e.g., $\pi_T = .005$) before (R1) becomes true. It may then be concluded that, while (R1) is not true, it is made true, and MAXCOV-HITMAX a sense 1 criterion of latent taxonicity, under reasonable side-conditions.

Is R2 True?

In a 1996 article, Miller questioned the usefulness of MAXCOV-HITMAX, claiming that, contrary to Meehl's beliefs, $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is *not* necessarily flat in the case of continuous latent structures. He did not address the question as to whether there exist continuous latent structures for which $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ is actually single-peaked. Obviously, if there did exist such structures, then (R2) would be false. Miller's counter-example, which involved continuous indicators, is worth reviewing. Let the distribution of (U, V, W) be $N_3(\mathbf{0}, I)$. The structure that Miller considers, featuring three indicators, $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, may be represented as:

Table 1
The Minimum $(p-2)$ which Produces Single-Peakedness of $C(\mathbf{X}_i, \mathbf{X}_j | H = h)$ for Selected Values of π_T and α (effect size in brackets)

$\pi_T \setminus \alpha$	1.1	1.5	2.0	2.5	3.0	4.0	4.5
.001	73 (.05)	18 (.25)	10 (.50)	8 (.75)	7 (1.0)	5 (1.5)	5 (1.75)
.05	32 (.05)	8 (.25)	5 (.50)	4 (.74)	3 (.99)	3 (1.5)	2 (1.77)
.100	24 (.05)	6 (.25)	4 (.49)	3 (.73)	2 (.98)	2 (1.5)	2 (1.79)
.150	19 (.05)	5 (.24)	3 (.48)	2 (.72)	2 (.96)	2 (1.5)	2 (1.81)
.200	15 (.05)	4 (.24)	2 (.48)	2 (.71)	2 (.95)	1 (1.5)	1 (1.83)
.800	55 (.05)	11 (.22)	5 (.42)	3 (.62)	2 (.85)	1 (1.5)	1 (2.17)
.850	69 (.05)	13 (.22)	7 (.42)	4 (.62)	3 (.84)	2 (1.5)	1 (2.21)
.900	87 (.05)	17 (.22)	8 (.42)	5 (.61)	4 (.83)	2 (1.5)	2 (2.25)
.950	118 (.05)	26 (.22)	11 (.41)	7 (.61)	5 (.82)	3 (1.5)	2 (2.29)
.999	273 (.05)	52 (.22)	25 (.41)	15 (.60)	10 (.82)	5 (1.5)	4 (2.33)

$$(34) \qquad \mathbf{X} = U,$$
$$\mathbf{Y} = .6 * U + .8 * V,$$
$$\mathbf{Z} = 2.5 * U + 2 * V + 3 * W + .6 * U^2 + .8 * V * U.$$

Miller notes that $C(\mathbf{Y}, \mathbf{Z}|\mathbf{X} = x) = .64 * x + 1.6$, which is certainly not flat, but rather linear, increasing. However, this counterexample is not relevant to a consideration of MAXCOV-HITMAX, for Equation 34 is not a UMLV structure. As is well known (see, e.g., Holland & Rosenbaum, 1986), if $\mathbf{X}$ contains the indicators of a UMLV structure, the distribution of $\mathbf{X}$ is conditionally associated. Hence, for any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, and any function $g(\cdot)$, $C[\mathbf{X}_1, \mathbf{X}_2|g(\mathbf{X}^*) = h] \geq 0$ for all values h. The function $C(\mathbf{Y}, \mathbf{Z}|\mathbf{X} = x) = .64 * x + 1.6$, on the other hand, assumes negative values. Hence, the structure Miller presents could be screened for by employing Holland and Rosenbaum's UMLV tests prior to the application of MAXCOV-HITMAX. The relevant question is whether there exist UMLV structures in which θ is distributed continuously, and for which $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is not flat, but instead single-peaked.

Consider, for a set of dichotomous indicators, the 1-parameter non-parameteric (Rasch) item response model,

$$(35) \qquad P(\mathbf{X} = x_*) = \int_{-\infty}^{\infty} \prod_i P(\mathbf{X}_i = 1|\theta)^{x_{i*}} \left[1 - P(\mathbf{X}_i = 1|\theta)\right]^{1 - x_{i*}} dF(\theta)$$

in which the $P(\mathbf{X}_i = 1|\theta)$, that is, the item characteristic curves, are of the form

$$\frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)},$$

and $F(\theta)$, the distribution function of θ, is arbitrary. Cressie and Holland (1983) showed that, if $\mathbf{X}$ is described by this UMLV model, then

$$(36) \qquad P(\mathbf{X} = x_*) = P(0) \prod_{i=1}^{p} f_i^{x_{*i}} E(\mathbf{Z}^{x+})$$

in which $P(0) = P(\mathbf{X} = 0), f_j = \exp(-b_j) > 0, \mathbf{Z} = \exp(\theta), x+ = 1'x_*$, and $E(\mathbf{Z}^{x+})$ is the $x+^{\text{th}}$ moment of $\mathbf{Z}$. Since $\mathbf{Z}$ is a positive random variate, these moments are non-negative, and satisfy a set of inequalities (see Cressie & Holland, 1983). For four manifest variates, Equation 36 involves 7 parameters:

$E(\mathbf{Z}^1) = \mu_1$, $E(\mathbf{Z}^2) = \mu_2$, $E(\mathbf{Z}^3) = \mu_3$, $E(\mathbf{Z}^4) = \mu_4$, f_2, f_3, and f_4. To overcome an indeterminacy in the model, f_1 may be set to 1. The 16 manifest probabilities may be represented as:

(37)
$$P(1111) = P(0)f_2f_3f_4\mu_4 \qquad P(1110) = P(0)f_2f_3\mu_3$$
$$P(1101) = P(0)f_2f_4\mu_3 \qquad P(1011) = P(0)f_3f_4\mu_3$$
$$P(0111) = P(0)f_2f_3f_4\mu_3 \qquad P(1100) = P(0)f_2\mu_2$$
$$P(1010) = P(0)f_3\mu_2 \qquad P(1001) = P(0)f_4\mu_2$$
$$P(0110) = P(0)f_2f_3\mu_2 \qquad P(0101) = P(0)f_2f_4\mu_2$$
$$P(0011) = P(0)f_3f_4\mu_2 \qquad P(1000) = P(0)\mu_1$$
$$P(0100) = P(0)f_2\mu_1 \qquad P(0010) = P(0)f_3\mu_1$$
$$P(0001) = P(0)f_4\mu_1 \qquad P(0000) = P(0)$$

Partition $\mathbf{X}$ so that $\mathbf{X}_1$ corresponds to the first variate, $\mathbf{X}_2$, the second variate, and $\mathbf{X}^*$ contains the remaining variates. It then follows from the Identities of 37 that

(38)
$$C(\mathbf{X}_1, \mathbf{X}_2 | H = 0) = \frac{f_2\mu_2 - P(0)f_2(f_2\mu_2 + \mu_1)(\mu_1 + \mu_2)}{\left[1 + \mu_1 + f_2(\mu_1 + \mu_2)\right]},$$

(39)
$$C(\mathbf{X}_1, \mathbf{X}_2 | H = 1) = \frac{f_2\mu_3(f_3 + f_4) - P(0)f_2(\mu_3 + \mu_2)(f_3 + f_4)\left[\mu_3(f_2f_3 + f_2f_4) + \mu_2(f_3 + f_4)\right]}{\left\{(f_3 + f_4)\left[\mu_1 + \mu_2(1 + f_2) + f_2\mu_3\right]\right\}},$$

and

(40)
$$C(\mathbf{X}_1, \mathbf{X}_2 | H = 2) = \frac{f_2\mu_4 - P(0)f_2f_3f_4(f_2\mu_4 + \mu_3)(\mu_3 + \mu_4)}{\left[f_2\mu_4 + \mu_3(1 + f_2) + \mu_2\right]},$$

Now, Equations 38, 39 and 40 are not necessarily equal, and, by varying the parameters within their admissible ranges, $C(\mathbf{X}_1, \mathbf{X}_2 | H = h)$ may be shown to take on many different shapes, including the single-peakedness Meehl claims is a characteristic only of taxonicity. Consider the case in which $P(0) = .001$, $b_1 = 0$, $b_2 = 2.3$, $b_3 = -.10$, $b_4 = -1.37$, and $\mathbf{Z}$ has a mean of 2, variance of 10, skewness of .85, and kurtosis of $-.14$ (i.e., is mildly positively skewed and platykurtic). In this case, $C(\mathbf{X}_1, \mathbf{X}_2 | H = 0) = .3$, $C(\mathbf{X}_1, \mathbf{X}_2 | H = 1) = .49$, and $C(\mathbf{X}_1, \mathbf{X}_2 | H = 2) = .01$. On the other hand, when $P(0) = .003$, $b_1 = 0$, $b_2 = 1$, $b_3 = 0$, $b_4 = .33$, and $\mathbf{Z}$ has a mean of 1.5, variance of 6, skewness

of 2.3, and kurtosis of 5.6, $C(\mathbf{X}_1, \mathbf{X}_2|H = 0) = .49$, $C(\mathbf{X}_1, \mathbf{X}_2|H = 1) = .75$, and $C(\mathbf{X}_1, \mathbf{X}_2|H = 2) = .17$. Note that in neither of these cases are the item characteristic curves steep. Other brands of non-flat conditional covariance function may also be produced. When $P(0) = .003$, $b_1 = 0$, $b_2 = 1$, $b_3 = 0$, $b_4 = .33$, and $\mathbf{Z}$ has a mean of 3, variance of 9, skewness of -1.8, and kurtosis of -1.9, $C(\mathbf{X}_1, \mathbf{X}_2|H = 0) = .55$, $C(\mathbf{X}_1, \mathbf{X}_2|H = 1) = .52$, and $C(\mathbf{X}_1, \mathbf{X}_2|H = 2) = .03$. Finally, when $P(0) = .011$, $b_1 = 0$, $b_2 = -1.1$, $b_3 = 3.5$, $b_4 = 1.05$, and $\mathbf{Z}$ has a mean of 1.3, variance of 2, skewness of .38, and kurtosis of 3.75, $C(\mathbf{X}_1, \mathbf{X}_2|H = 0) = .52$, $C(\mathbf{X}_1, \mathbf{X}_2|H = 1) = .25$, and $C(\mathbf{X}_1, \mathbf{X}_2|H = 2) = .76$. Hence, by counterexample, R2 has been shown to be false. Furthermore, there is no reason to believe that such continuous Rasch structures will, in practice, be seldom encountered. They are not particularly esoteric structures. Nor is it unreasonable to posit that more complex continuous UMLV structures, they containing a greater number of parameters, will also produce non-flat conditional covariance functions.

Discussion

It has been shown that, for the case of dichotomous indicators:

1. A latent taxonic structure does not necessarily produce a single-peaked conditional covariance function. That is, (R1) is not true. This is because [M1∩M2∩M3] does not imply that $P(\theta = T|H = h)$ crosses .5.

2. (R1) is made true given that a number of distinct side-conditions hold. In particular, it is made true given that: (a) $\mathbf{X}^*$ contains a large number of indicators; (b) the latent taxonic structure to be detected is characterized by a value of π_T close to .5; (c) The indicators in $\mathbf{X}^*$ are of good quality. When the indicators are of poor quality, a high base rate scenario (e.g., $\pi_T = .999$) will require many more indicators than a low base rate scenario (e.g., $\pi_T = .005$) before (R1) becomes true.

3. It is not true that continuous latent structures do not produce single-peaked conditional covariance functions. Certain Rasch structures, for example, can produce such curves. Hence, (R2) is not true.

It may then be concluded that, for the case of dichotomous indicators, MAXCOV cannot be employed to distinguish between continuous and taxonic UMLV structures, for it is, at best (given the side-conditions described previously), a sense 1 criterion. The researcher may, of course, employ MAXCOV as follows: If, for a large value of $(p-2)$, $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is not single-peaked, then the data is not in keeping with the hypothesis of taxonicity. However, if $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ *is* single-peaked, some other criterion would be needed to eliminate as candidates, the continuous UMLV

structures.

There exist a number of responses that might be made in regards either the tone or content of this conclusion. First, it might be re-emphasized that the conclusion does not constitute an indictment of MAXCOV per se, but only the use of dichotomous indicators. Researchers might then conclude from the present work that continuous indicators should instead be employed with MAXCOV. We wish only to note that, for the case of continuous indicators, neither (R1), nor (R2), have been proven, a point we revisit later in this section. Hence, there is no reason to believe that MAXCOV stands on firmer footing under this alternative employment. A second response might be to note that, even though MAXCOV is, at best, a sense 1 criterion, the same can be said of many other latent variable models. For example, the covariance structure $(\Lambda_2 \Lambda_2' + \Psi, \Psi$ diagonal and positive definite), an often employed sense 1 criterion for the 2-dimensional linear factor analytic structure, cannot be used to rule out as a possibility the cases of the unidimensional, quadratic factor structure, nor a range of latent profile structures. Finally, one might attempt to downplay the significance of the fact that (R2) was proven false by counterexample by arguing that a demonstration that "a few specific continuous structures can produce single-peaked covariance functions is a long way from proving that this will occur under conditions that the researcher can expect to encounter in practice." In our opinion, this rebuttal has a hollow ring to it. To speak of "the latent structures one will encounter in practice" betrays a puzzling overconfidence. Because inferences are being made about *latent structures*, the researcher does not *know* what he will encounter, this, presumably, the reason that decision-making machinery such as MAXCOV is needed. If MAXCOV only makes correct decisions about whether [M1∩M2∩M3] is the case given that such taxonic structures have properties $(l_1, ..., l_t)$, and, hence, its correct employment in a particular context requires prior knowledge as to whether [M1∩M2∩M3]∩$(l_1, ..., l_t)$ is the case, then MAXCOV's usefulness is severely limited.[6] Indeed, if the researcher is truly so confident in regards his knowledge of the latent structures that he will meet in applied research, then he has no need for MAXCOV.

There is no evidence to indicate that Meehl and his co-workers have appreciated the fact that, in order for $P(\theta = T|H = h)$ to be an increasing function of h, the joint distribution of H and θ must be monotone likelihood ratio dependent. Theorem 3 establishes that [M1∩M2∩M3] does, in fact,

---

[6] In the same sense that the usefulness of the Newman-Keuls procedure is vastly reduced by the fact that it can only be counted on to make correct decisions given *certain* patterns of population mean differences. If the researcher was in a position to know whether these patterns held, then why would Newman-Keuls be needed?

induce this strong brand of dependency. Interestingly, the Lemma which supports Theorem 3 also provides the basis for an alternative to Huynh's (1994) proof for the case of dichotomous indicators and continuous θ. For any two values of θ, say, $\theta_o < \theta_1$, simply apply the Lemma to the case in which, for all i, $a_i = P_i(\theta_o)$ and $a_i' = P_i(\theta_1)$, in which $P_i(z)$ is the item characteristic curve of indicator i evaluated at $\theta = z$.

Meehl and Yonce (1996, p. 1113) conjectured that (R2) would be violated if MAXCOV was faced with latent structures in which the items have the same difficulty levels and steep item characteristic curves. It is not clear how useful is this advice, for, as the counter-example employed in the current work shows, the situation with regards (R2) is a complex one. $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$ is determined by the joint distribution of $\mathbf{X}_1$, $\mathbf{X}_2$, and H, which, given that $\mathbf{X}$ conforms to a UMLV structure, is, in turn, determined not only by the form of the regressions (item characteristic curves) of the indicators on θ, but also by the distribution of θ. For example, the simple Rasch structure of Equation 42, and, hence, $C(\mathbf{X}_1, \mathbf{X}_2|H = h)$, is determined by fully $2p - 1$ parameters (p moments and $[p - 1]$ regression parameters). Since only UMLV structures are admissible, the moment parameters are, additionally, subject to a set of restrictions (Cressie & Holland, 1983). More complex continuous UMLV structures will involve a greater number of parameters and restrictions (see Cressie & Holland, 1983). The point then is that a useful analysis of (R2) must involve a detailed characterization of the continuous UMLV structures under consideration.

For neither (R1), nor (R2), has Meehl provided, for the case of continuous indicators, formal proofs, and the reader may wonder whether the results of the current work can be used to ground such proofs. For a set of continuous indicators, M1 remains as in Equation 1. Define an "indicator" of T to be a continuous random variate, $\mathbf{X}_i$, with the property that, after recoding,

(41) $$1 - F_{xi|\theta = T}(x) > 1 - F_{xi|\theta = T'}(x), \text{ for all values } x,$$

in which $F_{xi|\theta = t}(x) = P(\mathbf{X}_i \leq x|\theta = t)$. Condition M2C then states that there exists a set of p continuous indicators of the taxon, these indicators stored in the random vector $\mathbf{X}$. Finally, for the case of continuous indicators, M3C states, in analogy to M3, that

(42) $$f_{\mathbf{X}|\theta = t} = \prod_{i=1}^{p} f_{\mathbf{X}_i|\theta = t},$$

that is, the joint density of the indicators conditional on $\theta = t$ is a product of the individual conditional densities. Let [M1∩M2C∩M3C] symbolize the case of latent taxonicity with continuous indicators.

For (R1) to be true, it must, once again, be true that, for any partition $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ of $\mathbf{X}$, [M1∩M2C∩M3C] implies Properties 18-21. By substituting density functions for discrete mass functions in the proofs of 18 and 19 (Theorems 1 and 2), 18 and 19 are easily shown to follow from [M1∩M2C∩M3C]. Modifying the proof of Equation 20 in the same fashion results in the condition (analogous to Equation 29) that, for $P(\theta = T|H = h)$ to be an increasing function of h, that is, $P(\theta = T|H = h + \varepsilon) > P(\theta = T|H = h)$, $\varepsilon > 0$, it must be the case that

$$(43) \qquad f_{H|\theta = T}(h + \varepsilon)\, f_{H|\theta = T'}(h) > f_{H|\theta = T}(h) f_{H|\theta = T'}(h + \varepsilon)$$

in which $f_{H|\theta = t}(s)$ is the value, when $H = s$, of the conditional density of $H = 1'\mathbf{X}^*$ given $\theta = t$. This, once again, is the condition that H and θ are monotone likelihood ratio dependent. The question then is whether [M1∩M2C∩M3C] induces this brand of dependence. The answer is that it does not necessarily do so. Lemma 2 of Holland and Rosenbaum (1986) establishes that [M1∩M2C∩M3C] induces a weaker form of dependence on θ of functions of $\mathbf{X}^*$, namely that, for any increasing function, $g(\mathbf{X}^*)$, of $\mathbf{X}^*$, $E[g(\mathbf{X}^*)|\theta = t]$ is an increasing function of t. The mlrd property is not implied by [M1∩M2C∩M3C]. Recall that the mlrd property *was* implied by [M1∩M2∩M3] for the case of dichotomous indicators, in agreement with Holland and Rosenbaum's (1986) finding that dichotomous indicators of UMLV structures manifest stronger dependencies than do continuous indicators. Hence, for the case of continuous indicators, (R1) is not true.

As for the dichotomous case, it is possible that there exist reasonable side conditions under which (R1) is true, an obvious candidate being that of large $(p - 2)$. Now, as $(p - 2)$ becomes large, $f_{H|\theta = t}$ will converge to a normal distribution with, say, mean $\mu(t)$, and variance $\sigma^2(t)$ (Basawa & Rao, 1980; Holland, 1990), that is,

$$(44) \qquad f_{H|\theta = t} \to \frac{1}{\left[2\pi\sigma^2(t)\right]^{1/2}} \exp\left\{ -\frac{\left[H - \mu(t)\right]^2}{2\sigma^2(t)} \right\}$$

From lemma 2 of Holland and Rosenbaum (1986), $\mu(T) = E(H|\theta = T) > E(H|\theta = T') = \mu(T')$, and, hence, $\mu(T) = \mu(T') + \gamma$, $\gamma > 0$. Substituting the right member of 44 into 43, and taking the natural logarithm of both sides of the inequality results in the condition that

M. Maraun, K. Slaney, and L. Goddyn

$$(45) \quad \frac{\left[h+\varepsilon-\mu(T')\right]^2}{\sigma^2(T')} + \frac{\left[h-\mu(T')-\gamma\right]^2}{\sigma^2(T)} - \frac{\left[h+\varepsilon-\mu(T')-\gamma\right]^2}{\sigma^2(T)} - \frac{\left[h-\mu(T')\right]^2}{\sigma^2(T')}$$

must be positive. Expansion of Equation 45 shows that positivity is not guaranteed. Interestingly then, despite the doubts raised about the use of dichotomous indicators, their use more directly ensures the truth of (R1) than does the use of continuous indicators. For the special case in which $\sigma^2(T) = \sigma^2(T') = \sigma^2$, Equation 45 reduces to

$$\frac{2\gamma\varepsilon}{\sigma^2},$$

which *is* positive. This results, essentially, from the fact that, for this particular case, that is, large $(p-2)$, $\mu(T) > \mu(T')$, and $\sigma^2(T) = \sigma^2(T')$, $f_{H|\theta=t}$ is log-concave, that is, it is a strongly unimodal density (Lehmann, 1966). Hence, for a large value of $(p-2)$, and given that $\sigma^2(T) = \sigma^2(T')$, $P(\theta = T|H = h)$ is, for the case of continuous indicators, an increasing function of h. While conditional normality of H given $\theta = t$ under large $(p-2)$ is a reasonable side-condition, it is not easy to judge whether the same may be said of the requirement that $\sigma^2(T) = \sigma^2(T') = \sigma^2$. It is true, however, that for the linear factor model, the same property, that is, that $\sigma^2_{H|\theta=t} = \sigma^2(t) = 1'\boldsymbol{\Psi}1$, for all values t, follows from a defining feature of that model, namely that, for all values t, $C(\mathbf{X}|\theta = t) = \boldsymbol{\Psi}$, $\boldsymbol{\Psi}$ diagonal. We have not, as of yet, for the case of continuous indicators, been able to generate a reasonable counter-example to (R2).

References

Bartholomew, D. & Knott, M. (1999). *Latent variable models and factor analysis*. London: Arnold.

Basawa, I. & Rao, B. (1980). *Statistical inference for stochastic processes*. New York: Academic Press.

Cressie, N. & Holland, P. (1983). Characterizing the manifest probabilities of latent trait models. *Psychometrika, 48(1)*, 129-141.

Grayson, D. (1988). Two-group classification in latent trait theory: Scores with monotone likelihood ratio. *Psychometrika, 53(3)*, 383-392.

Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences*. London: Sage Publications.

Hemker, B., Sijtsma, K., Molenaar, I., & Junker, B. (1997). Stochastic ordering using the latent trait and the sum score in polytomous IRT models. *Psychometrika, 62(3)*, 331-347.

Holland, P. (1990). The Dutch identity: A new tool for the study of item response models. *Psychometrika, 55(1)*, 5-18.

Holland, P. & Rosenbaum, P. (1986). Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statistics, 14*(*4*), 1523-1543.

Huynh, H. (1994). A new proof for monotone likelihood ratio for the sum of independent bernoulli random variables. *Psychometrika, 59*(*1*), 77-79.

Junker, B. (1993). Conditional association, essential independence and monotone unidimensional item response models. *The Annals of Statistics, 21*, 1359-1378.

Karlin, S. & Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. 1. Multivariate totally positive distributions. *Journal of Multivariate Analysis, 10*, 467-498.

Lehmann, E. (1966). Some concepts of dependence. *Annals of Mathematical Statistics, 37*, 1137-1153.

Maraun, M. & Rossi, N. (2001). The extra-factor phenomenon revisited: Unidimensional unfolding as quadratic factor analysis. *Applied Psychological Measurement, 25*(*1*), 77-87.

McDonald, R. P. (1967). *Nonlinear factor analysis*. Richmond, VA: The William Byrd Press.

Meehl, P. E. (1965). *Detecting latent clinical taxa by quantitative indicators lacking an accepted criterion*. (Report PR-65-2). Minneapolis: University of Minnesota Department of Psychiatry.

Meehl, P. E. (1973). MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In P. E. Meehl (Ed.), *Psychodiagnosis: Selected articles* (pp. 200-224). Minneapolis: University of Minnesota Press.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60*, 117-174.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50*(*4*), 266-275.

Meehl, P. E. & Golden, R. R. (1982). Taxometric methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127-181). New York: John Wiley & Sons.

Meehl, P. & Yonce, L. (1996). Taxometric analysis: II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV PROCEDURE). [Monograph Supplement]. *Psychological Reports*, 1091-1227.

Miller, M. (1996). Limitations of Meehl's MAXCOV-HITMAX procedure. *American Psychologist, 51*, 554-556.

Molenaar, I. & von Eye, A. (1994). On the arbitrary nature of latent variables. In A. von Eye & C. Clogg (Eds.), *Latent variables analysis*. London: Sage.

Rosenbaum, P. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. *Psychometrika, 49*, 425-536.

Stanley, R. (1986). *Enumerative combinatorics*. New York: Wadsworth and Brooks.

Stout, W. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika, 55*, 293-325.

Tukey, J. (1958). A problem of Berkson, and minimum variance orderly estimators. *Annals of Mathematical Statistics, 29*, 588-592.

Waller, N. G. & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks: Sage.

M. Maraun, K. Slaney, and L. Goddyn

Appendix
Proof of Theorem 3

To prove Theorem 3, the following combinatoric lemma is required.

Lemma

Let $0 \le a_i \le 1$, $i = 1, ..., p$, and define the polynomial

$$\text{(46)} \qquad f(t) = f(a_1, ..., a_p; t) = \prod_{i=1}^{p} \left[a_i t + (1 - a_i) \right].$$

Then

$$f(t) = \sum_{j=0}^{p} A_j t^j$$

in which the coefficients, A_j, are equal to

$$\text{(47)} \qquad A_j = \sum_{S \in \left[\binom{(p)}{j} \right]} \prod_{i \in S} a_i \prod_{i \in (p) - S} (1 - a_i).$$

Here, S runs over the p choose j partitions of $(p) = (1, 2, ..., p)$ into j and $(p - j)$ elements. For example, if $(p) = (1, 2, 3)$, then $A_0 = (1 - a_1)(1 - a_2)(1 - a_3)$, $A_1 = a_1(1 - a_2)(1 - a_3) + (1 - a_1)a_2(1 - a_3) + (1 - a_1)(1 - a_2)a_3$, $A_2 = a_1 a_2 (1 - a_3) + a_1(1 - a_2)a_3 + (1 - a_1)a_2 a_3$, and $A_3 = a_1 a_2 a_3$. Similarly, let $0 \le a_i' \le 1$, $i = 1, ..., p$, and define A_j' via

$$f(a_1', ..., a_p'; t) = \sum_{j=0}^{p} A_j' t^j$$

to be

$$\text{(48)} \qquad A_j' = \sum_{S \in \left[\binom{(p)}{j} \right]} \prod_{i \in S} a_i' \prod_{i \in (p) - S} (1 - a_i').$$

If $0 < a_i \le a_i' < 1$, $i = 1, ..., p$, then

$$0 \le \frac{A'_o}{A_o} \le \frac{A'_1}{A_1} \le ... \le \frac{A'_p}{A_p},$$

or, equivalently, $A'_k A_m - A'_m A_k \ge 0$ for $0 \le m \le k \le p - 1$. These inequalities are strict if, for at least one i, $a_i \ne a'_i$.

Proof

First consider the special case in which $a'_i = a_i$, $i = 1, ..., (p - 1)$, and in which, for some $\alpha > 0$, $a'_p = a_p + \alpha$. Define $B_o, ..., B_{p-1}$ via

$$f(a_1, ..., a_{p-1}, 0; t) = \prod_{i=1}^{p-1} \left[a_i t + (1 - a_i) \right] = \sum_{j=0}^{p-1} B_j t^j,$$

so that

(49)
$$B_j = \sum_{S \in \left[\begin{smallmatrix} (p-1) \\ j \end{smallmatrix} \right]} \prod_{i \in S} a_i \prod_{i \in (p-1) - S} (1 - a_i).$$

Here, S runs over the $(p - 1)$ choose j partitions of $(p - 1) = [1, 2, ..., (p - 1)]$ into j and $(p - 1 - j)$ elements. By separating out those terms in the summation of Equation 47 for which $p \epsilon S$, it may be shown that, for $j = 1, ..., (p - 1)$, $A_j = a_p B_{j-1} + (1 - a_p)B_j$ and $A'_j = (a_p + \alpha)B_{j-1} + (1 - a_p - \alpha)B_j$. These equations are valid for $j = 0$ and $j = p$ provided B_{-1} and B_p are defined to be 0. By substituting and simplifying, it may be shown that, for $j = 0, ..., (p - 1)$,

(50)
$$A'_{j+1} A_j - A'_j A_{j+1} = \alpha(B_j^2 - B_{j-1} B_{j+1})$$

Since $\alpha > 0$, the expression on the left is positive provided that $B_j^2 > B_{j-1} B_{j+1}$, that is, the coefficients of $f(a_1, ..., a_{p-1}, 0; t)$ are log concave. But log-concavity follows from the fact that the roots of $f(a_1, ..., a_{p-1}, 0; t)$, namely

$$\left(\frac{a_i - 1}{a_i} : i = 1, ..., p - 1 \right),$$

are all real and nonpositive (Stanley, 1986). This proves the special case. The general case follows from the fact that

$$\frac{A_j''}{A_j'} \le \frac{A_{j+1}''}{A_{j+1}'} \ \& \ \frac{A_j'}{A_j} < \frac{A_{j+1}'}{A_{j+1}} \Rightarrow \frac{A_j''}{A_j} < \frac{A_{j+1}''}{A_{j+1}}. \quad \blacksquare$$

Proof of Theorem 1

Let $[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}^*]$ be any partition of $\mathbf{X}$. The $(p - 1)$ probabilities $P(H = h|\theta = t)$, $t = 0, ..., (p - 2)$, are each sums of the $(p - 2)$ choose h probabilities $P(\mathbf{X}^* = x: 1'x = h|\theta = t)$, that is,

$$(51) \qquad P(H = h|\theta = t) = \sum_{S \in \left[\binom{(p-2)}{h}\right]} \prod_{i \in S} P(\mathbf{X}_i = 1|\theta = t) \prod_{i \in (p-2) - S} \left[1 - P(\mathbf{X}_i = 1|\theta = t)\right].$$

Noting that, by M2, $P(\mathbf{X}_i = 1|\theta = T) > P(\mathbf{X}_i = 1|\theta = T')$ for all i, one may take $P(H = k|\theta = T)$ to be A_k', $P(H = k|\theta = T')$ to be A_k, $P(H = m|\theta = T)$ to be A_m', and $P(H = m|\theta = T')$ to be A_m. Since the requirements of the lemma are then met, it then follows that $P(H = k|\theta = T)P(H = m|\theta = T') > P(H = m|\theta = T)P(H = k|\theta = T')$, that is, that H and θ are mlrd. $\quad \blacksquare$