TAIL-MEASURABILITY IN MONOTONE LATENT VARIABLE MODELS

Jules L. Ellis
University of Nijmegen

Brian W. Junker
Carnegie Mellon University
Pittsburgh

Running Head: TAIL-MEASURABILITY

The authors made equivalent contributions to the results of this article. Ellis' research was supported by the Dutch Interuniversitary Graduate School of Psychometrics and Sociometrics. Junker's research was supported by ONR Grant N00014-87-K-0277, NIMH Grant MH15758, and a Carnegie Mellon University Faculty Development Grant. In addition Junker would like to acknowledge the hospitality of the Nijmegen Institute for Cognition and Information during his visit to the University of Nijmegen in August 5--10, 1993. Requests for reprints should be send to Jules L. Ellis, Department of Mathematical Psychology, University of Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, THE NETHERLANDS.

Abstract

We consider latent variable models for an infinite sequence (or universe) of manifest (observable) variables that may be discrete, continuous or some combination of these. The main theorem is a general characterization by empirical conditions of when it is possible to construct latent variable models that satisfy unidimensionality, monotonicity, conditional independence, and *tail-measurability*. Tail-measurability means that the latent variable can be estimated consistently from the sequence of manifest variables even though an arbitrary finite subsequence has been removed. The characterizing, *necessary and sufficient*, conditions that the manifest variables must satisfy for these models are conditional association and vanishing conditional dependence (as one conditions upon successively more other manifest variables). Our main theorem considerably generalizes and sharpens earlier results of Ellis and Van den Wollenberg (1993), Holland and Rosenbaum (1986), and Junker (1993). It is also related to the work of Stout (1990).

The main theorem is preceded by many results for latent variable models *in general* - not necessarily unidimensional and monotone. They pertain to the uniqueness of latent variables and are connected with the conditional independence theorem of Suppes and Zanotti (1981). We discuss new, definitions of the concepts of "true-score" and "subpopulation", which generalize these notions from the "stochastic subject", "random sampling", and "domain sampling" formulations of latent variable models (e.g., Holland, 1990; Lord & Novick, 1968). These definitions do not require the a priori specification of a latent variable model.

1. Introduction

In many testing and assessment settings it is reasonable to view the set of observed variables as a sample from an infinite *universe of admissible manifest (or observable) variables*. The variables in this universe are admissible in the sense that they satisfy some general substantive definition that specifies the field of research. For instance, the universe may be defined as "responses to mental rotation items of the class defined by Shepard et al." or "responses to items from personality questionnaires". Each investigator in the field will select a particular finite subset of variables from the universe, but with the pretention that the essential conclusions can the generalized to the whole universe. Often we hope to fit some statistical model to the universe. It is essential here that the model is claimed to hold for the whole universe, not merely for the particular finite subset of variables that the investigator selected. This view is immanent in image factor analysis (Guttman, 1953), alpha factor analysis (Kaiser & Caffrey, 1965), generalizability theory (Cronbach, Gleser, Nanda & Rajaratnam, 1972), domain sampling true-score theory (Lord & Novick, 1968, Ch. 8), the theory of specific objectivity in Rasch models (Fischer, 1987, 1995; Rasch, 1960), essential unidimensionality (Stout, 1990) and the work of many other test theorists (e.g., Mulaik & McDonald, 1978). The present paper elaborates on this.

We will consider an infinite sequence of manifest random variables $\mathbf{X} = (X_1, X_2, ...)$ on a common probability space. $\mathbf{X}$ may be seen as a universe of admissible manifest variables. At this point we assume that the variables in $\mathbf{X}$ are arranged in some fixed, artificial order. Later, in section 9, it will be shown that this assumption does not affect any of our conclusions. Some examples are: (1) X_i is the score variable on test item *i* of a specified pool; (2) X_i is the reaction time variable on trial *i* of a repeated recognition task; (3) X_i is the total score on test *i* of an infinite battery; (4) X_i consists of the ratings of judge *i* to a set of targets.

Latent variable models assume that there exists a (usually latent) random vector Θ such that the joint distribution of $(\mathbf{X}, \Theta)$ satisfies certain restrictions (e.g., McDonald, 1981). This paper introduces the restriction of *tail-measurability*. It will be shown that tail-

measurability is equivalent to the possibility of estimating Θ consistently from any "tail" of manifest variables X_n, X_{n+1},... - that is, even though observations on an arbitrary finite number of the manifest variables may be missing. Tail-measurability is a mathematization of what a psychologist means by the concept of "latent variable" in that a latent variable *should not* rely necessarily on any particular finite sequence of manifest variables and further *should* be determined by any tail of manifest variables, regardless of the removal of a finite subsequence of manifest variables. Tail-measurability is, therefore, a reasonable restriction to add to latent variable models. Junker (1993) explores this condition and argues that it is needed in every useful latent variable model for measurement. A few other new, but tangential, concepts are needed in this paper: The concepts of *tail-conditional independence* and *vanishing conditional dependence* will be introduced, and shown to be equivalent. Also, a *new definition of true-score variables* will be discussed.

Many latent variable models involve the restriction that $\mathbf{X}$ is conditionally independent given Θ. A fundamental theorem will be presented that concerns tail-measurability and conditional independence jointly. It can be derived from this theorem that tail-measurability and conditional independence jointly imply tail-conditional independence. The theorem also has consequences for the *uniqueness* of latent variables and true-scores.

We then focus on one of the most basic (classes of) latent variable models, which involves the following restrictions (Holland & Rosenbaum, 1986; Mokken, 1971):

1. Θ is unidimensional (or real-valued),
2. $\mathbf{X}$ is conditionally independent given Θ, and
3. monotonicity: each $P[X_i > x|\Theta]$ is nondecreasing in Θ.

Throughout Θ will be called a "random variable" only if it is unidimensional. We say that $\mathbf{X}$ *has a monotone latent variable (LV) representation* if a random variable Θ exists such that the above three conditions hold. In that case we will also say that $\mathbf{X}$ *satisfies the monotone LV model*. Some special cases of the model are: the 3-parameter logistic IRT model (Lord, 1980), graded response models (Andrich, 1978), and the linear 1-factor model with nonnegative loadings.

An important question is how the monotone LV model can be

characterized by manifest, and thus empirically verifiable, conditions. That is: Which conditions should **X** satisfy in order to conclude that a latent variable Θ exists such that **X** satisfies the monotone LV model with Θ? For binary manifest variables, Mokken (1971) showed that the model implies that the manifest variables have non-negative covariances, Holland (1981) showed that they are strongly positive orthant dependent, and Rosenbaum (1984) showed that they are *conditionally associated*. This was generalized to the non-binary case by Holland and Rosenbaum (1986). These conditions are merely necessary for the model. Several authors have tried to establish conditions that are both *necessary and sufficient* for the model. This has been achieved only for special versions of the model. Ellis and Van den Wollenberg (1993) stated the necessary and sufficient conditions for **X** to have a monotone LV representation in the case of binary variables X_i and with the additional restriction of local homogeneity - which presupposes stochastic subjects. Junker (1993) stated the necessary and sufficient conditions for the model in the binary case with an infinite sequence of items and with the additional restriction of locally asymptotically discriminating (LAD) items, which is a concept requiring reference to a latent variable and thus not a manifest condition.

The main theorem of this paper is a major generalization of the previous results:

> **X** satisfies the monotone LV model (i.e., there exists a monotone LV representation for **X**) with a latent variable Θ that is tail-measurable if and only if **X** has nonnegative conditional association and vanishing conditional dependence. This Θ is unique up to a strictly increasing transformation.

This characterization of the model does not require discrete manifest variables, stochastic subjects, or LAD items. The only additional constraint is that of tail-measurability - a fairly weak and reasonable restriction. Note that it need not be assumed in advance: tail-measurability is implied by the empirical conditions. The result is also discussed by Junker and Ellis (1995) in the context of mathematical statistics and exchangeability.

The main concepts and results of this paper are sketched in the following sections such that they hopefully can be understood and

appreciated by readers who are not familiar with probabilistic measure theory (not to be confused with the psychological theory of measurement). A rigorous development of the theory, however, cannot avoid measure theory. Therefore this is presented as an isolated appendix *Lemmas and Proofs* at the end. Lemmas to which other sections refer can be found there.

Overview of the paper. After (1) this introduction, we will subsequently (2) specify our notation and recall some elementary concepts; (3) introduce the concept of tail-measurability and its relationship with consistent estimators; (4) recall the definition of conditional independence and introduce the concepts of tail-conditional independence and vanishing conditional dependence; (5) discuss a new definition of true-score variables; (6) present the fundamental theorem concerning tail-measurability and conditional independence, and discuss its consequences; (7 and 8) discuss two variants of the characterization theorem for the monotone LV model; (9) extend the theory to permutations and infinite subsequences; (10) indicate some issues for future research; (11) discuss our theory, and (12) state and prove all needed lemmas.

2. Notation and Preliminary Concepts

We will consider an infinite sequence of manifest random variables

$$\mathbf{X} = (X_1, X_2, \ldots).$$

The variables X_i putatively address the empirical domain of interest. Note that they are arranged in some fixed, artificial order. In section 9 it will be shown that our results do not depend on that order, however. In addition, we consider some finite-dimensional random vector

$$\Theta = (\Theta_1, \ldots, \Theta_d).$$

At this point, we do not make further assumptions about Θ: The symbol Θ will be used for *any* random vector of which one might wonder whether it can serve as an "explanatory" set of variables for $\mathbf{X}$. It is useful to think of Θ as a *latent* random vector, but there is no mathematical need to

require this in advance.

Because this paper is concerned with the foundations of latent variable models, it is worthwhile to spell out a few elementary definitions about which confusion might arise. The theory will be confined to a single probability space with outcome space (sample space) Ω and probability measure P. Every random variable X is *a function of the outcomes in* Ω. If X is a random variable, and x a real number, then the event $[X \leq x]$ is defined as the set of outcomes v in Ω that satisfy $X(v) \leq x$. For more information, see e.g. Mood, Graybill and Boes (1974, Ch. 1-2).

Following Holland (1990), Ellis (1993) and Ellis and Van den Wollenberg (1993) discuss two conceptually different formulations of latent variable models that will receive attention in this paper too. In the *random sampling* formulation, Ω is a population of persons. Consequently, each person has a fixed score on each of the manifest variables. In the *stochastic subject* formulation, on the other hand, Ω is a set of (person, occasion) pairs. This allows for the possibility of random within-person variability of the manifest variables, as the sampled person is subjected to a randomly selected occasion.

An *event* is a subset of Ω. Thus in the random sampling formulation, an event is a "subpopulation" of persons. The main theorems of this paper involve a form of subpopulation invariance (Ellis, 1993; Meredith & Millsap, 1992). It is important therefore to know *which subpopulations can be selected on basis of a given random variable* X. For instance, consider $\Omega = \{u,v,w\}$ and let X be the random variable with $X(u) = X(v) = 0$ and $X(w) = 1$. Now the following subpopulations can be selected on basis of X: Ø, $\{u,v\}$, $\{w\}$ and Ω. This may be called the *event space* of X. The event space tells us how X partitions the population into subpopulations. It does not contain the event $\{u\}$, for instance, because one cannot distinguish u from v on basis of X alone.

It is a convention to denote the event space of a random variable X by $\sigma(X)$. A better understanding of event spaces may be obtained by the following observation: $\sigma(Y) \subseteq \sigma(X)$ if and only if $Y = f(X)$ for some function f. In that case we say that the event space of Y is *coarser* than the event space of X. A consequence of this is that two random variables have the same event space if and only if they are an 1-to-1 transformation of each other. Therefore, $\sigma(X)$ can be identified with the equivalence

class of all random variables that are a 1-to-1 transformation of X. Any such random variable induces the same partition of Ω.

Event spaces are defined more exactly and generally as "σ-fields" in a rigorous probabilistic measure theory treatment. The natural definitions of the central concepts of this paper therefore involve σ-fields. A fairly good intuitive understanding can be achieved on basis of the above description of event spaces. As a courtesy to the reader we recall the formal definition too: A collection *F* of events is called a σ-*field* if it satisfies these three conditions: (1) Ω is element of *F*, (2) F is closed under the formation of complements, and (3) *F* is closed under the formation of countable unions or, equivalently, countable intersections. By "closed" is meant that these operations, complement and countable union, if performed on events in *F*, will always result in an event that is again an element of *F*. Thus *F* can be seen as a collection of events under discussion, and logical operations with these events will never produce events outside this collection. The interested reader may consult the introductory statistics book of Mood, Graybill and Boes (1974, Ch. 1).

Two particular σ-fields are important here (for more information, see Billingsley, 1986):

1. $\sigma(\Theta)$, *the* σ-*field generated by* Θ. This is the event space of Θ. It is defined as the smallest σ-field that contains all events of the form $[\Theta_1 \leq y_1,...,\Theta_d \leq y_d]$. Less formally stated, $\sigma(\Theta)$ is the collection of events that can be described (without exotic constructions) by Θ. For example, if d = 1, a typical element of $\sigma(\Theta)$ is $[2 < \Theta \leq 3]$. Similar descriptions can of course be given for, say, $\sigma(\mathbf{X})$, the σ-field generated by the sequence of manifest variables **X**.

2. $\tau(\mathbf{X})$, *the tail-*σ-*field of* **X**. This is the event space that contains the "long-run" events of **X**. These are the events that can be defined by information that is "generalizable" in the sense that it is common to all tails of **X**. Elements of $\tau(\mathbf{X})$ are called *tail-events*. Tail-events are events that are determined solely by the *asymptotic* behavior of **X**.

The tail-σ-field, $\tau(\mathbf{X})$, plays a very important role in this paper because it ties the central definitions together. It is

formally defined as $\cap_n \sigma(X_n, X_{n+1}, ...)$. This means that an event is a tail-event if and only if it belongs to $\sigma(X_n, X_{n+1}, ...)$ for every n. Less formally stated, a tail-event is an event that can be described by every tail of **X**. Note that the σ-fields involved in this expression become coarser with n; each of these σ-fields is a subset of the preceding ones. It is therefore proper to say that $\sigma(X_n, X_{n+1}, ...)$ converges from above to $\tau(\mathbf{X})$ as n increases. The limit, $\tau(\mathbf{X})$, thus contains only very special events. It will not be totally empty since Ω and Ø belong to all σ-fields that were used in the intersection. Furthermore, $\tau(\mathbf{X})$ is a σ-field because any intersection of σ-fields is again a σ-field.

The remarkable characteristic of a tail-event is that *it is defined such that every tail of* **X** *contains sufficient information to determine wether the event occurs*. No particular X_i is needed for this and, moreover, no particular finite subsequence of X_i's is needed for this. The event can be described totally by **X** with any finite subsequence from it removed. This follows directly from the above definitions. There are some well-known examples of tail-events that may further clarify the concept:

a. In ordinary statistics, it is customary to let a sequence X_1, X_2, ... X_n represent the successive observations in a sample of size n. Assume that the X_i are i.i.d. and have finite expectation. The strong law of large numbers contends that, as n grows, the sample mean converges to the population mean with probability 1. *This involves a tail-event: The set of samples of which the sample mean converges to the population mean.* To judge whether a given sample belongs to this event, it suffices to know any tail of the sample. That is, for determining that the sample mean converges to the population mean, it is not necessary to know the values of the first 10 X_i's, or the first 100 X_i's or indeed the values of any finite number of X_i's. This is an essential feature of a tail-event. The present tail-event has probability 1 according to the law, ie., almost every sample belongs to this event.

b. Assume that the X_i are replicate measurements in the sense of classical test theory. That is, assume the existence of a latent

variabele T such that $E(X_i|T) = T$ and such that the X_i are conditionally i.i.d. given T. In each subpopulation [T = t] we may apply the strong law of large numbers to conclude that the average testscore, $\Sigma X_i/n$, converges to T with probability 1. *This again involves a tail-event: The set of subjects for which the average testscore converges to their score on* T. This is a tail-event because, to determine wether a given subject belongs to this event, it is not necessary to know, say, the first 100 X_i's or indeed any finite subsequence of X_i's.

In this example there is another important class of tail-events: *Every element of* $\sigma(T)$ *is a tail-event*. For example, the event $[T \leq 3]$ is typical element of $\sigma(T)$. Assume for simplicity that the average testscore converges to T for *all* subjects. Then $[T \leq 3]$ is identical to $[\lim_{n->\infty}\Sigma_{i=1}^{n}X_i/n \leq 3]$, which is a tail-event: Whether this event occurs depends on **X**, but not on any finite subsequence of **X**, since the contribution of such a subsequence vanishes if one passes to the limit.

This informal argument, that the tail-σ-field is closely related to the true score or latent trait, is central to our paper. See, for example, theorems 1 and 4 below. It also underlies the main results of our work, theorems 5 and 6 below.

The use of σ-fields allows a very general definition of conditional probabilities that includes conditioning on events of probability zero. We will first consider a formal definition, adapted from Billingsley (1986). The *conditional probability function* of an event A, given some σ-field *F*, is denoted by $P(A|F)$ and is uniquely defined by the following conditions:

1. $P(A|F)$ is a random variabele.
2. Every event of the form $[P(A|F) \leq r]$, with r a real number, is an event in *F*.
3. If S is an event in *F* with $P(S) > 0$ then $P(A|S) = E(P(A|F)|S)$. Here, $E(\ |S)$ is the conditional expectation in S according to the ordinary definition.

In particular, if Θ is a random variable and $F = \sigma(\Theta)$, the conditional probability function $P(A|\sigma(\Theta))$ is denoted as $P(A|\Theta)$. In that case condition (2) above may be replaced by

2'. $P(A|\Theta) = f(\Theta)$ for some measurable function f.

Here, measurability of f is a very weak regularity condition that needs no attention for an intuitive understanding.

A more informal description of conditional probability functions may be given as follows. Here, we assume for simplicity that there is a collection of "smallest events" E_i in F, such that the E_i form a partition of Ω and such that every event in F is an union of events E_i. For example, if $F = \sigma(\Theta)$, the smallest events are of the form $[\Theta = \downarrow]$. Conditional probability functions have the following properties.

1. $P(A|F)$ summarizes the values of $P(A|S)$ for all possible events S in F.
2. However, $P(A|F)$ does not contain directly all possible values $P(A|S)$. Instead, $P(A|F)$ will contain only the values of $P(A|E_i)$, thus the conditional probabilities given the smallest events in F. For example, $P(A|\Theta)$ will contain only the values $P(A|\Theta = \downarrow)$.
3. All other conditional probabilities $P(A|S)$ with $S \in F$ are obtained from $P(A|F)$ by integration of the probabilities $P(A|E_i)$. For example, $P(A|\Theta \leq t)$ is obtained as $\int_{-\infty}^{t} P(A|\Theta = \downarrow)dF(\downarrow)$.
4. $P(A|F)$ is a random variable; its domain is Ω, not F. This random variable contains the relevant conditional probabilities in the following way: If $v \in E_i$ and then $P(A|F)(v) = P(A|E_i)$. For example, if v is a subject with $\Theta(v) = 2$ then the "score" of v on the random variable $P(A|\Theta)$ is $P(A|\Theta = 2)$. The latter probability is the probability of A in the whole subpopulation of subjects with $\Theta = 2$.

Conditional independence of events A and B, given F, is the property that $P(AB|F) = P(A|F)P(B|F)$. This means that $P(AB|S) = P(A|S)P(B|S)$ if S is a smallest event in F. For other events S in F, the equality may not hold. For example, conditional independence given Θ means that $P(AB|\Theta = \downarrow) = P(A|\Theta = \downarrow)P(B|\Theta = \downarrow)$ but not that $P(AB|\Theta \leq \downarrow) = P(A|\Theta \leq \downarrow)P(B|\Theta \leq \downarrow)$. The formal definitions thus agree with the ordinary definitions of conditional probabilities and conditional independence.

3. Tail-measurability

We say that Θ is *tail-measurable* in $\mathbf{X}$ if

Tail-measurability

$\sigma(\Theta) \subseteq \tau(\mathbf{X})$.

That is: All events of Θ are determined by the long-run behavior of **X**. An equivalent definition (by Billingsley, 1986, Problem 13.6 and p. 184) is that for each $n \in$ IN there exists a measurable function $\mathbf{f}_n$ such that

$\Theta = \mathbf{f}_n(X_n, X_{n+1}, ...)$.

Here IN is the set op positive integers. That the $\mathbf{f}_n$ must be measurable is just a weak regularity condition that does not need attention at this point. The essential restriction is that Θ is fully determined by the manifest variables, even when an arbitrary finite subset of manifest variables is ignored. An example of a tail-measurable Θ would be a model where in fact $\Theta = \lim_{n->\infty} \Sigma_{i=1}^{n} X_i/n$.

That some X_i may be ignored is important if we assume that each investigator of the universe "samples" somehow a finite subsequence from **X**, and that the sample may be increased at will. This idea is underlying, for instance, alpha factor analysis (Kaiser & Caffrey, 1965). In sampling more and more manifest variables from **X** an investigator might eventually still "miss" some X_i. Thus the investigator can, in the limit, arrive at an infinite sample that is only a subsequence of **X**. It is reasonable to assume that the investigator omits at most a finite number of X_i's, because omission of an infinite subsequence would indicate a systematic bias in the sampling process (e.g., omission of "rotation items of the C-type" if the universe contains A, B, C, and D-type items). It is possible that the investigator obtains the X_i in an order that is different from how they arranged in **X**, but that does not matter - as will be shown in section 9. Tail-measurability now means that the investigator will obtain the same Θ as when he had used the whole universe **X**. Thus, although some X_i are omitted, no information about Θ is lost in the limit.

The idea of tail-measurability may be explained further by its relation to the existence of consistent estimators for Θ. We say that Θ *can be estimated consistently from* **X** if there exist measurable functions $\mathbf{f}_n$ such that

$\mathbf{f}_n(X_1, ..., X_n) -> \Theta$.

(almost surely, that is: $P\{v: \lim_{n->\infty} \mathbf{f}_n(X_1(v),...,X_n(v)) = \Theta(v)\} = 1$). Here, $\mathbf{f}_n(X_1,...,X_n)$ can be seen as a statistic based on a finite "sample" from the universe **X**. In the example $\Theta = \lim_{n->\infty} \Sigma_{i=1}^{n} X_i/n$, the finite score $\Sigma_{i=1}^{n} X_i/n$ would be such a statistic. More generally, if **X** is an item pool then $\mathbf{f}_n(X_1,...,X_n)$ is for each subject a profile of subtest scores based on the first n items from **X**. The above condition merely expresses that consistent estimators for Θ exist, not that they are known. The next theorem relates consistent estimation to tail-measurability. This makes clear why it it is reasonable to impose the constraint of tail-measurability to latent variable models. Such a Θ can still be estimated if, for example, not all items of a pool are selected in an adaptive test, or if not all judges are available in a rating task.

Theorem 1. Θ is tail-measurable in **X** if and only if Θ can be estimated consistently from every tail $(X_n, X_{n+1},...)$ of **X**.

Proof. By Lemma 12 and the definition of $\tau(\mathbf{X})$. [ ]

This shows that tail-measurability is a reasonable restriction: Θ should not contain information that is not accessible by **X**; e.g., Θ should not have a math and a reading component when **X** measures only math. Some examples of tail-measurability are: (1) If each X_i is binary and **X** satisfies the Rasch (1960) model with Θ, then various consistent estimators for Θ exist that are a function of $\Sigma_{i=1}^{n} X_i/n$ (e.g., Warm, 1989); so Θ is tail-measurable. (2) If **X** satisfies the model of essential unidimensionality and LAD with respect to Θ (Junker, 1991; Stout, 1990) then Θ is tail-measurable. (3) In item response models satisfying the regularity conditions of Junker (1991, Theorem 4.1), and either conditional independence or essential independence, the MLE based on any tail $(X_n,X_{n+1},...)$ is consistent for Θ, so Θ is tail-measurable.

4. Tail-conditional Independence

For any random vector Θ, we say that **X** is *conditionally independent* given Θ if

$$P(\cap_{i\in K}[X_i \leq x_i]|\Theta) = \Pi_{i\in K}P[X_i \leq x_i|\Theta]$$

for every set of real numbers x_i and every finite subset K of IN, where IN is the set of positive integers. This condition plays an important role in latent variable models because it means that all association in **X** can be explained by Θ (Lazarsfeld & Henry, 1968). All variation in **X** that is not attributable to the regression on Θ can then be seen as "noise" (e.g., independent errors of measurement or unique factors).

We say that **X** is *tail-conditionally independent* if **X** is independent given $\tau(\mathbf{X})$, i.e., given the long-run information of **X**. That is, replace Θ by $\tau(\mathbf{X})$ in the definition above of conditional independence (conditional independence given a σ-field was discussed in section 2). Remember that this means that **X** is independent given the "smallest" events of $\tau(\mathbf{X})$ and that **X** need not be independent given larger events. Tail-conditional independence is equivalent to the condition that

$(X_1,...,X_n)$ is conditionally independent given $(X_m,X_{m+1},...)$ whenever $n < m$

(by Lemma 10). In other words: Every finite subset of manifest variables is conditionally independent given any tail of the other manifest variables. Unlike conditional independence, the condition of tail-conditional independence does not involve any latent variable (Θ). In this sense, tail-conditional independence is completely manifest in nature; it is fully determined by the manifest variables alone. On the other hand, one might argue that tail-conditional independence is not really an "empirical" condition since it requires infinite subsequences of manifest variables. Thus it cannot be tested directly by the above definition. Therefore we will now show that independence will show up *approximately* in every finite subsequence that is sufficiently large.

The idea of tail-conditional independence can be explained further by the following equivalent concept that does not involve the rather opaque σ-field $\tau(\mathbf{X})$. Remember that a finite subsequence of **X** is independent iff for every partition (**Y**,**Z**) of the subsequence, and every two bounded measurable functions f and g, Cov(f(**Y**), g(**Z**)) = 0. In fact, it is

sufficient to require this for binary coordinatewise nondecreasing functions f and g only (Lemma 9; see also Esary, Proschan & Walkup, 1967). We say that **X** has *vanishing conditional dependence* if for every $n \in \mathbb{N}$, every partition (**Y**,**Z**) of $(X_1,...,X_n)$, and every two bounded measurable functions f and g,

$$\mathrm{Cov}(f(\mathbf{Y}), g(\mathbf{Z})|X_{n+1},...,X_m) \to 0 \quad \text{as } m \to \infty.$$

Again, it is sufficient to require this for binary coordinatewise nondecreasing f and g only. Thus the association in any finite subset of manifest variables should vanish with conditioning upon an increasing number of the other manifest variables.

Theorem 2. **X** is tail-conditionally independent if and only if **X** has vanishing conditional dependence.

Proof. See Lemma 10. [ ]

The conclusion is that, under tail-conditional independence, the association in any finite subsequence can be made arbitrarily small by conditioning on another finite subsequence that is sufficiently large. Thus, two investigators of the same universe, who both decide to assume tail-conditional independence if the conditional covariances drop below a certain bound, will arrive at the same conclusion if their sample from the universe becomes large enough. These investigators may nonetheless utilize different models and hold different assumptions with respect to conditional independence given the latent variable. In this sense, tail-conditional independence is an "asymptotically empirical" condition. We agree that this does not make tail-conditional independence a "completely empirical" condition of the same quality as, say, the condition $\mathrm{Cov}(X_2,X_3) = 0$.

Vanishing conditional dependence is such a natural condition that one may ask whether it is not always satisfied. Here is a counter-example: Suppose that $X_{2i} = X_{2i+1}$ for all i, and that the variables X_{2i} are independent of each other, and that $\mathrm{Var}(X_i) = c > 0$. The independence assures that the conditional covariance between X_{2i} and X_{2i+1} given any

sequence of other manifest variables is still c; so vanishing conditional dependence does not hold. Another example is the following: Suppose that the X_i satisfy the 2-parameter logistic model and that the discrimination parameter goes to zero if $i \to \infty$. In this case, it is possible that every possible test score variable becomes less and less reliable as estimator for Θ. The items simply do not contain sufficient information to estimate Θ consistently. Consequently, the association between X_1 and X_2 will not vanish with conditioning on successively more of the other items. This example illustrates that tail-conditional independence just means that the data in the limit provide the needed information to achieve conditional independence. Note that tail-conditional independence does not imply unidimensionality.

5. The Definition Of True-Scores

In this section we consider how a true-score theory can be founded on basis of the concepts developed here. True-score theory is often associated with a set of conventional assumptions and a certain view of test analysis (e.g., Fischer, 1974). None of this is assumed here, however, unless it is stated explicitly. Throughout this section we assume that $E(|X_i|) < \infty$ to assure the existence of the relevant true-score variables.

In true-score theory each manifest variable X_i is decomposed as

$$X_i = T_i + E_i$$

where E_i is a random variable that reflects errors of measurement and T_i is called a *true-score variable*. Different conceptualizations of true-score theory exist, depending on the definition of T_i (Lord and Novick, 1968, pp. 27-44). The two most important contemporary true-score concepts are of the form $T_i = E(X_i|F)$, with different choices for the conditioning σ-field F:

1. *The "stochastic-subject" true-scores* (e.g., Ellis & Van den Wollenberg, 1993; Lord & Novick, 1968, pp. 27-44). Assume the stochastic subject formulation and let V be a random variable that indicates the

persons. Then the stochastic-subject true-score variable T_i is defined to be the conditional expectation of X_i given V:

$$T_i = E(X_i|V).$$

This means that $T_i(v)$ is the expectation over occasions of X_i for the person with V = v.

2. *The "model-based" true-scores* (e.g., Stout, 1990). These true-scores may be defined in either the random sampling or the stochastic subject formulation. Assume that **X** satisfies some latent variable model with latent random vector Θ. Then the model-based true-score variable T_i is defined to be the conditional expectation of X_i given Θ:

$$T_i = E(X_i|\Theta).$$

If X_i is a binary item than T_i is merely the Item Response Function. To these definitions we add:

3. *The "tail-conditional" true-scores.* These also may be defined in either the random sampling or the stochastic subject formulation. Define

$$T_i = E(X_i|\tau(\mathbf{X})).$$

In the random sampling formulation, for instance, this means that the true-score of a person v, $T_i(v)$, is equal to the expectation of X_i in the subpopulation of persons that have comparable long-run behavior on **X**. The following theorem gives an equivalent definition and makes clear that, under tail-conditional independence, T_i can be approximated by the regression of X_i on a *finite* number of other manifest variables in **X**.

Theorem 3. For tail-conditional true-score variables T_i,

$$T_i = \lim_{n->\infty} E(X_i|X_n,X_{n+1},...).$$

If **X** is tail-conditionally independent then for all n > i,

Tail-measurability

$$T_i = E(X_i|X_n,X_{n+1},...) = \lim_{m->\infty} E(X_i|X_n,...,X_m).$$

Proof. The first and last equation follow by the backward and forward martingale convergence theorems, respectively (Billingsley, 1986, Theorem 35.7 and 35.5). The second equation holds because, under tail-conditional independence, (X_i,X_n) is conditionally independent given $(X_{n+1},X_{n+2},...)$ by Lemma 10; so $E(X_i|X_n,X_{n+1},...) = E(X_i|X_{n+1},X_{n+2},...)$. [ ]

It is worth noting that all reliability theory results in Lord and Novick (1968) remain valid if tail-conditional true-scores are used and experimental independence (uncorrelated error-variables) is replaced by tail-conditional independence. This is so because, for tail-conditional true-scores T_i and errors E_i defined by $X_i = T_i + E_i$, the "assumptions" of classical test theory hold: $E(E_i) = 0$; $Cov(T_i,E_i) = 0$; $Cov(T_i,E_j) = 0$; and, if tail-conditional independence holds, $Cov(E_i,E_j) = 0$ for i and j distinct.

It is worthwhile to compare the concept of tail-conditional true-scores with the other two. Consider the stochastic-subject true-scores first. The classical definition involves a hypothetical experiment where one could "brainwash" the persons and administer the test under identical circumstances again and again (Lazarsfeld, 1959; Lord & Novick, 1968). This would yield a series of *replicate measurements* X_i that are i.i.d. given V. The true-score of a subject is then defined to be the limiting value of the subject's average testscore, $\lim_{n->\infty}\Sigma X_i/n$. With Theorem 4 of the next section it can be shown that this true-score may be seen as a particular tail-conditional true-score, where the universe **X** consist of replicate measurements in the hypothetical experiment. The definition of tail-conditional true-scores is more flexible in that it may also be based on any other sequence of manifest variables. As Lord and Novick (p. 45, pp. 38-39) point out, stochastic subject true-scores depend on how the replication experiment is defined. Similarly, tail-conditional true-scores depend on how the universe **X** is defined.

Next, consider model-based true-scores. Here one may ask the question to which extent these true-scores are uniquely defined. It is possible that **X** simultaneously satisfies two different models with different latent random vectors Θ and Θ' that are not a 1-to-1 transformation of each other

(see the example in the next section preceding Corollary 2). Then the true-scores based on the one model will be different from those based on the other model. Corollary 3 of the next section, however, restricts this possibility considerably. It is concerned with the fairly broad class of models that satisfy both tail-measurability and conditional independence. The Corollary implies that two different models of this kind will yield the same true-scores, which are then equal to the tail-conditional true-scores. Thus, tail-conditional true-scores specialize to model-based true-scores whenever the model satisfies tail-measurability and conditional independence. Note that the concept of tail-conditional true-scores is more flexible in that it does not require the a priori specification of a model.

As noted above, tail-conditional true-scores also have a certain lack of unicity. A given manifest variable X_i that is element of the sequence **X** might equally well be embedded in some other sequence of manifest variables **X**' (Mulaik & McDonald, 1978). For instance, the item "Is it important for you to be thin" may be extended to a domain of items involving eating disorders, or a domain of items involving social desirability, or a domain of items involving fysical health. The tail-conditional true-scores based on one domain will be different from those based on an other domain, in general. Thus the definition of tail-conditional true-scores is unique only in the context of a well-defined universe **X** of admissible manifest variables (compare Cronbach et al., 1972, Ch. 1).

There is another kind of true-scores that we did not discuss here, namely the true-scores as defined in domain-sampling theory (Lord & Novick, 1968, Ch. 8). Here it is assumed that the variables X_i are randomly selected from a *domain*. The true-score of a subject is then defined as the expectation over all X_i in the domain. This will be called the *generic true-score* in the sequel. Lord & Novick discuss only parallel test-forms, but the definition of generic true-scores can of course be used anyhow - although it will not make much sense if the domain is multi-dimensional. The domain is comparable to the universe in our paper. The difference is that we do not explicitely assume a probability measure on the universe, which would be needed to define the generic true-score. Instead, we consider the variables X_i in a fixed, artificial order. Our

universe **X** may thus be seen as only one possible outcome of a domain-sampling procedure of infinite length (this distinction was also made by McDonald, 1977). In section 9 we will show that the particular order of the X_i does not affect any of our results. Although a domain is very similar to a universe, generic true-scores are quite different from tail-conditional true-scores. Given **X** and some probability measure on **X**, there is only a single generic true-score variable, while there is a different tail-conditional true-score variable for each X_i. Thus a generic true-score variable should rather be compared with Θ or a transformation of Θ. An important distinction then is that the present paper allows Θ to be multi-dimensonal, while the generic true-score variable is unidimensional by definition.

It is possible to define for each X_i the specific domain-sampling true-score variable as the conditional expectation of X_i given the generic true score. Are these comparable to tail-conditional true-score variables? We can distinguish three cases here. (1) If **X** does not satisfy the monotone LV model: Then the specific domain-sampling true-scores will in general be different from the tail-conditional true-scores. This is so because the generic true-score variable will fail to differentiate between subjects with systematically different long-run behavior. (2) If **X** satisfies the monotone LV model, but only with a Θ that is not tail-measurable: Then the opposite pattern is possible, namely that the generic true-score, which must be a monotone transformation of Θ, entails a differentiation that finds no basis in the manifest variables and that cannot be matched by $\tau(\mathbf{X})$. Again, the tail-conditional true-scores will not be identical to the specific domain-sampling true-scores. (3) If **X** satisfies the monotone LV model with a tail-measurable Θ: Then the generic true-score variable is a strictly monotone transformation of Θ and it will thus entail specific true-score variables that are identical to the tail-conditional true-score variables.

6. Tail-measurability and Conditional Independence

This section discusses a fundamental theorem that has interesting consequences for latent variable models *in general* - satisfying conditional independence and tail-measurability, but not necessarily

unidimensionality and monotonicity. This theorem forms a crucial step in the proof of our characterization theorems. This section is not needed to understand the meaning of those characterizations, but it may give more insight.

Here, it is necessary to introduce a more finely detailed set of true-scores than we have discussed in the previous section. The construction here is analogous to the practice of "dichotomizing" a continuous or polytomous variable to produce binary responses. Instead of one such dichotomization, though, we consider true-scores based on all possible dichotomizations. Also, the true-scores will initially be defined in an abstract form. To be exact, we define the true-score variables T_{ix} as

$$T_{ix} = P[X_i > x|F]$$

for some σ-field F that we do not specify at this point. If F is $\sigma(\Theta)$, these are the model-based true-score variables for the dichotomized manifest variables. If F is $\tau(\mathbf{X})$, these are the tail-conditional true-score variables for the dichotomized manifest variables. Next, define $\mathbf{T}_i = (T_{ix} : x \in \mathrm{IR})$, where IR is the set of real numbers. From $\mathbf{T}_i$ one can deduce the entire conditional distribution of X_i given F. Finally, define $\mathbf{T} = (\mathbf{T}_i : i \in \mathrm{IN})$, the collection of all T_{ix}.

Theorem 4. If $F \subseteq \tau(\mathbf{X})$ and $\mathbf{X}$ is conditionally independent given F then $F = \tau(\mathbf{X}) = \sigma(\mathbf{T}) = \tau(\mathbf{T})$.

Proof. See Lemma 3 of the appendix. [ ]

The remarkable conclusion that $F = \tau(\mathbf{X})$ means that conditional independence of $\mathbf{X}$ cannot hold given a σ-field that is strictly coarser that $\tau(\mathbf{X})$. In other words: If F contains *only* long-run events of $\mathbf{X}$ and explains *all* association in $\mathbf{X}$ (in the sense that conditional independence holds) then F must contain *all* long-run events of $\mathbf{X}$. Furthermore, under the hypothesis of this theorem, the abstract true-score variables $P[X_i > x|F]$ must in fact be the tail-conditional true-score variables $P[X_i > x|\tau(\mathbf{X})]$, i.e.,

$T_{ix} = P[X_i > x | X_n, X_{n+1}, ...]$ for $n > i$.

The conclusion that $\tau(\mathbf{X}) = \sigma(\mathbf{T})$ means that the events defined by the long-run behavior of **X** are also the events that can be described by the true-score variables in **T**. Finally, that $\sigma(\mathbf{T}) = \tau(\mathbf{T})$ means that these events can still be described by **T** if the true-score variables of some X_i are ignored. We will now consider four Corollaries of Theorem 4.

1. Tail-conditional independence.

Corollary 1. If **X** is conditionally independent given Θ and Θ is tail-measurable in **X**, then **X** is tail-conditionally independent.

Proof. With $F = \sigma(\Theta)$, Theorem 4 implies that $\sigma(\Theta) = \tau(\mathbf{X})$. So conditioning on Θ is the same as conditioning on $\tau(\mathbf{X})$. [ ]

Corollary 1 is important because it means that in the broad class of models that satisfy conditional independence and tail-measurability, **X** has to be tail-conditionally independent (and hence have vanishing conditional dependence by Theorem 2), regardless of other model restrictions such as dimensionality and monotonicity.

Suppes and Zanotti (1981) showed that, for a finite number of manifest variables with finite range there always exists a random variable Θ such that the manifest variables are conditionally independent given Θ. The Θ constructed in their proof is in fact manifest, since it encodes the total response pattern on **X**. The proof can be extended easily to infinitely many manifest variables (for binary X_i, define Θ as the real number for which **X** is the binary digit expansion). The Θ constructed by Suppes and Zanotti will not satisfy tail-measurability, however, because it requires knowledge from *every* X_i; no X_i may be omitted. Our result may be seen as a bound to the result of Suppes and Zanotti. While Suppes and Zanotti showed that one can always define a random variable Θ that satisfies unidimensionality and conditional independence, we show that it is not always possible to find a random vector Θ that is, in addition, tail-measurable. Thus, unidimensionality and conditional independence are

vacuous conditions if they are not accompanied by other conditions such as monotonicity (e.g., Holland & Rosenbaum, 1986). Unidimensionality, conditional independence and *tail-measurability* are not vacuous.

An intriguing question is how one can define a general concept of latent dimensionality with as little assumptions as possible. Our conclusion is that these three restrictions: Unidimensionality, conditional independence and tail-measurability, are sufficient to distinguish universes **X** that can be described by a unidimensional random variable Θ from universes **X** that require a multi-dimensional random vector Θ. This conclusion may be refined: Whenever **X** satifies the conditions of unidimensionality, conditional independence and tail-measurability with respect to a finite-dimensional Θ, **X** will also satisfy them with respect to some unidimensional Θ (it is a standard exercise in mathematical probability theory that the sigma-field of any finite-dimensional random vector can also be generated by some unidimensional variable). Therefore, the more proper conclusion is: The conditions of unidimensionality, conditional independence and tail-measurability are sufficient to distinguish universes **X** that can be described by a finite-dimensional Θ from universes **X** that require an infinite-dimensional Θ.

2. Uniqueness of latent variables. Tail-measurability and conditional independence have consequences for the uniqueness of latent random vectors for **X**. In general, even under such restrictive models as the 1-parameter logistic model or the linear 1-factor model, if a finite set of manifest variables satisfies the model with both Θ and Θ', Θ and Θ' need not be functions of each other (compare the factor indeterminacy problem: McDonald, 1977; Mulaik & McDonald, 1978; McDonald & Mulaik, 1979; Rozenboom, 1988; Steiger, 1979). Acknowledgement of this fact is crucial to appreciate our claims below, regarding the unicity of Θ. We present an example using the random sampling formulation.

Insert Table 1 about here

Example. Assume that the data in Table 1 form the population and that persons are sampled randomly. It is easily verified that

1. $P[X_i = 1|\Theta] = \Theta/(1 + \Theta)$ for all i, and (X_1,X_2,X_3,X_4) is conditionally independent given Θ; and

2. $P[X_i = 1|\Theta'] = \Theta'/(1 + \Theta')$ for all i, and (X_1,X_2,X_3,X_4) is conditionally independent given Θ'

(for instance, there are 256 persons with $\Theta = 1/3$, in the left half of the Table, of which 64 have $X_1 = 1$; so $P[X_1 = 1|\Theta = 1/3] = 64/256 = 1/4$). Thus (X_1,X_2,X_3,X_4) satisfies the 1-parameter logistic model with both Θ and Θ'. However, Θ is not a function of Θ'. The clue here is that, among those persons with response pattern 0011, some have a Θ of 1/3 while others have a Θ of 3. But one can as well postulate a latent variable Θ' for which the reverse is true.

Even for an infinite set of manifest variables, Θ and Θ' need not be functions of each other if, for example, the regressions of the manifest variables on Θ are all flat in a fixed open interval of the Θ scale. Tail-measurability helps to exclude such pathological cases from occuring: The next Corollary says that tail-measurability and conditional independence jointly determine Θ unique up to some measurable transformation - regardless of other restrictions imposed to the model. In other words, then Θ *is of at least nominal level of measurement*. This result can be seen as a generalization of the uniqueness results of McDonald (1977) and Mulaik & McDonald (1978) for linear factor models.

Corollary 2. Suppose Θ is tail-measurable in $\mathbf{X}$ and $\mathbf{X}$ is conditionally independent given Θ. If Θ' is another random vector (not necessarily of the same dimensionality) that satisfies these conditions, then there exists a measurable function $\mathbf{f}$ such that $\Theta' = \mathbf{f}(\Theta)$.

Proof. Theorem 4 implies that $\sigma(\Theta) = \sigma(\Theta')$, since both must be equal to $\tau(\mathbf{X})$. From this the conclusion can be shown (Billingsley, 1986, Problem 13.6.). [ ]

3. Uniqueness of model-based true-score variables.

Corollary 3. Suppose **X** is conditionally independent given Θ and Θ is tail-measurable in **X**. Then the conditional distribution of **X** given Θ is identical to the conditional distribution of **X** given $\tau(\mathbf{X})$.

Proof. Similar to Corollary 1. [ ]

Together with Theorem 3 this implies that in a tail-measurable conditionally independent model for **X**, the "latent" conditional probabilities $P[X_i > x|\Theta]$ are always equal to the "manifest" conditional probabilities $P[X_i > x|X_n,X_{n+1},...]$ and hence can be approximated by $P[X_i > x|X_n,...,X_m]$ as $m \to \infty$ and $n > i$ (compare Junker, 1993, Lemma 3.1). Furthermore, all such models for **X** generate the same model-based true-score variables and these are equal to the tail-conditional true-score variables.

4. Functional relationship between true-score variables.

Corollary 4. For tail-conditional true-score variables and all i:

1. If **X** is tail-conditionally independent, there exists a measurable function $\mathbf{f}_i$ such that $\mathbf{T}_i = \mathbf{f}_i(\mathbf{T}_j : j > i)$.

2. If Θ is tail-measurable in **X** and **X** is conditionally independent given Θ, there exists a measurable function $\mathbf{f}_i$ such that $\Theta = \mathbf{f}_i(\mathbf{T}_j : j > i)$.

Proof. 4.1: Put $F = \tau(\mathbf{X})$; then Theorem 4 implies that $\sigma(\mathbf{T}) = \tau(\mathbf{T})$. This implies that $\sigma(\mathbf{T}) = \sigma(T_{jy} : j > i, y \in \mathrm{IR})$, where IR is the set of real numbers. From this the first conclusion can be derived (Billingsley, 1986, Problem 36.3.). *4.2:* Analogously, using the fact that $\sigma(\Theta) = \tau(\mathbf{T})$ by Theorem 4. [ ]

Corollary 4 again and more exactly states the redundancy of true-score variables, pointed out under Theorem 4, and extends it to Θ. This can be used to illustrate what tail-measurability means in terms of the "response functions" $P[X_i > x|\Theta]$. In a tail-measurable conditionally independent

model, $T_{ix} = P[X_i > x|\Theta]$. Corollary 4.2 now implies that, for each open region of Θ, there must be infinitely many response functions $P[X_i > x_i|\Theta]$ that are not flat in that region.

7. Characterization Of The Monotone LV Model Using Conditional Association.

For any given unidimensional random variable Θ, we say that **X** is *monotone in* Θ if

$$P[X_i > x|\Theta]$$

is a nondecreasing function of Θ for each positive integer i and each real number x. This condition means that each X_i is stochastically increasing in Θ (Lehmann, 1986). If there exists a (usually latent) unidimensional random variable Θ such that **X** is conditionally independent given Θ and **X** is monotone in Θ, then we say that **X** satisfies the *monotone latent variable (LV) model* (with Θ). For more information about this model, see e.g. Ellis & Van den Wollenberg, 1993; Grayson, 1988; Holland & Rosenbaum, 1986; Junker, 1993; Mokken, 1971; Ramsay, 1991; Sijtsma, 1988.

Consider an important manifest condition implied by the model first (see Holland & Rosenbaum, 1986). Following Holland and Rosenbaum (1986), we say that a finite subsequence **X'** of **X** is *conditionally associated* if for every partition (**Y**,**Z**) of **X'**, every two bounded coordinatewise nondecreasing functions f and g, and every measurable vector-valued function **h**,

$$Cov(f(\mathbf{Y}), g(\mathbf{Y})|\mathbf{h}(\mathbf{Z})) \geq 0.$$

In fact, it is sufficient to require this for *binary* coordinatewise nondecreasing functions f and g only (Esary, Proschan & Walkup, 1967). We say that **X** is conditionally associated if every finite subsequence of **X** is (Junker, 1993). If **X** satisfies the monotone LV model, then **X** is conditionally associated (Holland & Rosenbaum, Theorem 6).

Theorem 5. The following conditions 5.1 and 5.2 are equivalent:

Tail-measurability

1. There exists a unidimensional random variable Θ that is tail-measurable in **X** and with which **X** satisfies the monotone LV model.

2. **X** is conditionally associated and has vanishing conditional dependence.

Uniqueness of Θ: If Θ' is another random variable that is tail-measurable in **X** and with which **X** satisfies the monotone LV model, then $\Theta' = f(\Theta)$ a.s. for some *strictly increasing* function f.

Proof. 5.1 implies 5.2: Conditional association follows by Theorem 6 of Holland & Rosenbaum (1986). Vanishing conditional dependence follows by Corollary 1 and Theorem 2.

5.2 implies 5.1: Tail-conditional independence must hold by Theorem 2. By Lemma 11, **X** must be associated given any nonnull event of $\tau(\mathbf{X})$ (see below). In Lemma 4 it is shown that these two conditions together imply that every pair (T_{ix}, T_{jy}) is comonotone, i.e., $T_{ix}(v) < T_{ix}(w)$ implies $T_{jy}(v) \leq T_{jy}(w)$. Next, in Lemma 5 it is shown that Θ may be constructed as a positively weighted sum of the T_{ix} with x rational, and that this Θ satisfies the required conditions.

Uniqueness: This is shown in Lemma 7. [ ]

Remarks. Conditional association in 5.2 may in fact be replaced by the weaker condition of conditional positive quadrant dependence (see proof of Lemma 4). It is remarkable that (T_{ix}, T_{jy}) is comonotone *even when* the true-scores are derived from dichotomizations of the same manifest variable (i = j): Two such dichotomizations are conditionally dependent, so their association says little about the underlying true-score variables. The crucial conclusion that they are nevertheless comonotone is obtained in Lemma 4 on basis of the redundancy of true-score variables, stated in Corollary 4.1. The construction of Θ as a positively weighted sum of true-score variables agrees with the more general observation under Corollary 4.2 that Θ must be a function of the true-score variables. The general - though weak - form of uniqueness stated in Corollary 2 is used to prove the specific and strong uniqueness result of

Lemma 7.

Theorem 5 provides necessary and sufficient conditions for the monotone LV model in 5.1 by the manifest conditions in 5.2. Thus it explains the empirical meaning of the model. In this sense it is comparable to the representation theorems in axiomatic measurement theory. A difference is that the manifest conditions are statistical in nature here. The main point of Theorem 5, that the model in 5.1 is equivalent to the manifest conditions in 5.2, answers for a very general class of manifest variables (binary, discrete, continuous or any mixture of these) a question posed by Holland and Rosenbaum (1986, p. 1534): Every sequence of manifest variables **X** that satisfies conditional association, and the additional requirement of vanishing conditional dependence, can be given a representation in the monotone LV model - with a Θ that is tail-measurable. Conversely, whenever such a representation exists, **X** must satisfy conditional association and vanishing conditional dependence. Junker (1993, Theorem 5.2) provides a partial characterization that is essentially a special case of Theorem 5.

The uniqueness result at the end of Theorem 5 is comparable to Stout's uniqueness result under essential unidimensionality and LAD (Stout, 1990, Theorem 3.3; Junker, 1991). It means that Θ *is of at least ordinal level of measurement*. Classical versions of this result, restricted to linear factor models, were obtained by McDonald (1977) and Mulaik & McDonald (1978).

Note that the empirical conditions in 5.2 imply tail-measurability of Θ here. This is mainly a consequence of tail-conditional independence. If Θ were not tail-measurable, then the variance of Θ given $(X_n,...,X_m)$ would not tend to 0 everywhere, and then the conditional covariances given $(X_n,...,X_m)$ would not all tend to 0 as $m \to \infty$. Thus Corollary 1 has a partial converse.

Theorem 5 can also be compared with de Finetti's characterization (e.g., Kingman, 1978) of *exchangeable* random variables as being manifest variables that are conditionally independent and identically distributed given a tail-measurable latent variable. The most familiar example of this in psychometrics is the beta-binomial true-score model (e.g., Lord & Novick, 1968). The monotone LV model generalizes the beta-binomial true-score model, and the restrictions in 5.2 generalize exchangeability. We

explore this connection more fully elsewhere (Junker & Ellis, 1995).

8. Characterization Of The Monotone LV Model Using Subpopulation Invariance.

Theorem 5 can be given a slightly different formulation that illuminates more of the theoretical structure of the monotone LV model. For this purpose we adopt the following definitions:

A *subpopulation* is an event in $\tau(\mathbf{X})$;
a *trait vector* is a random vector that is tail-measurable in $\mathbf{X}$;
experimental independence is another term for tail-conditional independence; and
local homogeneity means that the conditional distributions of $\mathbf{X}|\Theta$ and $\mathbf{X}|\tau(\mathbf{X})$ are identical (a finer trait than Θ will not yield a different conditional distribution).

We chose these definitions because they specialize to those of Ellis and Van den Wollenberg (1993) in the stochastic subject formulation if one assumes $\tau(\mathbf{X}) = \sigma(V)$, where V is a random variable that indicates the persons. The present definitions have the advantage that they are also very reasonable in the random sampling formulation. In particular, a subpopulation here is more than just an arbitrary event that can be selected on $\mathbf{X}$; a subpopulation is defined by the long-run behavior of $\mathbf{X}$ and does not depend on any finite number of manifest variables (which may involve non-generalizable "noise").

If A is a given non-null event, we say that $\mathbf{X}$ is *associated given* A (or *in* A) if for every finite subsequence $\mathbf{Y}$ and every two bounded coordinatewise nondecreasing functions f and g,

$$\mathrm{Cov}(f(\mathbf{Y}),g(\mathbf{Y})|A) \geq 0.$$

The difference with "conditional association" is that the event A is fixed and need not be determined by a *finite* subsequence $\mathbf{Z}$. In particular, A may be a subpopulation.

Note that with the above definitions, *conditional independence of* $\mathbf{X}$

given a trait vector Θ *implies local homogeneity*. This is so because $\sigma(\Theta)$ must be equal to $\tau(\mathbf{X})$, by Theorem 4. Using this and Lemma 11, Theorem 5 may be rephrased in a form that generalizes the characterization theorem of Ellis and Van den Wollenberg (1993) in that it presupposes neither stochastic subjects nor binary manifest variables:

> *Theorem 6*. There exists a unidimensional trait with which **X** satisfies the monotone LV model (and then local homogeneity holds) if and only if **X** is associated in every non-null subpopulation and **X** is experimentally independent.

The difference with Theorem 5 is that here, we consider association in *subpopulations* instead of events determined by finite subsequences. Such events are used only to approximate subpopulations. Consequently, not all events determined by finite subsequences are actually needed in the conditional association part of Theorem 5; only those events that approximate subpopulations are needed. All subpopulations are determined exactly by the tail-conditional true-score variables (Theorem 4) and thus can be approximated by finite regressions (Theorem 3). In summary, Theorem 6 is a more "narrow" - though less "empirical" - characterization than Theorem 5.

Subpopulation Invariance And The Interpretation of Models

The concept of subpopulation invariance is closely related to the substantive interpretation that is given to models. This less formal matter will be discussed now.

In the preceding sections we made a distinction between the stochastic subject interpretation and the random sampling interpretation, following Holland (1990). This distinction may be more than an arbitrary way of formalizing a latent variable model. In experimental psychology, data are frequently obtained and analyzed for each subject separately. This means that the replication experiment underlying the stochastic subject interpretation is actually performed (at least, the experimentator believes this). An intriguing question is wether the results from such an experiment are compatible with results from, say, a survey study where variables that are similar in content are measured only once and where

conditional probabilities are estimated by a latent variable model instead of replications. This is evidently not necessarily the case. For instance, if one starts with 10 items that represent different grey shades then the experimentator will extend these variables with replications of the same shades, while the surveyer may extend these variables with different grey shades or even different colors. The experimentator and the surveyer thus consider different universes, and may consequently need different models. The two methods will yield identical true-score variables only if the latent variable model of the survey method holds in the stochastic subject interpretation too.

Following existing ideas about factorial invariance (e.g., Byrne, 1989, Grayson, 1990; Thurstone, 1957), item bias (e.g., Lord, 1980; Mellenbergh, 1989), and sample-freeness (e.g., Fischer, 1974), Ellis (1993) and Ellis & Van den Wollenberg (1993) argue that a latent variable model allows a stochastic subject interpretation if and only if the model satisfies "subpopulation invariance" in the strict sense, i.e., if the basic model properties are invariant across all subsets of persons. Our present theory offers a generalization of this result: A conditionally independent latent variable model allows a latent random vector for which tail-measurability holds if and only if the model satisfies subpopulation invariance in the weak sense, i.e., if the basic model properties are invariant across all tail-events of the universe. Under any such model, the smallest subpopulations consist of subjects that have the same score on Θ and thus are indistinguishable in their long-run behavior (by Theorem 4). These smallest subpopulations may be considered as *stochastic meta-subjects*. Note that this definition is free of further assumptions such as monotonicity. Individual persons may be seen as independent replications of the stochastic meta-subject to which they belong. Consequently, tail-measurable true-scores may be seen as a particular kind of stochastic-subject true-scores, where the "subject" is in fact a subpopulation of persons that are indistinguishable in their long-run behavior.

Stochastic meta-subjects are the finest partition of persons that statistical modelling can reliably tell us about. Note that different universes may yield different meta-subjects. Thus in any particular experiment or investigation, the operationalization of a theoretical concept by means of response variables $\mathbf{X}$ determines what the meta-subjects

are in that experiment or investigation. This agrees with the intuition that, for instance, in a perception experiment one may conceive persons as "equivalent" who are perhaps not conceived so in some clinical psychology experiment.

Our conclusion is that subpopulation invariance, in either the strict of the weak sense, forms an important aid in the substantive interpretation of latent variable models. If subpopulation invariance holds in the strict sense, then a stochastic subject interpretation is allowed. If subpopulation invariance holds in the weak sense then a stochastic meta-subject interpretation is allowed.

9. Extension To Permutations And Infinite Subsequences

The theory as it is formulated up to here pertains to a sequence **X** wherein the variables X_i are arranged in a particular order. In many applications, however, there is no natural order of manifest variables. Would a permutation of the X_i change the meaning of the concepts tail-measurability and tail-conditional independence? Lemma 13 in the appendix states that if **X**' is a (possibly infinite) permutation of **X**, then $\tau(\mathbf{X}') = \tau(\mathbf{X})$. Therefore, *all concepts and theorems in the preceding theory are independent of the order in which the manifest variables are arranged.*

The central concepts in the theory we have developed allow that any finite subsequence of **X** may be ignored. This is important because in sampling manifest variables randomly from **X** one might "miss" some X_i. What happens if the sampling process is biased and an *infinite* subsequence of **X** is missed? In general, the theory developed will not apply anymore then: If Θ is tail-measurable in **X** and **X**' is an infinite subsequence of **X** that misses infinitely many X_i, then Θ need not be tail-measurable in **X**'. In other words, *the properties of tail-measurability and tail-conditional independence are not inherited by all infinite subseqences of* **X**.

It is possible to strengthen the theory such that the main concepts are inherited by all infinite subsequences of **X** whenever they hold for **X**. We will first consider three reasons why this may be worthwhile; next, we will develop the stronger theory.

1. In tail-measurability it is assumed that, in adding more and more

manifest variables, all manifest variables $X_n, X_{n+1}, \ldots$ from a certain n on will be "sampled" from the universe. However, if the sampling process is such that one can miss X_3, why would one not miss X_{n+1}? Or, indeed, why not an infinite subsequence? Adding more manifest variables will eventually lead to an infinite subsequence, but there is little reason to believe - in general - that this subsequence cannot contain infinitely many "gaps". So if the conditional covariances do not tend to zero as the number of manifest variables increases, it is not clear whether this is due to a failure of vanishing conditional dependence or to gaps in the conditioning sequence.

2. Tail-measurability allows certain pathological cases that one may wish to exclude from occuring. For instance, suppose that the even manifest variables in **X** satisfy the 1-parameter logistic model with Θ, while the odd manifest variables are independent Bernouilli variables resulting from coin flips. Then **X** satisfies the tail-measurable monotone LV model. However, the odd variables have a negative contribution to the reliability of estimates for Θ whenever they are used. Similarly, if the odd variables satisfy the 2-parameter logistic model with Θ while their discrimination parameters tend to zero, they have a negative contribution to the reliability of $\Sigma_{i=1}^{n} X_i$ from a certain n on; then they add more error than true-score variance. Although such variables are not inconsistent with the tail-measurable monotone LV model, there seems little reason to include them. Indeed, such variables would be deleted in classical test analysis and presumably in Mokken's (1971) scale analysis.

3. The stronger theory leads to a generalization of the idea of "specific objectivity" (Rasch, 1977; Fischer, 1987) - as will be discussed below.

We will develop the stronger theory briefly now. For this purpose we define the σ-field $\Sigma(\mathbf{X})$ as the intersection of all σ-fields that are generated by infinite subsequences of **X**. We say that Θ is *subtail-measurable* in **X** if $\sigma(\Theta) \subseteq \Sigma(\mathbf{X})$. This means that Θ can be estimated consistently from every infinite subsequence of **X**. We say that **X** is *subtail-conditionally independent* if **X** is conditionally independent given

$\Sigma(\mathbf{X})$. This means that every finite subsequence of $\mathbf{X}$ is conditionally independent given any infinite subsequence of $\mathbf{X}$ that is disjoint from it. Some ideas of McDonald (1977) may be seen as a classical form of subtail-measurability in linear factor models. A reviewer suggested that subtail-measurability may imply a restriction on the dimensionality of Θ, but this will be explored elsewhere.

Lemma 14 states some elementary facts that imply

1. If Θ is subtail-measurable then Θ is tail-measurable.
2. Θ is subtail-measurable in $\mathbf{X}$ if and only if Θ is subtail-measurable in every infinite subsequence of $\mathbf{X}$.
3. If $\mathbf{X}$ is conditionally independent given some σ-field $F \subseteq \Sigma(\mathbf{X})$ then $F = \Sigma(\mathbf{X}) = \tau(\mathbf{X})$ (compare Theorem 4).
4. $\mathbf{X}$ is subtail-conditionally independent if and only if every infinite subsequence of $\mathbf{X}$ is subtail-conditionally independent.

The consequence of this is that all preceding theory remains true if tail-measurability is replaced by subtail-measurability, tail-conditional independence by subtail-conditional independence, etc. The ensuing stronger theory applies to every infinite subsequence of $\mathbf{X}$ whenever it applies to $\mathbf{X}$.

A special case of the subtail-measurable monotone LV model is the item response theory model of Rasch (1960). This model has received much attention because it is claimed to have the property of "specific objectivity" (Fischer, 1987, p. 567):

> "Rasch furthermore required that, based on the observed reactions, each two subjects S_v and S_w should be comparable to each other. The result of such a comparison should be unique and was termed "specifically objective" if it does not depend on the particular subset of items chosen from the universe $\{I_i\}$."

We now weaken this and require that only comparisons based on *infinite* subseqences of items do not depend on the particular subsequence chosen - which will be called *asymptotic specific objectivity* of the comparison of persons. This is of course much more realistic, since our Example above Corollary 2 shows that even the Rasch model does not really satisfy specific objectivity for a finite number of items. Indeed, derivations of

the Rasch model from specific objectivity implicitly assume infinitely many manifest variables by regarding the individual stochastic-subject response probabilities as fully identified (e.g., Fischer, p. 567; Roskam & Jansen, 1984, pp. 298-299).

The subtail-measurable monotone LV model implies asymptotic specific objectivity of the ordinal comparison of persons: Whenever two infinite subsequences $\mathbf{X}_i$ and $\mathbf{X}_j$ yield latent variables Θ_i and Θ_j, respectively, it holds that $\Theta_i = f(\Theta_j)$ for some strictly increasing function f, by Theorem 5. So the ordering of persons based on Θ_i is identical to the one based on Θ_j - thus independent of the subsequence of manifest variables.

10. Issues for Future Research

Our work answers many important questions, which makes it possible to pose even more new questions:

Consistent estimators. Tail-measurability of Θ assures that consistent estimators exist. Another question is *which* estimators are consistent. Stout (1990) and Junker (1991) show that consistent estimators can be based on the mean score $\Sigma_{i=1}^{n} X_i/n$ if $\mathbf{X}$ satisfies the model of essential unidimensionality and LAD. This suggests that the easiest way to obtain consistent estimates is to order the manifest variables such that the sequence is LAD, e.g., by adding new items only if they increase the psychometric reliability of $\Sigma_{i=1}^{n} X_i/n$.

Conditioning on estimators. Assessment of tail-conditional independence requires determination of conditional covariances given $(X_{n+1},...,X_m)$ - which is basically used to approximate Θ. An interesting question therefore is whether $(X_{n+1},...,X_m)$ can be replaced by any consistent estimator for Θ, as in Stout (1990), and whether Theorems 2 and 5 remain valid then.

Oscillation of conditional covariances. Theorem 2 allows the conditional covariances given $(X_{n+1},...,X_m)$ to oscillate as $m \to \infty$. Testing tail-conditional independence would be enhanced if it is known under which conditions all or some conditional covariances go monotonically to 0. This seems reasonable since adding more manifest variables implies that Θ can be estimated more reliably.

Finite item pools. It would be interesting to have a characterization

of the monotone LV model for finite sequences of manifest variables. Our characterization for infinite sequences may help to solve this problem. If such a characterization cannot be found, however, it may be more worthwhile to consider instead which restriction should be added to the monotone LV model in order to assure that a finite sequence can always be extended to an infinite sequence $\mathbf{X}$ that satisfies the monotone LV model with tail-measurable Θ (as in Junker, 1993).

Pairwise conditional independence. Many results in test theory (e.g., Ellis, 1993; Lord & Novick, 1968; Stout, 1990) can be derived using only pairwise conditional independence instead of full conditional independence. Perhaps this can be done for many of the results in this paper too. For this it would be needed to generalize Kolmogorov's zero-one law (see Lemma 3).

Multi-facetted universes. As remarked, $\mathbf{X}$ can be seen as a universe of admissible manifest variables, and it is treated as a one-facetted universe here (see Cronbach et al., 1972). It may be worthwhile to develop a similar theory for multi-facetted universes. A two-facetted universe would consist of manifest variables X_{nk}, where n refers to the first facet (e.g., judges) and k to the second (e.g., items). Latent variable models may then be formulated analogously to three-mode factor analysis (e.g., Tucker, 1966).

11. Discussion

In this paper, we have sketched a test theory for a general class of manifest latent variables (including binary, discrete and continuous), based on the notions of *tail-measurable* latent variables. Tail-measurability was introduced by Junker (1993, Proposition 2.1), who discussed it as a fundamental requirement of latent variable models in the context of measurement purposes. Tail-measurable latent variables may be consistently estimated from the manifest variables, and their estimation does not depend on any particular manifest variable (Theorem 1).

Tail-measurable latent variables contain *only* long-run information from the manifest variables. In a fundamental result (Theorem 4) we showed that this, together with conditional independence, implies that they must moreover embody *all* long-run information from the manifest variables. On

the basis of this result we derived four general properties that all tail-measurable, conditionally independent latent variable models must satisfy. These models may be multidimensional and non-monotone. The first property is tail-conditional independence (Corollary 1), which is equivalent to *vanishing conditional dependence* (Theorem 2): All covariances in any finite sequence of manifest variables tend to zero as one conditions on more and more of the remaining manifest variables. The second property pertains to the *uniqueness of the latent variables* (Corollary 2): All latent random vectors that fit the model are functions of each other; i.e., they have at least nominal level of measurement (an interesting example shows that this need not to be true outside the class of tail-measurable conditionally independent models). The third property is *uniqueness of the model-based true-score variables*, which must then coincide with the tail-conditional true-score variables (Corollary 3), which can be approximated by finite manifest regressions (Theorem 3). The fourth property is *redundancy of the true-score variables*: All true-score variables must be functions of each other, and the latent variables must be functions of the true-score variables (Corollary 4).

Building on the work of Ellis and Van den Wollenberg (1993), Holland and Rosenbaum (1986), and Junker (1993), we then established a manifest characterization of the tail-measurable monotone latent variable model (Theorem 5). The conditions required of the manifest variables are conditional association (Holland & Rosenbaum) and vanishing conditional dependence. These conditions *assure the existence of a tail-measurable, unidimensional latent variable such that the manifest variables satisfy monotonicity and conditional independence with respect to it*. Conversely, if such a representation holds, then the two manifest conditions hold. The latent variable is unique up to a *monotone* transformation, it is estimatable from the manifest variables, and it does not depend on what order we put the manifest variables - nor whether a finite number of them are omitted. In a second, more theoretical version of this characterization (Theorem 6) we generalized the result of Ellis and Van den Wollenberg. This Theorem can also be used to obtain a characterization of congenerity if the definitions and proofs of Ellis (1993) are adapted to the present context.

Our theory has several philosophical aspects. First, it is based on

the idea of an *infinite universe of manifest variables*. This idea is not uncommon in theoretical psychometrics (e.g., Cronbach et al. 1972, Ch. 1 and 11; Fischer, 1974, Ch. 19; Guttman, 1945, p. 265-266; Stout, 1990) and it is necessary in mathematical statistics whenever asymptotic properties such as consistency are considered. Our results illustrate once again the power of it. From a strictly empiricists point of view, however, one might reject the idea. We do not agree with that point of view. Although the actual sample of manifest variables if always finite, the *universe of generalization* may well be infinite. This is comparable with the assumption of an infinite population of persons (Cronbach et al., p. 9, p. 20) - which is a useful theoretical assumption even though it is not actually true.

Second, we discussed the consequences for a *lower-bound of latent dimensionality*. Under Corollary 1 we compared our result with that of Suppes and Zanotti (1981), who showed that there always exist a unidimensional latent variable such that conditional independence holds. If the restriction of tail-measurability is added, such a latent variable does not always exist, but this restriction is still not strong enough to distinguish unidimensional universes from finite-dimensional universes without further model restrictions such as monotonicity.

Third, we introduced *tail-conditional true-scores* and a new definition of *subpopulations*, both based on the long-run information of manifest variables. These concepts generalize the corresponding concepts of the stochastic-subject formulation (Lord & Novick, 1968), the model-based random sampling formulation (Holland, 1990), and the specialization of the latter in the domain-sampling formulation (Lord & Novick, 1968, Ch. 8; Mulaik & McDonald, 1978; McDonald & Mulaik, 1979). Our definitions are model-free and do not necessarily involve a hypothetical replication experiment. In this sense, they are more flexible than both special cases. Related to this, we introduced the notion of a *stochastic meta-subject*, which is a set of persons that are indistinguishable in their long-run behavior as far as the given universe of manifest variables is concerned. A substantive interpretation of a latent variable may be given in terms of either stochastic subjects or stochastic meta-subjects, depending on whether the model satisfies a strict or a weak form of subpopulation invariance.

Finally, one may ask whether the restriction of tail-measurability should perhaps be replaced by the stronger restriction of *subtail-measurability*, which means that the latent variable can be estimated consistently from *every* infinite subsequence of manifest variables. This would agree more with the idea of "specific objectivity" (Rasch, 1977; Fischer, 1987) as it leads to *asymptotic specific objectivity:* The comparison of persons does not depend on the infinite subsequence of items used.

12. Appendix: Lemmas and Proofs

It is unavoidable to assume in this section that the reader is familiar with measure theory. As a standard reference we use the textbook of Billingsley (1986). Measure theory is particularly dominant in Propositions 4-6, Lemmas 1-3, 8, 11 and 12. Readers who are not familiar with measure theory may probably still understand most of the other proofs, with some effort, if they use the intuitive descriptions of the main text. A better intuition may often be achieved by considering the special cases $F = \sigma(\Theta)$ or $F = \tau(\mathbf{X})$.

It is worthwhile to point out here that in most interesting cases of our theory, Ω *must be uncountable*, because an infinite sequence of manifest variables has uncountably many possible outcomes (for example, if the X_i are binary, Ω must contain an element for every possible infinite sequence of 0s and 1s). This forces us in the proofs to take care of certain mathematical anomalies, such as nonempty events with probability 0. This was suppressed in the main text.

Notation and general facts.

All random variables, events, σ-fields and probabilities under consideration are supposed to be defined in a common probability space $(\Omega, \mathbf{P}, P)$. A, A_n, and B are events; F and F_n are σ-fields; $\mathbf{X}$ is a collection of random variables X_i; $\mathbf{Y}$ and Θ are random vectors.

$\mathbf{1}(A)$ is the indicator function of A (1 in A, 0 outside A).
$A \Delta B = (A-B) \cup (B-A)$, the symmetric difference of A and B.
$A_n \uparrow A$ means that $A_n \subseteq A_{n+1}$ and $\cup_n A_n = A$.
$A_n \downarrow A$ means that $A_{n+1} \subseteq A_n$ and $\cap_n A_n = A$.
$F_n \uparrow F$ means that $F_n \subseteq F$ and $\cup_n F_n$ *generates* F.
$F_n \downarrow F$ means that $F_{n+1} \subseteq F_n$ and $F = \cap_n F_n$.
$\sigma(\mathbf{X})$ is the σ-field generated by $\mathbf{X}$.
$\mathbf{X} \in F$ means that $\mathbf{X}$ is measurable F, i.e., $\sigma(\mathbf{X}) \subseteq F$.
$\perp\mathbf{X}|F$ means that $\mathbf{X}$ is conditionally independent given F.
$\perp\mathbf{X}|\Theta$ means $\perp\mathbf{X}|\sigma(\Theta)$.
$\alpha\mathbf{X}|A$ means that $\mathbf{X}$ is associated given A (see main text).
$\mu\mathbf{X}|\Theta$ means that $\mathbf{X}$ is monotone in Θ (see main text).

Tail-measurability

Proposition 1. If **Y** is measurable $\sigma(\mathbf{X})$, then there exists a measurable function **f** such that $\mathbf{Y} = \mathbf{f}(\mathbf{X})$ (by Billingsley, Problem 36.1 and p. 184).

Proposition 2. If **Y** is measurable $\sigma(\mathbf{X})$ then **Y** is measurable $\sigma(X_i : i \in S)$ for some countable subset S of indices (Billingsley, Theorem 36.3 and Problem 36.3).

Proposition 3. If $A_n \uparrow A$ then $P(A_n|F) \uparrow P(A|F)$ a.s. If $A_n \downarrow A$ then $P(A_n|F) \downarrow P(A|F)$ a.s. (Billingsley, p. 458).

Proposition 4. $\cup_n \sigma(X_1,...,X_n)$ is a field that generates $\sigma(X_1, X_2,...)$ (analogous to Billingsley, pp. 509-510.). Thus $\sigma(X_n,...,X_m) \uparrow \sigma(X_n, X_{n+1},...)$ as $m \to \infty$.

Proposition 5. If **A** is a semiring that contains Ω, then the field generated by **A** consists of the finite disjoint unions of elements of **A** (Billingsley, Problem 11.3.a).

Proposition 6. Let *F*, *G* and *H* be σ-fields. If G_0 is a field that generates *G*, and $P(A|F)$ is a.s. measurable *H* for every G_0-event A, then $P(A|F)$ is a.s. measurable *H* for every *G*-event A.

Proposition 7. Suppose U is a random variable that is bounded by $C \in \mathrm{IR}$. If $P(A \Delta B) = \delta > 0$ then $|E(U|A) - E(U|B)| \leq 2CP(A)^{-1}\delta$.

Proof of Proposition (6). For any set $A \in G$ there exist G_0-events C_{ni} such that, if $C = \cup_n \cap_i C_{ni}$, then $C \subseteq A$ and $P(A-C) = 0$ (Billingsley, problem 11.8). For any finite subset K of IN, let $D_{Km} = \cap_{n \in K} \cap_{i=1}^{m} C_{ni}$ and $D_K = \cap_m D_{Km}$. Then $D_{Km} \downarrow D_K$ as $m \to \infty$, so $P(D_{Km}|F) \downarrow P(D_K|F)$, which shows that $P(D_K|F)$ is a.s. measurable *H*. Define $C_n^* = \cup_{i=1}^{n} D_{\{i\}}$. Then $P(C_n^*|F)$ is a.s. measurable *H* since it can be written as a finite sum of conditional probabilities of the form $P(D_K|F)$ by the inclusion-exclusion formula. But $C_n^* \uparrow C$, so $P(C_n^*|F) \uparrow P(C|F)$; thus $P(C|F)$ is a.s. measurable *H*; but $P(C|F) = P(A|F)$ a.s. [ ]

Proof of Proposition (7). $|E(U|A) - E(U|B)| \leq$

$C \int |\mathbf{1}_A P(A)^{-1} - \mathbf{1}_B P(B)^{-1}| \leq$

$C \int (|\mathbf{1}_A - \mathbf{1}_B| P(A)^{-1} + \mathbf{1}_B |P(A)^{-1} - P(B)^{-1}|) =$

$C(\delta P(A)^{-1} + |P(B) - P(A)| P(A)^{-1}) \leq$

$C(\delta P(A)^{-1} + \delta P(A)^{-1})$. [ ]

Lemmas Concerning Conditional Independence

Consider a fixed sequence of random variables $\mathbf{X} = (X_1, X_2, ...)$ and a fixed σ-field F. Let $T_{ix} = P[X_i > x|F]$ for all $i \in \mathbb{N}$, $x \in \mathbb{R} \cup \{-\infty, \infty\}$. Denote $\mathbf{T} = (T_{ix} : (i,x) \in \mathbb{N} \times \mathbb{R})$. Let

$\tau(\mathbf{X}) = \cap_n \sigma(X_i : i \geq n)$ and
$\tau(\mathbf{T}) = \cap_n \sigma(T_{ix} : i \geq n, x \in \mathbb{R})$,

the tail σ-fields of $\mathbf{X}$ and $\mathbf{T}$, respectively.

Lemma 1. If $A \in \sigma(\mathbf{X})$ and $\perp\mathbf{X}|F$, then $P(A|F) \in \sigma(\mathbf{T})$.

Proof. Define $\mathbf{A}_n$ as the collection of events A of the form $A = \cap_{i=1}^n [x_i < X_i \leq y_i]$ where $x_i, y_i \in \mathbb{R} \cup \{-\infty, \infty\}$. Conditional independence of $\mathbf{X}$ given F implies that for such A, $P(A|F) = \Pi_{i=1}^n (T_{ix(i)} - T_{iy(i)})$. Thus $P(A|F) \in \sigma(T)$ for all $A \in \mathbf{A}_n$.

$\mathbf{A}_n$ is a semiring (analogous to Billingsley, example 11.5) and contains Ω. Let G_n be the field generated by $\mathbf{A}_n$. For $A \in G_n$, write $A = \cup_k A_k$ with the A_k disjoint $\mathbf{A}_n$-events, as in Proposition 5; then $P(A|F) = \Sigma_k P(A_k|F)$. Thus $P(A|F) \in \sigma(\mathbf{T})$ for all $A \in G_n$.

Each G_n is a field that generates $\sigma(X_1,...,X_n)$, so by Proposition 6, $P(A|F) \in \sigma(\mathbf{T})$ for all $A \in \cup_n \sigma(X_1,...,X_n) \equiv G$. But G is a field that generates $\sigma(\mathbf{X})$, so by Proposition 6, $P(A|F) \in \sigma(\mathbf{T})$ for all $A \in \sigma(\mathbf{X})$. [ ]

Lemma 2. If $F \subseteq \sigma(\mathbf{X})$ and $\perp\mathbf{X}|F$, then $F = \sigma(\mathbf{T})$ up to null events.

Proof. $\sigma(\mathbf{T}) \subseteq F$ because each $T_{ir} \in F$. It remains to prove that $F \subseteq \sigma(\mathbf{T})$. If $A \in F$, then by Lemma 1, $\mathbf{1}(A) = P(A|F) \in \sigma(\mathbf{T})$ and by standard arguments this is enough to show $A \in \sigma(\mathbf{T})$ and hence $F \subseteq \sigma(\mathbf{T})$. [ ]

Lemma 3. If $F \subseteq \tau(\mathbf{X})$ and $\perp\mathbf{X}|F$, then $F = \sigma(\mathbf{T}) = \tau(\mathbf{T}) = \tau(\mathbf{X})$ up to null events.

Proof. Since $(X_n, X_{n+1}, ...)$ has the same tail σ-field as $\mathbf{X}$, Lemma 2 implies that $F = \sigma(T_{ix} : i \geq n, x \in \mathbb{R})$ for each $n \in \mathbb{N}$. So $F = \sigma(\mathbf{T}) = \tau(\mathbf{T})$. It remains to prove that $\tau(\mathbf{X}) \subseteq F$. For every $A \in \tau(\mathbf{X})$, since $\perp\mathbf{X}|F$,

Kolmogorov's zero-one law (Billingsley, Theorem 22.3) implies that $P(A|F)$ assumes only the values 0 and 1. Then $P(A|F)$ is the indicator variable of the F-event $B = [P(A|F) = 1]$, and it is easily shown that $P(A \Delta B) = 0$.[ ]

Lemmas Concerning the Monotone LV Model

A pair of variables (S,R) is *comonotone* if there exists an event C in **P** with $P(C) = 1$ and for all $v, w \in C$: $S(v) > S(w)$ implies $R(v) \geq R(w)$ (Ellis & Van den Wollenberg, 1993). Then C will be called a *comonotonicity event* of (S,R), and (S,R) is said to be *strictly comonotone* on C.

Lemma 4. If $F \subseteq \tau(\mathbf{X})$, $\perp\mathbf{X}|F$ and $\alpha\mathbf{X}|A$ for every nonnull event $A \in F$, then every pair (T_{ix},T_{jy}) is comonotone.

Proof. For $i \neq j$, by generalizations of Theorems 4b and 5 of Ellis and Van den Wollenberg (1993; replace V by F), $\mathrm{Cov}(T_{ix},T_{jy}|A) = \mathrm{Cov}(\mathbf{1}[X_i > x],\mathbf{1}[X_j > y]|A) \geq 0$ for all nonnull $A \in F$. This implies that (T_{ix},T_{jy}) is comonotone by the Lemma of Ellis and Van den Wollenberg.

For $i = j$, by Propositions 1, 2, and Lemma 3, there is a countable subset S of indices (j,z) with $j \geq i+1$ and $z \in \mathrm{IR}$ such that T_{ix} and T_{iy} are functions of $(T_{jz} : (j,z) \in S)$ on some event A with $P(A) = 1$. Let B be the intersection of A and the comonotonicity events of all pairs (T_{ix},T_{jz}) and (T_{iy},T_{jz}) with $(j,z) \in S$. Then $P(B) = 1$ because S is countable. For every $v,w \in B$, if $T_{ix}(v) > T_{ix}(w)$, there is at least one (j,z) in S such that $T_{jz}(v) \neq T_{jz}(w)$. But then, using the argument for $i \neq j$, $T_{jz}(v) > T_{jz}(w)$ and hence, by the same argument, $T_{iy}(v) \geq T_{iy}(w)$. [ ]

Lemma 5. If $F \subseteq \tau(\mathbf{X})$, $\perp\mathbf{X}|F$ and $\alpha\mathbf{X}|A$ for every nonnull event $A \in F$, then there exists a random variable $\Theta \in F$ such that $\mu\mathbf{X}|\Theta$ and $\perp\mathbf{X}|\Theta$. Moreover, $F = \sigma(\Theta) = \tau(\mathbf{X})$.

Proof. By Lemma 4, all (T_{ix},T_{jy}) are comonotone. Define $\Theta = \Sigma_{i,q} w_{iq} T_{iq}$ where q extends across the rational numbers and the w_{iq} are positive real numbers such that $\Sigma_{i,q} w_{iq} < \infty$. Then $\Theta \in F$ because each $T_{iq} \in F$. Each T_{iq} is a nondecreasing function of Θ because of comonotonicity and the fact that each w_{iq} is positive. For arbitrary $x \in \mathrm{IR}$, T_{ix} can be shown to be a nondecreasing function of Θ by approximating x with a monotone sequence of

rational numbers, using Proposition 3 of the appendix. This implies that T_{ix} satisfies the functional equation that defines $P[X_i > x|\Theta]$ (see Billingsley, p. 451). Thus $T_{ix} = P[X_i > x|\Theta]$ and hence $\mu\mathbf{X}|\Theta$. This, together with Lemma 3, implies that $\sigma(\mathbf{T}) \subseteq \sigma(\Theta) \subseteq F = \tau(\mathbf{X}) = \sigma(\mathbf{T})$. Thus $\sigma(\Theta) = F = \tau(\mathbf{X})$, which proves $\perp\mathbf{X}|\Theta$. [ ]

Lemma 6. If $F \subseteq \tau(\mathbf{X})$ and there exists a random variable $\Theta \in F$ such that $\mu\mathbf{X}|\Theta$ and $\perp\mathbf{X}|\Theta$, then $\perp\mathbf{X}|F$ and $\alpha\mathbf{X}|A$ for every nonnull set $A \in F$.

Proof. By Lemma 3, $\sigma(\Theta) = F = \tau(\mathbf{X})$, so $\perp\mathbf{X}|F$. Consider a nonnull event $A \in F$. Then $A \in \sigma(\Theta)$, so $\mu\mathbf{X}|\Theta$ and $\perp\mathbf{X}|\Theta$ do also hold conditionally on A. By Theorem 6 of Holland & Rosenbaum (1986), $\mathbf{X}$ is associated conditionally on A, hence $\alpha\mathbf{X}|A$. [ ]

Lemma 7. If $F \subseteq \tau(\mathbf{X})$, and Θ_1 and Θ_2 satisfy the conditions of Θ in Lemma 6, then there exists an event C with $P(C) = 1$ and such that Θ_2 is a strictly increasing function of Θ_1 on C.

Proof. By Lemmas 6 and 5, $\sigma(\Theta_1) = \sigma(\Theta_2) = \sigma(\mathbf{T})$. By Propositions 1 and 2, there is a countable set S of indices (i,x) and an event C with $P(C) = 1$ such that Θ_1 and Θ_2 are functions of $(T_{ix} : (i,x) \in S)$ on C, Θ_1 and Θ_2 are functions of each other on C, and all pairs (Θ_1, T_{ix}) and (Θ_2, T_{ix}) with $(i,x) \in S$ are strictly comonotone on C. For $v, w \in C$, if $\Theta_1(v) > \Theta_1(w)$, then there is a (j,y) in S such that $T_{jy}(v) > T_{jy}(w)$, but then $\Theta_2(v) > \Theta_2(w)$. [ ]

Lemmas Concerning Conditional Covariances

Denote $\mathbf{X}_{nm} = (X_n,...,X_m)$ and $\mathbf{X}_{n\infty} = (X_n,X_{n+1},...)$.

Lemma 8. Suppose U, V and UV are random variables and that $E(|U|)$, $E(|V|)$ and $E(|UV|)$ are finite.

1. $\mathrm{Cov}(U,V|\mathbf{X}_{nm}) \to \mathrm{Cov}(U,V|\mathbf{X}_{n\infty})$ as $m \to \infty$.
2. $\mathrm{Cov}(U,V|\mathbf{X}_{nm},\tau(\mathbf{X})) \to \mathrm{Cov}(U,V|\mathbf{X}_{n\infty})$ as $m \to \infty$.
3. $\mathrm{Cov}(U,V|\mathbf{X}_{n\infty}) \to \mathrm{Cov}(U,V|\tau(\mathbf{X}))$ as $n \to \infty$.

Proof. It holds that $\sigma(\mathbf{X}_{nm}) \uparrow \sigma(\mathbf{X}_{n\infty})$ as m -> ∞, $\sigma(\mathbf{X}_{nm}, \tau(\mathbf{X})) \uparrow \sigma(\mathbf{X}_{n\infty})$ as m -> ∞, and $\sigma(\mathbf{X}_{n\infty}) \downarrow \tau(\mathbf{X})$ as n -> ∞. Put $\mathrm{Cov}(U,V|G) = E(UV|G) - E(U|G)E(V|G)$ for the appropriate σ-fields *G* and apply the (backward) martingale convergence theorems (Theorems 35.5 and 35.7 of Billingsley). [ ]

Lemma 9. The following conditions are equivalent, and the same is true conditionally upon any σ-field *G*.

1. $\mathbf{X}_{1n}$ is independent.
2. Cov(f(**Y**), g(**Z**)) = 0 for every partition (**Y**,**Z**) of $\mathbf{Z}_{1n}$ and every two bounded functions f and g.
3. Same, but with bounded coordinatewise nondecreasing functions f and g.
4. Same, but with f and g of the form f(**Y**) = **1**[**Y** > **y**] and g(**Z**) = **1**[**Z** > **z**] (here > is taken coordinatewise and the coordinates of **y** and **z** are allowed to be -∞).

Proof. We will consider the unconditional case only. That 9.1 implies 9.2 implies 9.3 implies 9.4 is trivial. *9.4 implies 9.1*: Cov(**1**[**Y** > **y**], **1**[**Z** > **z**]) = 0 means that P([**Y** > **y**] ∩ [**Z** > **z**]) = P[**Y** > **y**]P[**Z** > **z**]. Apply this with **Y** = $X_1, X_2, ..., X_{n-1}$, respectively, then one obtains $P[\mathbf{X}_{1n} > \mathbf{x}] = \Pi_i P[X_i > x_i]$. [ ]

Lemma 10. The following conditions are equivalent.

1. $\perp \mathbf{X} | \tau(\mathbf{X})$.
2. $\perp \mathbf{X}_{1n'} | \mathbf{X}_{nm}, \tau(\mathbf{X})$ whenever $1 \le n' < n \le m$.
3. $\perp \mathbf{X}_{1n'} | \mathbf{X}_{n\infty}$ whenever $1 \le n' < n$.
4. For all n', n, m ∈ IN with $1 \le n' < n$, for every partition (**Y**,**Z**) of $\mathbf{X}_{1n'}$ and all functions f and g of the form specified in Lemma 9.1, 9.2 or 9.3, $\mathrm{Cov}(f(\mathbf{Y}), g(\mathbf{Z})|\mathbf{X}_{nm}) \to 0$ as m -> ∞.

Proof. 10.1 implies 10.2: $\perp \mathbf{X} | \tau(\mathbf{X})$ means $\perp \mathbf{X}_{1m} | \tau(\mathbf{X})$ for all m, and this implies 10.2. *10.2 implies 10.3*: By Lemmas 9 and 8.3. *10.3 implies 10.4*: By Lemmas 9 and 8.1. *10.4 implies 10.1*: By Lemmas 8.1, 8.3 and 9 one obtains $\perp \mathbf{X}_{1n'} | \tau(\mathbf{X})$ for all $n' \ge 1$. [ ]

Tail-measurability

Lemma 11. Suppose $1 \leq n' < n$. The following conditions 11.1 and 11.2 are equivalent, and both imply 11.3.

1. $\alpha\mathbf{X}_{1n'}|A$ for every nonnull event $A \in \sigma(\mathbf{X}_{n\infty})$.
2. $\alpha\mathbf{X}_{1n'}|A$ for every nonnull event $A \in \cup_n\sigma(\mathbf{X}_{nm})$
3. $\alpha\mathbf{X}_{1n'}|A$ for every nonnull event $A \in \tau(\mathbf{X})$.

Proof. 11.1 implies 11.2 because $\cup_n\sigma(\mathbf{X}_{nm}) \subseteq \sigma(\mathbf{X}_{n\infty})$, and 11.1 implies 11.3 because $\tau(\mathbf{X}) \subseteq \sigma(\mathbf{X}_{n\infty})$. *11.2 implies 11.1*: $\cup_n\sigma(\mathbf{X}_{nm})$ is a field that generates $\sigma(\mathbf{X}_{n\infty})$ (Proposition 4). So if A is an nonnull event in $\sigma(\mathbf{X}_{n\infty})$, there exists for every $\delta > 0$ an event B_δ in $\cup_n\sigma(\mathbf{X}_{nm})$ such that $P(A \Delta B_\delta) < \delta$ (Billingsley, Corollary on p. 167). B_δ must be a nonnull event if δ is sufficiently small. If U and V are coordinatewise nondecreasing functions of $\mathbf{X}_{1n'}$ that are bounded by C_U and C_V, respectively, then $|\mathrm{Cov}(U,V|A) - \mathrm{Cov}(U,V|B_\delta)| \leq 6C_UC_VP(A)^{-1}\delta$ by Proposition 7. However, $\mathrm{Cov}(U,V|B_\delta) \geq 0$ for all B_δ, thus $\mathrm{Cov}(U,V|A) \geq 0$. [ ]

Lemmas Concerning Consistent Estimation

Lemma 12. For a random vector Θ, the following conditions 12.1 and 12.2 are equivalent.

1. $\Theta \in \sigma(\mathbf{X})$.
2. There exist Borel-measurable mappings $\mathbf{f}_n$ such that $\mathbf{f}_n(\mathbf{X}_{1n}) \to \Theta$ a.s.

Proof. 12.1 implies 12.2: It is sufficient to prove this for real-valued Θ, which can be done by standard approximation arguments, e.g., using Billingsley, Theorem 13.5, Problem 11.7 and the first Borel-Cantelli lemma.

12.2 implies 12.1: Since $\mathbf{f}_n(\mathbf{X}_{1n}) \in \sigma(\mathbf{X})$ for each n. [ ]

Lemma Concerning Order Independence.

Lemma 13. If $\mathbf{X}'$ is a permutation of $\mathbf{X}$ then $\tau(\mathbf{X}') = \tau(\mathbf{X})$.

Proof. Consider the σ-field $\tau'(\mathbf{X}) = \cap_{K \subseteq IN \text{ finite}} \sigma(X_i : i \in IN\text{-}K)$, where the intersection extends across all finite subsets K of IN. Since $\tau'(\mathbf{X})$ does not depend on the order of the X_i in **X**, the lemma can be proved by showing that $\tau'(\mathbf{X}) = \tau(\mathbf{X})$. Remember that, by definition, $\sigma(X_i : i \in IN\text{-}\{1,...,n\})$ is the intersection of all σ-fields with respect to which all X_i with $i \in IN\text{-}\{1,...,n\}$ are measurable, which includes $\sigma(X_i : i \in IN\text{-}K\}$ if $K \subseteq \{1,...,n\}$. So $\sigma(X_i : i \in IN\text{-}\{1,...,n\}) \subseteq \cap_{K \subseteq \{1,...,n\}} \sigma(X_i : i \in IN\text{-}K\}$. But since the latter intersection includes $K = \{1,...,n\}$, we also have $\cap_{K \subseteq \{1,...,n\}} \sigma(X_i : i \in IN\text{-}K\} \subseteq \sigma(X_i : i \in IN\text{-}\{1,...,n\})$; thus

$$\sigma(X_i : i \in IN\text{-}\{1,...,n\}) = \cap_{K \subseteq \{1,...,n\}} \sigma(X_i : i \in IN\text{-}K\}.$$

Therefore, $\tau(\mathbf{X}) = \cap_n \sigma(X_i : i \in IN\text{-}\{1,...,n\}) = \cap_n \cap_{K \subseteq \{1,...,n\}} \sigma(X_i : i \in IN\text{-}K\} = \cap_{K \subseteq IN \text{ finite}} \sigma(X_i : i \in IN\text{-}K\} = \tau'(\mathbf{X})$. [ ]

Lemma Concerning Infinite Subsequences.

Define the σ-field $\Sigma(\mathbf{X})$ as the intersection of all σ-fields that are generated by an infinite subsequence of **X**.

Lemma 14.
1. $\Sigma(\mathbf{X}) \subseteq \tau(\mathbf{X})$.
2. If **X'** is an infinite subsequence of **X** then $\Sigma(\mathbf{X}) \subseteq \Sigma(\mathbf{X}')$.
3. $F \subseteq \Sigma(\mathbf{X})$ if and only if $F \subseteq \Sigma(\mathbf{X}')$ for every infinite subsequence **X'** of **X**.
4. If $F \subseteq \Sigma(\mathbf{X})$ and $\perp\mathbf{X}|F$ then $F = \Sigma(\mathbf{X}) = \tau(\mathbf{X})$.
5. If $\perp\mathbf{X}|\Sigma(\mathbf{X})$ and **X'** is an infinite subsequence of **X**, then $\Sigma(\mathbf{X}) = \Sigma(\mathbf{X}')$
6. $\perp\mathbf{X}|\Sigma(\mathbf{X})$ if and only if $\perp\mathbf{X}'|\Sigma(\mathbf{X}')$ for every infinite subsequence **X'** of **X**.

Proof. 14.1, 14.2, and 14.3 follow from the definitions by elementary set theory. Using 14.1 and Lemma 3, one obtains 14.4. Then 14.5 follows, since **X'** is conditionally independent given $\Sigma(\mathbf{X})$ and $\Sigma(\mathbf{X}) \subseteq \Sigma(\mathbf{X}')$, by application of 14.4 to **X'**. 14.5 is a direct consequence of 14.4. [ ]

Tail-measurability

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43,* 561-573.

Billingsley, P. (1986). *Probability and measure.* New York: Wiley.

Byrne, B.M. (1989). Multigroup comparisons and the assessment of equivalent construct validity across groups: Methodological and substantive issues. *Multivariate Behavioral Research, 24,* 503-523.

Cronbach, L.J., Gleser, G.C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Ellis, J.L. (1993). Subpopulation invariance of patterns in covariance matrices. *British Journal of Mathematical and Statistical Psychology, 46,* 231-254.

Ellis, J.L. & Van den Wollenberg, A.L. (1993). Local homogeneity in latent trait models. A characterization of the homogeneous monotone IRT model. *Psychometrika, 58,* 417-429.

Esary, J.D., Proschan, F. & Walkup, D.W. (1967). Association of random variables, with applications. *Annals of Mathematical Statistics, 38,* 1466-1474.

Fischer, G.H. (1974). *Einfuhrung in die Theorie psychologischer Tests.* Bern: Huber.

Fischer, G. H. (1987). Applying the principles of specific objectivity and of generalizability to the measurement of change. *Psychometrika, 52,* 565-587.

Fischer, G.H. (1995). Derivations of the Rasch model. In Fischer, G.H. & Molenaar, I.W., (Eds), *Rasch models. Foundations, recent developments and applications,* Ch. 2. New York: Springer-Verlag.

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

Grayson, D.A. (1990). Selection effects and unidimensionality in item response theory. *British Journal of Mathematical and Statistical Psychology, 43,* 207-216.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10,* 255-282.

Guttman, L. (1953). Image theory for the structure of quantitative

variates. *Psychometrika, 18,* 227-296.

Holland, P.W. (1981). When are item response models consistent with observed data? *Psychometrika, 46,* 79-92.

Holland, P.W. (1990). On the sampling theory foundations of item response theory models. *Psychometrika, 55,* 577-601.

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

Junker, B.W. (1991). Essential independence and likelihood-based ability estimation for polytomous items. *Psychometrika, 56,* 255-278.

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

Junker, B.W., & Ellis, J.L. (1995). *A characterization of monotone unidimensional latent variable models.* Submitted.

Kaiser, H., & Caffrey, J. (1965). Alpha factor analysis. *Psychometrika, 30,* 1-14.

Kingman, J.F.C. (1978). Uses of exchangeability. *Annals of Probability, 6,* 183-197.

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

Lazarsfeld, P.F. (1959). Latent structure analysis. In S. Koch (Ed.), *Psychology: A study of science, Vol. 3* (pp. 476 - 543). New York: McGraw Hill.

Lehmann, E.L. (1986). *Testing statistical hypotheses.* New York: Wiley.

Lord, F.M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

McDonald, R.P. (1977). The indeterminacy of components and the definition of common factors. *British Journal of Mathematical and Statistical Psychology, 30,* 165-176.

McDonald, R.P. & Mulaik, S.A. (1979). Determinacy of common factors: A nontechnical review. *Psychological Bulletin, 86,* 297-306.

McDonald, R.P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.

Mellenbergh, G.J. (1989). Item bias and item response theory. *International Journal of Educational Research, 13,* 127-143

Meredith, W., & Millsap, R.E. (1992). On the misuse of manifest variables in the detection of measurement bias. *Psychometrika, 57,* 289-311.

Mokken, R.J. (1971). *A theory and procedure of scale-analysis.* The Hague: Mouton.

Mood, A.M., Graybill, F.A., & Boes, D.C. (1974). *Introduction to the theory of statistics.* New York: McGraw-Hill.

Mulaik, S.A., & McDonald, R.P. (1978). The effect of additional variables on factor indeterminacy in models with a single common factor. *Psychometrika, 43,* 177-192.

Ramsay, J.O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika, 56,* 611-630.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Nielson & Lydiche.

Rasch, G. (1977). On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. In M. Blegvad (Ed.), *The Danish yearbook of philosophy.* Copenhagen: Munksgaard.

Rosenbaum, P.R. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. *Psychometrika, 49,* 425-435.

Roskam, E.E., & Jansen, P.G.W. (1984). A new derivation of the Rasch model. In E. Degreef & J. Van Buggenhaut (Eds), *Trends in mathematical psychology.* Amsterdam: Elsevier.

Rozenboom, W.W. (1988). Factor indeterminacy: The saga continues. *British Journal of Mathematical and Statistical Psychology, 41,* 209-226.

Sijtsma, K. (1988). *Contributions to Mokken's nonparametric item response theory.* Amsterdam: Free University Press.

Steiger, J.H. (1979). Factor indeterminacy in the 1930's and the 1970's some interesting parallels. *Psychometrika, 44,* 157-161.

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

Suppes, P. & Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese, 48,* 191-199.

Thurstone, L.L. (1957). *Multiple-factor analysis.* Chicago: University of Chicago Press.

Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31*, 279-311.

Warm, T.A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427-250.

TABLE 1. Example of a distribution for $(\Theta,\Theta',X_1,X_2,X_3,X_4)$. Here *pattern* is the score pattern on (X_1,X_2,X_3,X_4), and *F* is the frequency.

Θ	Θ'	pattern	F	Θ	Θ'	pattern	F
1/3	1/3	0000	81	3	3	0000	1
1/3	1/3	0001	27	3	3	0001	3
1/3	1/3	0010	27	3	3	0010	3
1/3	**3**	**0011**	**9**	**3**	**1/3**	**0011**	**9**
1/3	1/3	0100	27	3	3	0100	3
1/3	1/3	0101	9	3	3	0101	9
1/3	1/3	0110	9	3	3	0110	9
1/3	1/3	0111	3	3	3	0111	27
1/3	1/3	1000	27	3	3	1000	3
1/3	1/3	1001	9	3	3	1001	9
1/3	1/3	1010	9	3	3	1010	9
1/3	1/3	1011	3	3	3	1011	27
1/3	1/3	1100	9	3	3	1100	9
1/3	1/3	1101	3	3	3	1101	27
1/3	1/3	1110	3	3	3	1110	27
1/3	1/3	1111	1	3	3	1111	81