



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

**“VALIDACIÓN Y EVALUACIÓN ESTADÍSTICA DEL
EXAMEN DE COLOCACIÓN AL CURSO EN LÍNEA
PARA LA COMPRESIÓN DE LECTURA EN
INGLÉS”**

T E S I S A
QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN MATEMÁTICAS
APLICADAS Y COMPUTACIÓN.

PRESENTA

DANIEL SÁNCHEZ DELGADO

ASESORA: MTRA. ELVIRA BEATRIZ CLAVEL DÍAZ

NAUCALPAN, EDO. MEX.

MAYO 2008



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Quiero agradecer primeramente a mis Padres, por que nunca escatimaron esfuerzo alguno y se sacrificaron gran parte de su vida para formarme y educarme.

A mis Hermanos, de quiénes siempre recibí un apoyo incondicional y a quienes les dedico este trabajo. Gracias por confiar y creer en mí...

A la Mtra. Beatriz Clavel, por ser una excelente asesora y profesora, mi respeto y admiración. Gracias por tenerme paciencia y por su disposición para escucharme y animarme.

A Dios, por la vida, salud y fortaleza que me brinda para levantarme siempre que tropiezo.

Y por último, Agradezco al amor de mi vida, Elizabeth, por estar conmigo en las buenas y en las malas, y gracias a u su constante apoyo logré culminar este trabajo que forma parte de nuestros proyectos de vida.

ÍNDICE

INTRODUCCIÓN	vi
CAPÍTULO 1. ANTECEDENTES	
1.1 CONTEXTO DEL PROBLEMA DE INVESTIGACIÓN: EVALUACIÓN DE CONOCIMIENTOS DE COMPRENSIÓN DE LECTURA.	2
1.2 PSICOMETRÍA	5
1.2.1 Teoría clásica	6
1.2.2 Teoría respuesta al ítem	8
1.3 MODELOS ESTADÍSTICOS DE ANÁLISIS CAUSA-EFECTO	11
CAPÍTULO 2. MÉTODOS ESTADÍSTICOS PARA LA VALIDACIÓN DE ITEMS	
2.1 FIABILIDAD	13
2.2 MEDIDAS DE FIABILIDAD	15
2.3 MODELOS DE FIABILIDAD TCT	16
2.3.1 Modelo alfa de Cronbach	16
2.3.2 Modelos de dos mitades	17
2.3.3 Modelos de formas equivalentes	19
2.4 MODELOS DE FIABILIDAD TRI	20
2.4.1 Supuestos de la TRI	21
2.4.1.1 Curva Característica del ítem	21
2.4.1.2 Parámetros	22
2.4.1.3 Unidimensionalidad	23
2.4.1.4 Independencia local	23
2.4.2 Modelos	24
2.4.2.1 Modelo logístico de un parámetro	24
2.4.2.2 Modelo logístico de dos parámetros	25
2.4.2.3 Modelo logístico de tres parámetros	25
2.4.3 Estimación de parámetros	26
2.4.4 Ajuste del modelo	26
2.4.5 Invarianza de las estimaciones y función de información	28
2.5 VALIDEZ	30
2.5.1 Validez de Contenido	30

2.5.2	Validez de Criterio	31
2.5.3	Validez de Constructo	32

CAPÍTULO 3. TÉCNICAS MULTIVARIABLES Y MODELOS CAUSALES

3.1	INTRODUCCIÓN	34
3.2	MODELO CAUSAL	34
3.2.1	Tipos de relaciones causales	37
3.3	ANÁLISIS DE REGRESIÓN MÚLTIPLE	38
3.4	ANÁLISIS FACTORIAL	39
3.4.1	Requerimientos teóricos	40
3.4.2	Supuestos	41
3.5	ANÁLISIS PATH O DE SENDEROS	42
3.6	MODELOS DE ECUACIONES ESTRUCTURALES	46
3.6.1	Etapas del modelado de ecuaciones estructurales	47
3.6.1.1	Especificación de modelo	48
3.6.1.2	Identificación de modelo	51
3.6.1.3	Estimación de parámetros	52
3.6.1.4	Evaluación del modelo	53
3.6.1.5	Interpretación del modelo	56

CAPÍTULO 4. EVALUACIÓN Y VALIDACIÓN DEL EXAMEN DE COLOCACIÓN AL CURSO EN LÍNEA PARA LA COMPRESIÓN DE LECTURA EN INGLÉS

4.1	DESCRIPCIÓN DEL PROBLEMA	58
4.1.1	Objetivo General	59
4.1.2	Objetivos Particulares	59
4.1.3	Marco contextual	59
4.1.4	Selección y descripción de la muestra	60
4.1.5	Instrumentos para la recolección de datos	60
4.1.6	Selección de variables	61
4.2	DESARROLLO DE LA SOLUCIÓN	64
4.2.1	Evaluación de los supuestos de la TRI	64
4.2.2	Modelo TRI de tres parámetros	70
4.2.3	Estimación del coeficiente de consistencia interna α	77
4.3	ANÁLISIS DE ECUACIONES ESTRUCTURALES	80
4.3.1	Requerimientos teóricos	81

4.3.2	Especificación del modelo	83
4.3.3	Identificación del modelo	86
4.3.4	Estimación de parámetros	87
4.3.5	Evaluación del modelo	90
4.4	ANÁLISIS DE RESULTADOS	96
4.4.1	Interpretación del modelo	96
4.4.2	Limitaciones	99
	CONCLUSIONES	100
	BIBLIOGRAFÍA	103
	ANEXO A. Texto del examen de colocación	106
	ANEXO B. Análisis de fiabilidad del examen de colocación	108
	ANEXO C. Análisis factorial confirmatorio (AFC)	122

INTRODUCCIÓN

El presente trabajo es una investigación de apoyo a un proyecto PAPIME de la FES Acatlán, el cual tiene como objetivo general desarrollar un curso en línea para la comprensión de lectura en inglés. La presente investigación colabora en el análisis de fiabilidad de un examen de colocación diseñado para dicho curso, así como en la validez de constructo de un modelo hipotético que representa las habilidades de los alumnos para la comprensión de lectura en inglés.

Generalmente para poner en evidencia los conocimientos de un sujeto, se aplican tests o exámenes –que en el caso de esta investigación son exámenes de colocación– los cuales permiten indagar las preconcepciones del alumno y que ponen en evidencia su nivel de conocimientos. Usar un test o cuestionario de examen elaborado para el diagnóstico de conocimientos como instrumento de recolección de datos lleva necesariamente a evaluar la confiabilidad y validez del mismo. En un sentido general, Nunnally (1991) señala que la validez de un instrumento de medición ocurre cuando éste cumple satisfactoriamente con el propósito para el cual fue construido. La confiabilidad se refiere a la capacidad que un instrumento de medición tiene para medir de forma consistente y precisa las características que pretende medir. Con el análisis de confiabilidad también pretende que el instrumento de medición esté libre de errores de medida.

Aparte de conocer la validez y confiabilidad del examen de colocación, utilizado como instrumento para la obtención de datos, en la presente investigación es de interés evaluar la validez de constructo del mismo, partiendo de un modelo hipotético previamente especificado por los docentes del área, en el que se establecen y especifican las relaciones causa-efecto y algunas de las dimensiones que representan la habilidad de los estudiantes para elegir y aplicar las estrategias de lectura adecuadas para comprender de manera eficiente textos académicos escritos en inglés. Los modelos estadísticos multivariados nos permiten explorar el comportamiento conjunto de variables y sus relaciones para evaluar la validez de constructo. En este sentido la técnica estadística de ecuaciones estructurales, apoyándose en los principios

estadísticos del modelo general de regresión lineal múltiple, el análisis factorial y un diagrama de senderos, permiten explicar las relaciones causa-efecto hipotetizadas.

Por lo anterior, en este trabajo se presenta el proceso estadístico que se sigue para la evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura de textos en inglés, así como un modelo de ecuaciones estructurales, que permita evaluar y observar las dimensiones pedagógicas previamente hipotetizadas que lo explican.

Es importante tener una idea clara del rumbo que se desea seguir en una investigación, por lo que es necesario dar una visión general antes de abordar un tema en específico. Para tal efecto este trabajo se estructura de la siguiente manera:

El propósito del Capítulo 1. Antecedentes, es el de introducir al lector con los elementos que se pueden validar estadísticamente de un examen de colocación de un curso en línea enfocado a la comprensión de lectura en inglés, definiendo algunos de los factores que determinan un aprendizaje eficiente en la comprensión de los textos de acuerdo a diversos enfoques pedagógicos, que nos permiten determinar las variables endógenas, exógenas, latentes y observables, que afectan y definen dichas habilidades. De la misma forma, se presentan las ideas fundamentales de las teorías psicométricas: teoría clásica de los tests y teoría de respuesta al ítem. Por último se incorpora al texto una breve introducción de modelos causales y ecuaciones estructurales.

Una vez expuesto el panorama en torno a lo que se pretende evaluar del examen de colocación, y habiendo detallado algunos de los elementos teóricos que intervienen en el proceso de enseñanza de la comprensión y lectura de textos en inglés, se procede en el Capítulo 2. Métodos estadísticos para la evaluación de ítems, a explicar los conceptos de fiabilidad (confiabilidad) y validez; se profundiza en la teoría clásica de los test (TCT) y la teoría de respuesta al ítem (TRI), detallando los modelos matemáticos más representativos de ambas teorías que permiten la evaluación del test. Por último, se habla sobre la validez de contenido, de constructo y de criterio.

En el Capítulo 3. Técnicas multivariantes y modelos causales, como sustento para evaluar la validez de constructo, se abordan métodos estadísticos multivariados. Así, para estimar los

parámetros, el capítulo inicia con la exposición de los modelos causales y de técnicas estadísticas tales como: el análisis de regresión múltiple, el análisis factorial, el análisis path o de senderos y por último el modelo de ecuaciones estructurales, especificando sus características y el tipo de variables que se pueden analizar.

Ya que el objetivo de este trabajo es validar el constructo de un modelo hipotético a través de modelos de ecuaciones estructurales, una gran parte de este capítulo está dedicada al análisis de ecuación estructurales, comenzando con la definición de este último y especificando sus principales componentes: las variables latentes, el error de medición, la fiabilidad de las medidas en las variables latentes. Se revisan con detalle sus cinco etapas de modelación: i) especificación, ii) identificación, iii) estimación, iv) evaluación, e v) interpretación.

El Capítulo 4. Evaluación estadística y validación del examen de colocación al curso en línea para la enseñanza de comprensión de lectura en inglés, se enfoca a la descripción y desarrollo de la aplicación, se describe el problema, objetivos y sus limitaciones, así como el diseño de los instrumentos de medición que nos permiten evaluar dicho curso. Posteriormente se presentan las variables que intervienen en el planteamiento del modelo causal y la muestra con la que se trabaja.

Después se procede al análisis de los ítems, para lo cual se evalúa el cumplimiento de los supuestos y se aplican las técnicas estadísticas necesarias para realizar el análisis de fiabilidad y validez. También se plantea el modelo causal para la comprensión de lectura en inglés, mismo que se representa por medio de un diagrama de senderos y es estimado mediante la técnica de ecuaciones estructurales. Por último, se presentan y discuten los resultados.

Las conclusiones abordan el alcance y las limitaciones que implica esta investigación.

CAPÍTULO 1

ANTECEDENTES

1.1 CONTEXTO DEL PROBLEMA DE INVESTIGACIÓN: EVALUACIÓN DE CONOCIMIENTOS DE COMPRENSIÓN DE LECTURA

El aprendizaje es un proceso de obtención de información, por medio de la cual se pueden formar juicios que conducen a la toma de decisiones (TenBrink, 2002). De acuerdo con Reimers (2003), la evaluación del aprendizaje es una línea particular de investigación educativa, que se ocupa de describir y explicar diversos aspectos que inciden en las oportunidades de aprendizaje, para lo cual se relacionan distintos resultados educativos con los factores que influyen en el aprendizaje de los alumnos, y estas relaciones se utilizan para proponer o establecer la efectividad de programas o políticas que podrían afectar estas condiciones.

Por otra parte, la evaluación busca recoger evidencias en relación con las características y propiedades de los fenómenos que estudia y para ello, por lo general, utiliza una serie de instrumentos y herramientas que le permiten contar con una mejor descripción de las propiedades relevantes de su objeto de estudio (Fernández-Ballestreos, 1997). Estas propiedades o atributos de los objetos se expresan a través de la definición de constructos¹ que pueden ser identificados en la respuesta de los estudiantes en una prueba (Martínez y Moreno, 2002), o en la interacción de los alumnos con eventos u objetos en situaciones específicas, o bien, en los resultados esperados al diseñar y aplicar programas educativos. Cuando se asignan valores (puntuaciones numéricas) a esas características o propiedades representadas por los constructos (indicadores), entonces se dice que se está midiendo un comportamiento.

Los instrumentos elaborados para medir un factor o rasgo determinado deben recoger evidencias confiables acerca de las propiedades y atributos definidos teóricamente en tal instrumento, de modo que la medición de los resultados requiere evaluar los criterios de la validez, de cada uno de los reactivos del instrumento, tanto como los constructos implicados en el mismo. Actualmente existen dos tipos de evaluaciones que consideran tal evaluación: la adaptativa y la generativa de respuestas.

¹ Se entiende por constructo variables latentes, no observables, que reflejan un comportamiento, conducta o aptitud.

De acuerdo con Castañeda y sus colaboradores (1993) señalan que “La evaluación adaptativa tiene la característica de que las pruebas no se diseñan con un carácter fijo para todas las personas sino que, más bien, los reactivos se seleccionan de un banco a partir de la información inicial, misma que se obtiene mediante el nivel de dominio de habilidad o conocimiento que se pretende evaluar”. Mientras que la evaluación generativa de respuestas, supone una aproximación al diseño de pruebas desde la validación de constructos, donde los ítems son generados conjuntamente con la estimación de sus parámetros psicométricos en función de los intereses de la evaluación. Martínez y Moreno (2002).

El proceso de hacer explícita una palabra abstracta en términos de variables observables se conoce como validación del constructo, así la validez de constructo se refiere a la garantía de que las variables empíricas son indicadores adecuados del concepto o modelo teórico a ser evaluado.

De acuerdo a los planteamientos de Martínez y Moreno (2002), se puede considerar que la elaboración y aplicación de instrumentos para medir y evaluar el aprendizaje se encuentra relacionada con la teoría que subyace a dichos instrumentos y a los constructos que se pretenden medir, e implica tres condiciones:

1. Delimitación teórica del constructo a medir a partir de la abstracción de alguna propiedad de los objetos o fenómenos bajo estudio.
2. Desarrollo de un diseño de observación o recolección que operativice la noción teórica respecto de dichos constructos mediante la construcción de un instrumento de medición.
3. Aplicación de los modelos matemáticos para obtener medidas de la propiedad o constructos a partir de la ejecución de determinadas tareas o respuestas recogidas con los instrumentos.

Las condiciones anteriores son importantes para aportar indicadores valiosos en la evaluación de la comprensión de lectura. Este estudio en particular aborda el punto número

tres donde a parte de hacer una evaluación de constructo, se realiza el análisis de fiabilidad de los instrumentos de obtención de datos, que para este caso es el examen de colocación.

En el contexto de medición educativa, Castañeda (1993) desarrolló un marco de referencia que permite identificar los componentes del proceso de responder a diversos tipos de examen. Este marco emplea un análisis funcional de la ejecución crítica, que permite identificar la macroestructura de la medición a partir de la descripción integral de los elementos de competencia; esto es, de los conocimientos, habilidades y actitudes necesarias para lograr el resultado deseado al realizar una tarea. Una vez identificada la macroestructura de la medición, el marco emplea un análisis cognitivo de tareas (ACT) consistente en un análisis recursivo que descompone una tarea en los componentes constitutivos esenciales que conforman los constructos a ser evaluados.

El ACT identifica las operaciones cognitivas solicitadas, los tipos de conocimiento a evaluar y los contextos en lo que serán medidas las habilidades y tareas de forma gradual. Una vez identificado lo anterior, el procedimiento permite reconocer los mecanismos del proceso de responder exámenes para poder validar los constructos hipotetizados durante la medición, al asumir que la ejecución mostrada por el examinado constituye un ejemplo representativo de aquello que se busca medir.

A partir del marco analítico descrito, para la investigación sobre validación de constructo aquí presentada, se identificaron, de acuerdo a los responsables de la construcción de los exámenes de colocación García Jurado, R. (2003), tres mecanismos responsables de la ejecución eficiente durante la comprensión de lectura, los dos primeros definidos a partir del modelo interactivo de lectura propuesto por Van Dijk y Kintsch (1983); el primero se refiere a la **microestructura del texto**, que corresponde al nivel local del discurso e incluye la estructura de las proposiciones individuales -en forma de oraciones-, la segunda corresponde a la **macroestructura** que indica los significados globales -a partir de párrafos- que caracterizan al texto como un todo. El tercer mecanismo corresponde al **procesamiento mental (cognitivo)** que mide la capacidad de analizar y organizar ideas, inferir, evaluar y sintetizar información.

Así el modelo causal que represente la habilidad de los estudiantes para elegir y aplicar las estrategias de lectura adecuadas para comprender de manera eficiente textos académicos escritos en inglés, estará conformado por la macroestructura, microestructura y proceso cognitivo.

1.2 PSICOMETRÍA

La psicometría nos brinda un cuerpo de teoría y métodos para la medición de constructos en ciencias sociales. Uno de sus propósitos principales es el desarrollo de técnicas de aplicación empírica que permitan construir instrumentos de medición, indicadores, de alta confiabilidad y validez. Estas técnicas y métodos se basan en enfoques cuantitativos y utilizan conceptos, procedimientos y medidas derivadas de la estadística.

Un instrumento psicométrico se caracteriza por que intenta representar el constructo por medio de un puntaje numérico derivado de la aplicación de un conjunto de reactivos o estímulos a los sujetos de interés. En su forma más usual el instrumento está compuesto por una serie de ítems o preguntas, cada uno de los cuales es calificado o respondido por el individuo de acuerdo con una cierta escala de medición, dicha escala puede ser: *nominal*, *ordinal*, *de intervalo*, *de razón*.

Los contenidos de la psicometría se articulan en tres bloques:

- **Teoría de la medición**, que se refiere a la fundamentación teórica de la medida;
- **Teoría de los Test**, que hace referencia a la construcción , validación y aplicación de los tests;
- **Escalamiento**, que incluye los métodos para la elaboración de escalas psicofísicas y psicológicas.

De los elementos anteriores, para el presente trabajo es de gran importancia la teoría de los tests, que a su vez, se divide en dos ramas: la teoría clásica de los tests y la teoría de respuesta a los ítems.

A menudo el docente realiza la construcción de tests que pretenden evaluar determinados constructos no susceptibles a un proceso de medición directa, como lo es la habilidad de comprensión lectora, para lo cual se han elaborado teorías estadísticas que permiten inferir el nivel de rasgo a partir del rendimiento observado de la persona.

1.2.1 Teoría clásica de los tests (TCT)

Se denomina Teoría Clásica de los Tests (TCT) al conjunto de principios teóricos y métodos cuantitativos que fundamentan la construcción, aplicación, validación e interpretación de distintos tipos de tests, que permiten derivar escalas estandarizadas aplicables a una población (Hambleton, 1991). Los principios en que se basa son relativamente simples y se aplican tanto a las pruebas de desempeño y a las de pruebas aptitud.

La teoría clásica de los tests propone un modelo formal, denominado como modelo clásico o modelo lineal clásico, fundamentado en diversos supuestos² a partir de los cuales se extraen determinadas consecuencias de aplicabilidad práctica para determinar el grado en que un test informa de los niveles de rasgo.³

La TCT es un modelo de puntuación verdadera como valor esperado, es decir, la puntuación verdadera del alumno en un test sería el promedio aritmético de las puntuaciones empíricas obtenidas en infinitas aplicaciones (Muñiz, 1998). La hipótesis fundamental de la TCT es que la puntuación observable de una persona en un test es una función de dos componentes: su puntaje verdadero y el error implícito en la prueba.

Según Murat (1973), “la puntuación verdadera del sujeto no significa que se pueda lograr la medición exacta del atributo mediante la aplicación de un test, sino que es un concepto teórico que se definiría como la puntuación promedio que obtendría ese sujeto en infinitas aplicaciones de una misma prueba”.

² Son vistos más adelante

³ Se entiende por rasgo aquella variable no observable que en general da cuenta de una serie de conductas

A partir de lo anterior se puede inferir que dada la puntuación empírica de un sujeto X en un test ,

- a) $X = V + \xi$, indica que la puntuación empírica directa de una persona en un test (X) está compuesta de dos componentes hipotéticos: el nivel de rasgo o puntuación verdadera de la persona (V) y un error de medida (\mathcal{E}) que se comete al medir el rasgo con el test.
- b) $\mathcal{E} = X - V$, es el error de medida se establece como la diferencia entre la puntuación empírica y la verdadera

El problema es que \mathcal{E} y V resultan en principio desconocidos, por lo que para poder obtener información de ellos se definen los siguientes supuestos:⁴

- 1) $V = E[X]$, se define la puntuación verdadera de una persona como el valor esperado de las posibles puntuaciones empíricas que puede obtener en el test.
- 2) $\rho(V, \mathcal{E}) = 0$, quiere decir que no existe correlación entre las puntuaciones verdaderas de los sujetos en un test y los respectivos errores de medida.
- 3) $\rho(\mathcal{E}_j, \mathcal{E}_k) = 0$, los errores de medida de los sujetos en un test no correlacionan con sus errores de medida en otro test distinto.

A todo lo anterior, existe una extensión de la TCT en el que diversas mediciones del mismo individuo pueden variar tanto por efecto de una variación en lo que se mide como el error de medición. Dicha extensión es conocida como Teoría Generalizada (TG), en esta teoría las decisiones sobre la bondad de un instrumento se basan en estudiar las fuentes y tipos de error, utilizando el análisis de varianza, lo que pretenderá analizar las fuentes de error de medida de un modo sistemático y desglosado, frente al tratamiento globalizado de la teoría clásica.

⁴ Muñiz, José (1998), *Teoría Clásica de los Tests*, Pág. 29.

A partir de la TG, existió un parte aguas en la teoría de los tests, pues se hacen notar dos problemas generales de la TCT, por un lado, la medición de las variables no era independiente del instrumento de medida utilizado, y por otro lado, las propiedades de los instrumentos dependían del tipo de sujetos utilizados para establecerlas, es decir, no eran propiedades puras de los instrumentos, si no que surgían de la interacción de éstos con los sujetos. Dado que la teoría clásica esta orientada hacia todo tipo de test, en esta teoría se produce la paradoja de que un ítem es “fácil” o “difícil”, es decir, puede suceder que, si el test es "difícil", el examinado parecerá tener un nivel bajo de habilidad, si el test es "fácil", el examinado parecerá tener un mayor nivel de habilidad. De esta manera, surge la Teoría de Respuesta al ítem (TRI) que pretende resolver estos problemas que de fondo tiene la TCT.

1.2.2 Teoría de Respuesta al Ítem

La Teoría de Respuesta al Ítem (TRI) constituye un nuevo enfoque en psicometría que permite superar algunas de las limitaciones de la Teoría Clásica de los Tests (TCT). Entre las principales limitaciones de la TCT se pueden exponer las siguientes:

- a) Condición de invarianza: Cuando se llevan a cabo las mediciones de una determinada variable, se esperaría que el resultado no dependa del instrumento utilizado, es decir, el test puede arrojar la medida que sea, pero esta no ha de estar en función del test empleado, por ejemplo: Si se aplica un test de inteligencia A y B a un sujeto en tiempos distintos, la medida que obtenga en ambos debería de ser la misma, sin embargo no siempre o casi nunca es así.
- b) Otra limitación es que en la TCT las propiedades de instrumentos de medida dependen en exceso de los sujetos a los que se les aplican los tests. Por ejemplo, el índice de dificultad de un ítem⁵ dependerá de que el grupo de sujetos utilizado para calcularlo sea competente o no; para el primer caso el test será fácil, y en el segundo será difícil.

⁵ Se entiende por índice de dificultad (ID) de un ítem a la proporción de sujetos que lo aciertan de aquellos que han intentado resolverlo

La TRI pretende abordar los puntos ya vistos en la TCT, desde un enfoque probabilístico, y considera al ítem como unidad básica del test. De acuerdo a Hambleton y colaboradores (1991), las principales características de la TRI como una alternativa a la teoría clásica de los tests son:

1. Las características de los reactivos no dependen del grupo del cuál fueron obtenidas.
2. Los puntajes que describen la habilidad del examinado (sujeto) no dependen del test en su conjunto.
3. El modelo se expresa al nivel del reactivo más que al nivel del test.
4. El modelo no requiere de pruebas paralelas⁶ para determinar el índice de confiabilidad.
5. Provee una medida de la precisión de cada índice de habilidad.

Otra ventaja de la TRI es que los sujetos no necesitan contestar el mismo conjunto de ítems a fin de ser comparados con una misma escala.

En ese sentido de acuerdo con Muñiz (1997) los objetivos generales de la TRI son:

- a) “Proporcionar mediciones de las variables psicológicas y educativas que no estén en función del instrumento utilizado para obtenerlas, es decir, que sean invariantes respecto de los instrumentos de medida utilizados”.
- b) “Disponer de instrumentos de medida cuyas propiedades no dependan de los objetos medidos, que sean invariantes respecto a las personas evaluadas”.

Los postulados básicos de la teoría de la respuesta al ítem son:

⁶ Por el momento entiéndase por pruebas paralelas como la elaboración de dos o más test que tiene la misma finalidad y miden lo mismo, pero que son aplicados a la misma muestra en tiempos distintos.

- a) El resultado de un examinado en un ítem puede ser explicado por un conjunto de factores llamados rasgos latentes o aptitudes que se simbolizan por θ .
- b) La relación entre la respuesta de un sujeto a un ítem y el rasgo latente que subyace puede describirse como una función monótona creciente que recibe el nombre de función característica del ítem o curva característica del ítem (CCI). Esta función especifica que a medida que la aptitud aumenta la probabilidad de una respuesta correcta al ítem también aumenta.
- c) Las estimaciones de la aptitud (θ) obtenidas con distintos ítems sería iguales y las estimaciones de los parámetros de los ítems obtenidos en distintas muestras de examinados serán iguales. Es decir existe la propiedad de invariancia entre los parámetros.

Los supuestos de la TRI son:

1. Unidimensionalidad, sólo puede medirse una habilidad por el reactivo que forma parte del examen.
2. La función característica del reactivo refleja la relación real que hay entre las variables no observables (habilidades) y las variables observables (las respuestas a los reactivos).

Al igual que la TCT, la TRI tiene algunas dificultades, en primera instancia, la TRI suele ser un modelo más complejo con respecto a la parcimonia de la TCT, por ejemplo, cuando se tienen más de dos parámetros en estudio.

En general, la elección de una teoría y otra depende de la cantidad de parámetros en estudio y los niveles de facultad de los test que el investigador diseñe. Sin embargo, para el presente estudio se utilizaran modelos de las dos teorías.⁷

⁷ En el capítulo 2 se mencionan los modelos tanto de la TCT como de la TRI, donde se abunda más sobre la TRI

1.3 MODELOS ESTADÍSTICOS DE CAUSA-EFECTO

En la mayoría de las investigaciones científicas está presente, de forma explícita o implícita, la idea de causalidad (o relación funcional), entendiendo por causalidad al hecho de que A causa a B siempre que sucede A y acontece a B, y nunca sucede B sin antes haber sucedido A (Bisquerra, 1989). En ese sentido, los modelos estadísticos de causa-efecto están integrados en lo que se conoce como análisis causal, que se refiere a un conjunto de estrategias que permiten validar modelos causales elaborados por el investigador basándose en una teoría.

En todas las ciencias básicas, incluso en la psicología, hay interés en establecer relaciones funcionales entre variables importantes para el fenómeno en estudio (Nunnally, 1995), en ese sentido la validez de constructo, que mide atributos psicológicos (atributos abstractos y no observables), es una relación funcional que refleja la consecución causa-efecto del fenómeno, así, la validez de constructo es el proceso de hacer explícita una palabra abstracta en términos de variables observables.

Para lo anterior, los modelos de ecuaciones estructurales, nos permitirán presentar dichas relaciones funcionales, analizando todas las relaciones existentes a la vez, cosa que otras técnicas estadísticas como la regresión múltiple, el análisis multivariante de la varianza, el análisis de discriminante, no lo pueden hacer.

CAPÍTULO 2

MÉTODOS ESTADÍSTICOS PARA LA VALIDACIÓN DE ITEMS

Para que los datos obtenidos con diferentes tipos de instrumentos de medición puedan usarse en situaciones prácticas, estos deberán de satisfacer determinadas condiciones como: la validez, que implica que el instrumento de medida que se está usando para un determinado propósito, es decir, debe medir el rasgo que pretende medir; por otro lado, la fiabilidad (confiabilidad), se refiere a que el instrumento de medidas confiables, de manera que se obtenga el mismo resultado al volver a medir el rasgo bajo condiciones similares del sujeto. En general, la confiabilidad se refiere a la significación y reproducibilidad de los datos. En este capítulo abordaremos ambas condiciones, y explicaremos los métodos para la evaluación de dichas condiciones.

2.1 FIABILIDAD

La fiabilidad de una prueba de rendimiento, se refiere a la capacidad que el instrumento tiene para medir de forma consistente y precisa las características que pretende medir. Con el análisis de fiabilidad también se pretende que el instrumento de medición esté libre de errores de medida.

Podemos hablar de dos tipos de errores: los errores aleatorios que, como tales, no pueden ser controlados y no se pueden predecir; y los errores sistemáticos que son controlables y pueden ser explicados por alguna fuente de variación. De ambos errores los que más interesan a la teoría de la fiabilidad son los errores aleatorios.

La fiabilidad de un instrumento de medida debe de ser entendida de las tres formas siguientes:¹

- a) Aludiendo a la estabilidad temporal de las medidas que proporciona.
- b) Haciendo referencia al grado en que diferentes partes del test miden un rasgo de manera consistente (congruencia o consistencia interna).
- c) Enfatizando el grado de equivalencia entre dos formas paralelas.

¹ Abad, Francisco, *et.al.* (2006), *Introducción a la psicometría*, Pág. 41

La primera parte se refiere al hecho de replicabilidad de la prueba a la misma muestra, es decir, si hoy se aplicase un test a n personas, después de transcurrido un tiempo, volvemos a medir a las mismas n personas en el mismo test, cabe suponer que siendo el test altamente fiable, deberíamos obtener una correlación de Pearson elevada entre ambas mediciones. Para este caso se usa el método del test-retest, mismo que no se abordará en este capítulo, pues esta condición no se cumple para el presente estudio, ya que lo que se pretende medir con el examen son habilidades que ubican a los alumnos en un determinado nivel de inglés del curso, y ya una vez iniciado el curso y transcurrido un determinado tiempo, teóricamente el sujeto habrá adquirido más conocimiento (habilidades), así que es de esperarse que al aplicar el mismo examen a las mismas n personas, muy probablemente su calificación sea más alta a la obtenida inicialmente.

En el segundo inciso, la precisión o fiabilidad de un test se puede entender también como el grado en que diferentes subconjuntos de ítems miden un rasgo o comportamiento homogéneo; es decir, el grado en que covarían, correlacionan o son consistentes entre sí diferentes partes del cuestionario. En un instrumento que incluya distintos factores, se esperaría que aquellos reactivos que remiten al mismo factor estuvieran midiendo también lo mismo y por ende estarían altamente correlacionados.

Lo más usual es obtener la consistencia entre dos mitades del test (método de dos mitades) o entre tantas partes como elementos tenga la prueba (consistencia interna).² Los procedimientos para calcular la congruencia interna podemos dividirlos en dos métodos principales:

- Métodos basados en la división del instrumento en dos mitades
- Métodos basados en la correlación de los ítems (coeficiente alpha de Cronbach)

El último punto nos indica que cuando se diseña un test y una segunda versión del mismo, denominada forma paralela, que intenta evaluar o medir lo mismo que el test original

² Abad, Francisco, *et.al.* (2006), *Introducción a la psicometría*, Pág. 43

pero con diferentes ítems, se pretende que la media y la varianza sean probabilísticamente similares.

2.2 MEDIDAS DE FIABILIDAD

Las medidas de fiabilidad intentan cuantificar qué cantidad de la variabilidad de las mediciones obtenidas en una escala se debe a la variabilidad de las puntuaciones verdaderas y qué cantidad se debe a los errores de medida. De acuerdo a la teoría clásica de los test (Muñiz, 1998), se asume que los errores de medida son independientes de las puntuaciones verdaderas, así la varianza de las puntuaciones observadas puede descomponerse de esta manera:

$$\sigma_x^2 = \sigma_v^2 + \sigma_e^2$$

Donde:

σ_x^2 : Es la varianza de la puntuación empírica obtenida.

σ_v^2 : Es la varianza de la puntuación verdadera.

σ_e^2 : Es la varianza del error de medida.

El coeficiente de fiabilidad se define como la proporción de varianza de las puntuaciones observadas que es atribuible a la variabilidad de las puntuaciones verdaderas.

$$\rho_{xx} = \frac{\sigma_v^2}{\sigma_x^2}$$

Pero

$$\sigma_v^2 = \sigma_x^2 - \sigma_e^2$$

Así que

$$\rho_{xx} = 1 - \sigma_e^2 / \sigma_x^2$$

Suele denominarse índice de fiabilidad ρ_{xv} a la correlación entre las puntuaciones empíricas de un test y las verdaderas, siendo igual a la raíz cuadrada del coeficiente de fiabilidad³

$$\rho_{xv} = \frac{\sigma_v}{\sigma_x}$$

Y el error típico de medida está dado por:

$$\sigma_e = \sigma_x \sqrt{1 - \rho_{xx}}$$

2.3 MODELOS DE FIABILIDAD (TCT)

2.3.1 Modelo Alfa (α , alpha de Cronbach)

Como se mencionó anteriormente, dentro de los métodos basados en la covarianza de los ítems, el más utilizado es el coeficiente alpha de Cronbach. El modelo de consistencia interna de Cronbach asume que el instrumento de medida está compuesto por elementos homogéneos que miden la misma característica y que la consistencia interna del instrumento puede evaluarse mediante la correlación existente entre todos sus elementos.

Algo que es fundamental al interpretar el coeficiente alpha es que el valor de éste, se ve afectado por el número de ítems. De manera que entre dos tests que tuvieran la misma confiabilidad, el que tuviera menos ítems daría un menor coeficiente alpha. Esto se debe a que al aumentar la longitud del instrumento la varianza verdadera aumenta en mayor proporción que la varianza de error.

El coeficiente de alfa depende del número de elementos del instrumento de medición (n) y del cociente entre la covarianza promedio de los elementos y su varianza promedio, α

³ Muñiz, José (1998), *Teoría clásica de los tests*, Pág. 35

refleja el grado en que covarían los ítems que constituyen el test, es, por tanto, un indicador de la consistencia interna del test. Su fórmula viene dada por:⁴

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_x^2} \right)$$

Donde:

n : Número de ítems del test

$\sum \sigma_j^2$: Suma de varianzas de los n ítems

σ_x^2 : Varianza de las puntuaciones del test

En cuanto a cuál debe ser el valor mínimo del coeficiente alpha, no hay un criterio uniforme. Podríamos decir que un parámetro aceptable es un mínimo de .70, pero esto hay que tomarlo con precaución, ya que habría que considerar el número de ítems del instrumento. Si su longitud es reducida, valores menores a .70, podrían estar indicando una buena confiabilidad y si su longitud fuera extensa, un valor de .70 podría no ser tan buen indicador de confiabilidad.

De manera que, posiblemente, lo más recomendable al interpretar el coeficiente de confiabilidad sea considerar cuan próximo está al valor máximo (uno), evaluar el margen de error y considerar las distintas circunstancias que involucren el uso del instrumento y, fundamentalmente, tener en cuenta el número de ítems que contenga el test.

2.3.2 Modelo de dos mitades

A través de este método se verifica la consistencia interna de una prueba, es decir, el grado en que diferentes partes del test miden la misma variable.

⁴ Muñiz, José (1998), *Teoría clásica de los tests*, Pág. 54

El modelo de dos mitades asume que el test está constituido por dos partes de igual longitud. Ambas mitades pueden sumarse para obtener la puntuación total en el test. En general, este método es útil cuando se aplica el método del test-retest y se desea evaluar la estabilidad de las medidas entre ambas mediciones; también es muy usada cuando se tienen dos o más formas paralelas de la misma escala y se desea saber si realmente son dos escalas equivalentes.

El cálculo de un coeficiente de confiabilidad de división por mitades por lo general implica tres pasos.⁵

1. Dividir la prueba en mitades equivalentes: si los elementos del test están ordenados por sus nivel de dificultad, es recomendable seleccionar primero los elementos pares y después los impares.⁶
2. Calcular una r de Pearson entre las puntuaciones en las dos mitades de la prueba.
3. Ajustar la confiabilidad de la mitad de la prueba.

Para el paso tres se proponen tres coeficientes, la elección de cada uno de ellos, depende del escenario en el que se obtuvieron los datos. El primero de ellos es el coeficiente de correlación entre las puntuaciones totales de las dos mitades X_1 y X_2 .

$$r_{X_1X_2} = \frac{(\sigma_X^2 - \sigma_{X_1}^2 - \sigma_{X_2}^2)/2}{\sigma_{X_1}^2 \sigma_{X_2}^2}$$

Un segundo coeficiente es el estadístico de dos mitades de Guttman, que sirve para evaluar la fiabilidad de la escala total, asumiendo que la varianza de todos los elementos son iguales.

$$r_{Guttman} = \frac{2(\sigma_X^2 - \sigma_{X_1}^2 - \sigma_{X_2}^2)}{\sigma_X^2}$$

⁵ Cohen, Ronald, et al. (2000). *Pruebas y evaluación psicológicas*. Pág. 161

⁶ Abad, Francisco, et.al. (2006), *Introducción a la psicometría*, Pág. 43

Por último tenemos el coeficiente de Spearman-Brown para la fiabilidad del test total (el test de doble longitud), asumiendo que los sub-test son de igual longitud, tiene la misma fiabilidad y que sus varianzas son iguales.

$$r_{\text{Spearman-Brown}} = \frac{2r_{X_1X_2}}{1 + r_{X_1X_2}}$$

2.3.3 Modelo de formas equivalentes

A través de este método, también llamado de formas alternativas o paralelas, se puede evaluar, tanto la consistencia interna como la estabilidad temporal de un conjunto de puntuaciones.

Formas equivalentes se refiere al hecho de tener dos o más test que miden el mismo rasgo pero que son distintos en cuanto a los ítems que los conforman. Esta prueba parte del hecho que debe comprobarse que las formas alternas son efectivamente paralelas o equivalentes. Para que las pruebas sean equivalentes, deben de cumplir ciertos requisitos: haber sido construidas con las mismas tablas de especificaciones, tener medias y desviaciones estándar semejantes, tener coeficientes de correlación elevados entre ambas formas, y que cada una de las formas incluya una muestra representativa de ítems que mida la misma variable.

Haciendo uso de los procedimientos empleados en estadística inferencial, si disponemos de datos muestrales podemos realizar los contrastes oportunos para determinar, con cierta probabilidad, si dos formas son o no paralelas.

Para muestras relacionadas, el contraste sobre diferencia de medias se plantea como:

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

El estadístico de contraste⁷ es:

$$T = \frac{\bar{Y} - \mu_D}{S_D / \sqrt{n}}$$

Este estadístico se distribuye de acuerdo al modelo t de Student con n-1 grados de libertad que permite conocer la probabilidad asociada a los diferentes valores de \bar{Y}_D que se obtiene en muestras aleatorias de tamaño n.

El contraste sobre diferencia de varianzas, para muestras relacionadas, se plantea como:

$$H_0 : \sigma_1 - \sigma_2 = 0$$

$$H_1 : \sigma_1 - \sigma_2 \neq 0$$

Siendo el estadístico de contraste:

$$T = \frac{(S_1^2 - S_2^2) \sqrt{N - 2}}{2S_1 S_2 \sqrt{1 - r_{1,2}^2}}$$

Que sigue la distribución "t" con $N - 2$ grados de libertad, donde $r_{1,2}$ es la correlación de Pearson entre X_1 y X_2 .

Existen otros métodos para la evaluación de la confiabilidad de los test, pero no son aplicables al presente caso de estudio, tales como lo es el método del test-retest.

2.4 MODELOS DE FIABILIDAD (TRI)

En el capítulo 1 se dio una breve introducción a la teoría de respuesta al ítem, sin abundar mucho sobre sus conceptos fundamentales. En esta sección se trata de exponer dichos conceptos.

⁷ Pardo Antonio (2002). *SPSS 11 Guía para el análisis de datos*. Pág.266.

Recordemos que la Teoría de Respuesta al Ítem (TRI) constituye un enfoque en psicometría que permite superar algunas de las limitaciones teóricas de la Teoría Clásica de los Tests (TCT), principalmente el hecho de la invarianza del test. Pues bien, para poder atender dicho problema, la TRI incorpora un enfoque probabilístico que permite resolver dicho problema.

2.4.1 Supuestos de la TRI

En el capítulo 1 se mencionaron los supuestos de la TRI, mismos que retomamos en esta sección y describimos a continuación.

2.4.1.1 La curva característica del ítem (CCI)

Indica la probabilidad que tienen de acertarlo las personas que se enfrentan a él. Esta probabilidad depende, naturalmente, de cual sea el nivel de la persona en la variable medida.

Las CCI más habituales adoptan la forma de S como las de la figura 1.1 En el eje “x” aparecen los valores de la variable de medida ($-\infty < \theta < \infty$) y en eje de las ordenadas aparece la probabilidad de acertar el ítem. Cabe señalar que las puntuaciones de θ no son las puntuaciones de los sujetos en el test, si no una estimación de dichas puntuaciones (como se vera más adelante).

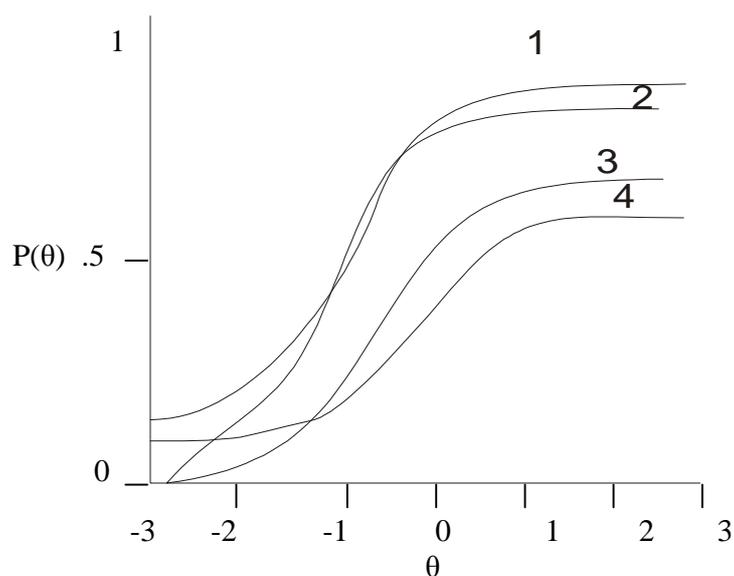


Fig. 1.1 Curva característica con diferentes valores de los

2.4.1.2 Parámetros

- **Parámetro a:** Se denomina *índice de discriminación*⁸ y su valor es proporcional a la pendiente de la recta tangente a la CCI en su punto de máxima pendiente, como se puede apreciar en la figura 1.2. Cuanto mayor sea la pendiente, obviamente mayor será el poder de discriminación del ítem. Cuando la variable θ sigue una distribución $N(0,1)$, el valor de a se puede aproximar a:

$$a \equiv \frac{r_b}{\sqrt{1-(r_b)^2}}, r_b$$

Que es el índice de discriminación clásico.

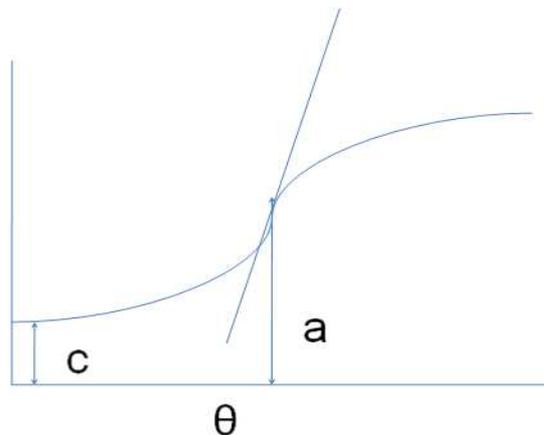


Figura 1.2 Muestra los parámetros

- **Parámetro b:** Es el *índice de dificultad del ítem* y su valor está dado por θ para el punto de máxima pendiente de la CCI. Su valor aproximado cuando θ es $N(0,1)$ se define como:

$$b \equiv \frac{-Z_p}{r_b},$$

⁸ Significa que distingue entre los que saben y no saben

Z_p es la puntuación típica correspondiente al índice de dificultad

- **Parámetro c:** Es la probabilidad de acertar al azar determinado test, es decir, expresa la oportunidad de acertar un ítem cuando se desconoce la respuesta correcta.

2.4.1.3 Unidimensionalidad

La unidimensionalidad se refiere a que sólo puede medirse una habilidad por el reactivo que forma parte del test, por ejemplo, supongamos que para el presente estudio se construye un ítem que mida el nivel de vocabulario de inglés, la probabilidad de acertarlo depende de los valores "a", "b" y "c" del ítem y del nivel de vocabulario inglés de la persona (θ). La CCI excluye que el rendimiento en el ítem dependa de los niveles de la persona en otros rasgos más o menos relacionados con el de vocabulario de inglés (como el nivel de inglés hablado, el nivel de gramática inglesa,...). De tener en consideración otros rasgos, en la fórmula aparecerían los elementos θ_2 , θ_3 , etc..., es decir, los niveles de la persona en esos otros rasgos. En otras palabras, el rendimiento en un ítem depende del nivel de la persona en un sólo rasgo o dimensión.

Este supuesto nunca se cumple totalmente porque el rendimiento en un test está influido por variables cognitivas y de personalidad, como la motivación, ansiedad, etc. Por lo que en la práctica es una cuestión de grado, y no puede afirmarse categóricamente si un conjunto de ítems son o no unidimensionales. No obstante hay varios métodos para comprobar la unidimensionalidad. El propuesto por Reckase (1979) se basa en el análisis factorial y consiste en estudiar la varianza explicada por el primer factor extraído de la matriz de correlaciones entre ítems. En la práctica, cuando el primer factor explica más de un 25% de la varianza total, tras haber eliminado los ítems con saturaciones inferiores a 0.10, se considera que se cumple el supuesto de unidimensionalidad.

2.4.1.4 Independencia Local

Existe independencia local entre los ítems de un test si la respuesta que una persona da a uno de ellos, no depende de las respuestas que da a los otros.

La independencia local se deriva de la unidimensionalidad porque, simplemente, significa que la respuesta a un ítem sólo depende de sus parámetros y de θ , y no está influida por el orden de presentación de los ítems, las respuestas que ya se hayan dado, etc.

Matemáticamente puede expresarse diciendo que la probabilidad de que un sujeto acierte " n " ítems es igual al producto de las probabilidades de que acierte cada uno de ellos por separado.

$$P(U_1, U_2, \dots, U_n / \theta) = \prod_{i=1}^n P(U_i / \theta)$$

2.4.2 Modelos

La mayoría de los estudios de la TRI se han centrado en dos tipos de funciones para la CCI, la función logística y la función normal acumulada, donde por principios de parcimonia y dado a la extensa gama de paquetes computacionales para la estimación de parámetros, se han desarrollado modelos logísticos de unos, dos y tres parámetros. En los tres casos se asume que la respuesta a los ítems es dicotómica, es decir o se acierta o se falla el ítem, independientemente del número de alternativas que se tienen para cada ítem.

2.4.2.1 Modelo Logístico de un parámetro

Este es el modelo más simple de todos. Se le llama también modelo de Rasch. La probabilidad de acertar un ítem depende solamente del nivel de dificultad de dicho ítem y del nivel del sujeto en la variable medida (nivel de rasgo o habilidad).

La fórmula general de la función logística está dada por:

$$Y = \frac{e^x}{1 + e^x}$$

La función anterior, tiene la ventaja con respecto a la curva normal, de ser más tratable. Mediante el uso de una constante ($D=1.7$) en la función logística, los valores de ésta se aproximan a los de una Normal acumulada. Así el modelo Rasch tiene la siguiente estructura

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}}$$

Donde:

$P(\theta)$: Probabilidad de acertar el ítem si el nivel de rasgo es θ .

θ : Nivel de habilidad del sujeto(valores de la variable medida).

b_i : Índice de dificultad del ítem.

e : Base de los logaritmos neperianos (2.718)

D : Constante ($D = 1.7$)

2.4.2.2 Modelo Logístico de dos parámetros

Este modelo añade al anterior un segundo parámetro (a_i) que indica la capacidad discriminativa del ítem, y cuya expresión matemática es la siguiente:

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

Normalmente los valores de "a" oscilan entre 0.3 y 2.5, y se suelen considerar ítems "discriminativos" los que tienen valores "a" mayores de uno.

2.4.2.3 Modelo logístico de tres parámetros

Este modelo añade el *parámetro c*, que como ya se mencionó, es la probabilidad de acertar un ítem cuando no se conoce la respuesta correcta, esto es $P(\theta)$ cuando $\theta \rightarrow -\infty$

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} .$$

2.4.3 Estimación de los parámetros

Una vez seleccionado un modelo, hay que aplicar el test a una muestra amplia y estimar los parámetros de cada ítem y de la variable de medida θ de cada sujeto, a partir de la matriz de respuestas obtenidas. La estimación de parámetros es el paso que nos permite llegar de las respuestas conocidas de las personas a los valores desconocidos de los parámetros de los ítems.

El método más usual para la estimación es el de Máxima verosimilitud, sin embargo, como se sabe este proceso se realiza por medio de aproximaciones sucesivas, lo cual implica realizar bastantes iteraciones, y su cálculo resulta ser muy laborioso. Pues bien, para evitar todos estos cálculos “manuales”, se han desarrollado programas computacionales que facilitan estos cálculos, destacando los programas BICAL, XCALIBRE, RASCAL, ASCAL entre otros.⁹

2.4.4 Ajuste del Modelo

El ajuste del modelo, nos dirá que tan bueno es el modelo propuesto, en contraste con los datos obtenidos a través de un test, es decir, si no hay diferencias importantes entre los resultados pronosticados por el modelo y los obtenidos empíricamente, al mismo tiempo que el modelo cumple con los objetivos generales de la TRI expuestos anteriormente. Para ello Muñiz (1998) propone las pruebas chi-cuadrado, el análisis de los residuos y la comparación de las distribuciones de las puntuaciones.

⁹ Para el presente estudio, se hizo uso del software XCALIBRE, que permite la estimación de tres parámetros.

➤ **Pruebas Chi-Cuadrado**

En general, el método consiste en comparar los valores pronosticados por el modelo con los obtenidos empíricamente. El ajuste será perfecto cuando ambos valores coincidan, por el contrario, si las diferencias de dichos valores aumentan, entonces el ajuste es peor. Lo que nos indica Chi-cuadrado es si esas diferencias son estadísticamente significativas.

$$\chi^2 = \sum \frac{n_j [P(\theta_j) - P_e(\theta_j)]^2}{[P(\theta_j)][1 - P(\theta_j)]}$$

Donde:

χ^2 : Se distribuye con $k-p$ grados de libertad, siendo K el número de categorías en el que se divide θ y p es el número de parámetros.

n_j : Número de sujetos dentro de cada categoría.

$P(\theta_j)$: Valor de la CCI dado por la fórmula del modelo con los parámetros estimados para la categoría j

$P_e(\theta_j)$: Proporción de sujetos que de hecho superan al ítem dentro de cada categoría j .

La prueba anterior, es para cuando se está evaluando el ajuste de un ítem, sin embargo si se quisiera tener una idea general de todo el test, el estadístico a usar sería el siguiente.

$$\chi^2 = \sum_{j=1}^k \sum_{i=1}^n n_j [P(\theta_{ji}) - P_e(\theta_{ji})]^2 / [P(\theta_{ji})][1 - P(\theta_{ji})]$$

Donde χ^2 se distribuye con $(k - p)(n - 1)$ grados de libertad.

➤ **Análisis de los residuos**

Al igual que en el caso anterior, hay que dividir a θ en varias categorías, que generalmente son entre 10 y 15, y calcular para cada una de ellas el residuo estandarizado.

$$RE = \frac{P(\theta_j) - P_e(\theta_j)}{\sqrt{P(\theta_j)Q(\theta_j)/n_j}}$$

Donde

n_j : Número de sujetos dentro de cada categoría.

$P(\theta_j)$: Valor de la CCI dado por la fórmula del modelo con los parámetros estimados para la categoría j

$P_e(\theta_j)$: Proporción de sujetos que de hecho superan al ítem dentro de cada categoría j

$$Q(\theta_j) = 1 - p(\theta_j)$$

Cada vez que los residuos en valor absoluto se alejen más de cero, peor es el ajuste del modelo.

2.4.5 Invarianza de las estimaciones y función de información

La invarianza de la puntuación de θ de cada sujetos en distintos tests, que midan la misma variable, puede comprobarse por medio de la correlación de las puntuaciones obtenidas en los diversos tests. Si al hacer esta correlación, resulta elevada y lineal, podemos asumir dicha invarianza. En otras palabras, la gráfica de correlación debe acercarse a una línea recta y los valores de θ no deben de coincidir en los distintos tests.

Por otra parte, una vez aplicado un conjunto de ítems y estimado el nivel de habilidad de un sujeto, la TRI nos permite calcular el error típico de estimación (Se) de esa persona en

el test aplicado. Esto es una diferencia fundamental con la TCT, que asume que el error es el mismo para todos los sujetos.

El error típico de estimación nos dice la precisión con que hemos estimado θ . A mayor error, menos precisión. Su tamaño depende de varios factores:¹⁰

- 1 . Número de ítems aplicado: En general, al aumentar la longitud del test disminuye Se .
- 2 . La capacidad discriminativa de los ítems: Al aumentar el parámetro "a" disminuye Se .
- 3 . La diferencia entre "b" y θ : Cuanto más próximo a θ esté el índice de dificultad de los ítems (b), menor será Se .

La varianza de las puntuaciones θ estimadas, $Var(\theta)$, se obtiene mediante la expresión siguiente:

$$Var(\theta) = \sum_{i=1}^n \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} .$$

Donde:

n : Número de ítems en el test

$P_i(\theta)$: Valores de la CCI de los ítems

$Q_i(\theta)$: $1 - P_i(\theta)$

$P'_i(\theta)$: Derivada de $P_i(\theta)$

El error típico de medida es la raíz cuadrada de la parte derecha de la fórmula anterior. El error típico de estimación permite obtener el intervalo de confianza en el que, con una probabilidad predeterminada, se ha de encontrar el nivel de habilidad de la persona. En concreto, si a la " θ " estimada de una persona le sumamos y restamos (1.96) multiplicado por

¹⁰ Abad, Francisco, *et.al.* (2006), *Introducción a la psicometría*, Pág. 143

Se, obtenemos los extremos del intervalo en el que, con probabilidad 0.95, se encontrará su verdadero nivel de rasgo.

2.5 VALIDEZ

Una cosa es que el test mida de manera precisa o estable (esta cualidad se refiere a su fiabilidad), y otra diferente es la cuestión de qué es lo que auténticamente está evaluando. En un sentido general, un instrumento de medición es valido si cumple satisfactoriamente con el propósito para el cual fue construido (Nunnally, 1995).

La validación es un proceso continuo, que incluye procedimientos diferentes para comprobar si el cuestionario mide realmente lo que dice medir. Dicho de otro modo, tiene que ver con el tipo de conclusiones o inferencias que pueden realizarse a partir de las puntuaciones obtenidas en el test.

Existen diferentes tipos de validez, dentro de los mas utilizados encontramos las siguientes: validez de constructo, validez de criterio y validez de contenido. Es importante tener en cuenta que distintas formas de validez, pueden ser adecuadas para propósitos diferentes y que cada una permite inferencias distintas, que no pueden cambiarse entre sí.

2.5.1 Validez de Contenido

La validez de contenido en un test se refiere a la correspondencia que existe entre el rasgo que se pretende medir y el contenido de la muestra de ítems que compone al test. Un test presenta validez de contenido si los ítems que lo integran tienen relevancia y además representan al rasgo en cuestión.

Un ítem tiene relevancia si permite recavar la información relacionada directamente con el rasgo que pretende medir. Por otro lado, los ítems de un test son representativos si permiten recoger información sobre todos los aspectos que se desea evaluar, de forma que se cubra todo lo que se pretende medir.

La validez de contenido, estará avalada, por personas expertas en el tema, quienes deberán de ser capaces de detectar problemas relacionados con la falta de relevancia y representatividad de los ítems.

Uno de los métodos para la cuantificación de la validez de contenido es el propuesto por Lawshe(1975), quien propuso una fórmula simple para cuantificar la validez pidiendo a un grupo de expertos responder a la pregunta ¿La habilidad o conocimiento medido por este reactivo es: a) esencial b) útil pero no esencial c) no necesaria? Para cada reactivo se anota el número de miembros del jurado que afirma que el reactivo es esencial. La fórmula es la siguiente.

$$CVR = \frac{n_e - N/2}{N/2}$$

Donde CVR es la razón de validez de contenido, n_e es el número de expertos que indican esencial y N el número total de expertos.

Así pues, el tipo de fundamento para la validez de contenido es más teórico y lógico que empírico, puesto que se evalúa la margen de las repuestas emitidas por una muestra representativa de personas.

2.5.2 Validez de criterio

La validez de criterio persigue un fin más práctico, no se trata únicamente de establecer que se mida adecuadamente un constructo, sino fundamentalmente relacionar las puntuaciones del instrumento con otras variables a las que comúnmente se les denomina variables criterio.

Dentro de la validez de criterio se habla de validez concurrente y validez predictiva. La diferencia entre ambas formas de validez, radica en la temporalidad del criterio. Si las puntuaciones del test se utilizan para predecir alguna medida del criterio que se va a realizar a

futuro, sería validez predictiva. Si por el contrario relacionamos las puntuaciones del test con alguna medida del criterio tomada en el mismo momento sería validez concurrente.

2.5.3 Validez de Constructo

Gran parte de las variables manejadas en el ámbito educativo, son variables hipotéticas no observables directamente, que forman parte de las diversas teorías que tratan de explicar la conducta humana, dichas variables son conocidas como constructos. Estos constructos al no poder ser observados directamente, deben ser operacionalizados, es decir, deben indicarse los procedimientos de medida para observar el constructo.

Los instrumentos psicométricos se refieren a constructos hipotéticos, siendo el instrumento una forma de operacionalizar los mismos. De esta manera la validez de constructo, consiste en tratar de probar que las conductas que registra el test, pueden ser consideradas indicadores válidos del constructo al cual refieren. En otras palabras, la validación de un test implica la obtención de pruebas a favor de la existencia del constructo de interés, así como la demostración de que el test es un instrumento adecuado para medir dicho constructo.

Para la validez de constructo se cuentan con diversas herramientas estadísticas, destacando: Análisis Factorial, Modelos de Ecuaciones Estructurales y la Matriz Multirasgo-Multimétodo. Donde para el presente estudio nos enfocaremos principalmente a los Modelos de Ecuaciones estructurales, mismo que serán abordados en el siguiente capítulo.

CAPÍTULO 3

TÉCNICAS MULTIVARIABLES Y MODELOS CAUSALES

3.1 INTRODUCCIÓN

A menudo cuando se realiza una investigación se intenta dar una explicación de la realidad, la cual involucra relaciones funcionales (causales) que el investigador estudia. Estos procesos involucran relaciones de causa-efecto entre variables manifiestas y variables latentes. Por ejemplo, la depresión se “construye” a partir de la falta de apetito, el reporte de tristeza, baja actividad corporal y el exceso de sueño¹.

Para explicar dichos fenómenos causales como el anterior se construye un sistema de ecuaciones donde las causas son representadas por variables independientes, mismas que pueden ser variables latentes o manifiestas.

Las relaciones causales pueden ir desde las más simples como lo es un modelo causal de depresión hasta las más complejas como lo son modelos para la explicación accidentes en una empresa. Estas relaciones se analizan a través de técnicas estadísticas que nos permiten aumentar la capacidad explicativa y la eficacia estadística del modelo propuesto por el investigador. Así, si en una relación causal de un fenómeno que involucra varias causas, que repercuten en un solo efecto, el método estadístico a emplear podría ser el análisis de regresión múltiple. Sin embargo, si el número de causas del fenómeno fueran tan grandes, tal vez lo más conveniente sería resumir el número de causas o factores antes de plantear las relaciones causales, para esto el análisis factorial ofrece una técnica adecuada. Ahora bien, si se tuviese de ante mano un modelo empírico apoyado en investigaciones previas sobre los factores y sus relaciones, de tal forma que se puedan expresar en un diagrama Path, entonces se puede explicar el fenómeno por medio de un modelo de ecuaciones estructurales.

3.2 MODELO CAUSAL

Los modelos causales también conocidos como modelos estructurales, sistemas de ecuaciones simultáneas o modelos de ecuaciones estructurales, tienen como objetivo explicar las posibles relaciones causa-efecto de un conjunto de variables que explican un fenómeno. Aun cuando

¹ Corral Verdugo, Víctor., et al. *Análisis cuantitativo de variables latentes*, Pág. 15.

los modelos causales no son técnicas puramente estadísticas, emplean a la estadística para la estimación de sus parámetros.

Todo modelo causal debe de cumplir las siguientes características²:

1. El modelo causal debe de contener todas las variables que definen perfectamente el comportamiento del fenómeno en estudio.
2. Cuando se incluyan las variables en el modelo, debe de hacerse tal y como funcionan en la realidad.
3. La construcción del modelo con sus respectivas ecuaciones determinará el tipo de datos que hay que recoger y el método por el cual serán analizados. Las ecuaciones se construyen a partir del funcionamiento y antecedentes del fenómeno en estudio.

Las variables que involucra un modelo causal pueden ser variables medidas o variables latentes. Las variables medidas, también llamadas variables observadas, manifiestas o indicadores, son variables que pueden ser observadas y medidas directamente al aplicar algún instrumento de medida. Las variables latentes son variables que no pueden ser observadas directamente, y tienen que ser inferidas a partir de sus efectos en las variables observadas. Las variables latentes reciben también el nombre de constructos, factores o simplemente variables no observadas.

Ambos tipos de variables, observadas y latentes, pueden ser a su vez endógenas y exógenas. Las *variables endógenas* son variables dependientes que vienen explicadas por otras variables incluidas en el modelo y su variabilidad pueden ser a su vez otras variables endógenas o exógenas. Las *variables exógenas* no están explicadas por otras variables incluidas en el modelo, y su variabilidad se atribuye a causas externas del mismo. Así cada tipo de variable se representa como sigue³:

² Álvaro Page, Mariano. *Hacia un modelo causal del rendimiento académico*, Pág. 151.

³ Mangin Lévy, Jean-Pierre., et al. *Análisis multivariable para las ciencias sociales*. Pág. 771.

- $\xi(xi)$ Variable latente exógena.
- $\eta(eta)$ Variable latente endógena
- X Variable manifiesta exógena.
- Y Variable manifiesta endógena

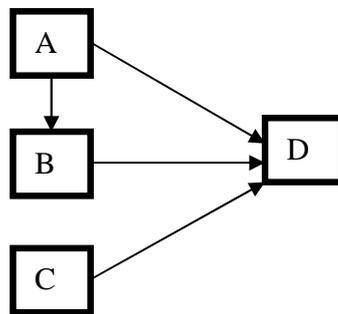


Fig. 3.1. Modelo recursivo para tomar una decisión en D en función de A,B y C.

Cuando en un modelo causal se presentan relaciones unidireccionales el modelo se conoce como recursivo, el cual podemos ver en la figura 3.1; si presenta relaciones bidireccionales se denomina no recursivo (fig 3.2). Los modelos recursivos pueden ser saturados o no saturados. Un modelo recursivo está saturado cuando se han introducido todos los efectos posibles que no sean recíprocos o bidireccionales (fig 3.3)

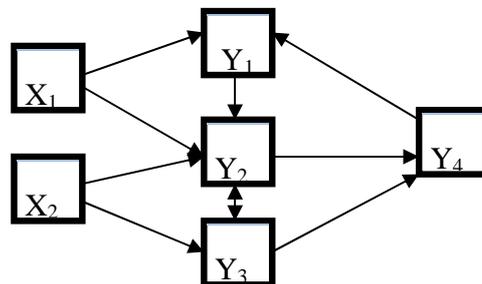


Fig. 3.2. Modelo no recursivo.

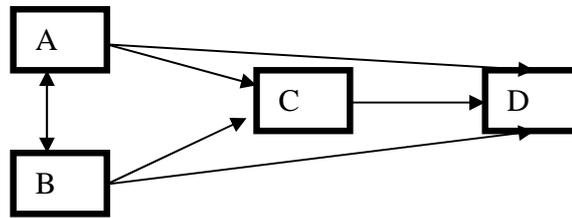


Fig. 3.3. Modelo recursivo saturado

3.2.1 TIPOS DE RELACIONES CAUSALES

Muchas veces los tipos de relaciones causales no solo se generan a partir de las características de los componentes y la forma de la relación, sino que también dependen en gran parte de la objetividad y perspectiva del investigador. En general, podemos destacar dos tipos de fenómenos causales, simples y múltiples. Los procesos simples incluyen sólo un factor que contribuye a la explicación del efecto, mientras que los fenómenos múltiples admiten la presencia de varios factores influyendo en la producción del efecto.

En la figura 3.4⁴ podemos describir otro tipo de relación causal, como se puede observar, por un lado el conjunto de predictores conductuales I,R, IR (información, Reforzamiento y la interacción de estos) predicen a una variable dependiente P (puntaje, en la ejecución de una tarea), por otro lado el predictor cognoscitivo E (Expectancia) también predice a P. Entonces podemos decir que hay un efecto directo de E a P, y que también hay un efecto indirecto de E en P.

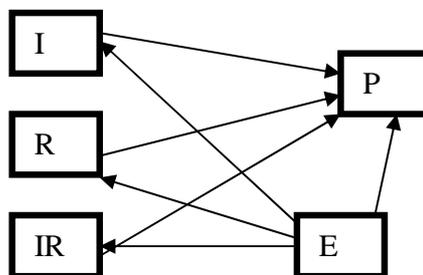


Figura 3.4. Efectos causales directos e indirectos.

⁴ Corral Verdugo, Víctor., et al. *Análisis cuantitativo de variables latentes*, Pág. 27.

3.3 ANÁLISIS DE REGRESIÓN MÚLTIPLE

El análisis de regresión múltiple se basa en el concepto de dependencia estocástica entre variables, de las cuales una será tratada como variable dependiente (variable a explicar) y tendrá carácter cuantitativo, y el resto serán variables independientes utilizadas en la explicación del comportamiento de la independiente. El objetivo de este método estadístico consiste en predecir la única variable dependiente, por medio de las variables independientes, cuyos valores son conocidos.

La siguiente relación muestra la formulación de la regresión múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n .$$

Además de predecir de una variable dependiente dadas varias variables independientes, el análisis de regresión múltiple determina el grado en que las variables independientes contribuyen a la predicción para determinar su importancia.

Es muy importante considerar los supuestos necesarios antes de hacer especificación del modelo de regresión múltiple, ya que dichos supuestos nos permiten garantizar la validez del modelo, de no considerarse muy probablemente se presentarán serios problemas con el tratamiento de los datos que nos llevarían a formular un modelo incorrecto. Dichos supuestos son⁵:

1. *Linealidad.* La media de Y para cada combinación específica de X_1, X_2, \dots, X_n , es una función lineal de X_1, X_2, \dots, X_n ; es decir, la variable dependiente es la suma de un conjunto de elementos: el origen de la recta, una combinación lineal de variables independientes o predictoras (X_1, X_2, \dots, X_n) y los residuos componente de error (E) que refleja la diferencia entre una respuesta observada Y de un individuo y la respuesta promedio verdadera; esto es:

⁵ Kleinbaum, David G. *Applied regression analysis and other multivariate methods*. Págs. 107-109.

$$Y = \beta_0 + \beta_1 + \dots + \beta_n + \varepsilon$$

El incumplimiento del supuesto de linealidad suele denominarse **error de especificación**.

2. *Independencia*. Los residuos ε son independientes entre sí y constituyen una variable aleatoria.
3. *Homocedasticidad*. La varianza de Y ha de ser homogénea para cualquier combinación de X_1, X_2, \dots, X_n .

$$\text{Var}(Y | X_1, X_2, \dots, X_n) = \sigma^2$$

4. *Normalidad*. Para cada combinación de los valores de las variables independientes, la variable Y está normalmente distribuida.

$$Y \sim N(\mu_{Y|X_1, X_2, \dots, X_n}, \sigma^2)$$

5. *No-Colinealidad*. No existe relación lineal exacta entre ninguna de las variables independientes.

3.4 ANÁLISIS FACTORIAL

El análisis factorial (**AF**) es una técnica de análisis multivariante que se utiliza para la reducción de dimensiones encontrando grupos homogéneos de variables a partir de un conjunto numeroso de variables. Esos grupos homogéneos se forman con las variables que correlacionan mucho entre si, procurando que los grupos sean independientes unos de otros. Así, el **AF** una técnica de reducción de la dimensionalidad de los datos. Su propósito consiste en buscar el número mínimo de dimensiones capaces de explicar el máximo de información contenida en los datos.

Podemos hablar de dos enfoques distintos del análisis factorial; cuando el investigador no tiene una idea clara de cuantas dimensiones subyacentes existen para explicar los datos, entonces el Análisis Factorial Exploratorio (AFE) puede utilizarse como una forma de conocer el número mínimo de factores hipotéticos que pueden resumir la información. Por otra parte, dependiendo de los conocimientos del investigador, el método puede utilizarse para probar hipótesis específicas; donde de acuerdo a estudios previos, se tiene identificado el número de factores. Entonces al uso del AF se le denomina Análisis Factorial Confirmatorio (AFC) y se resuelve con la aplicación de método de Ecuaciones Estructurales.

Para el presente estudio, Se va a emplear el segundo método, sin embargo, para ambos casos (AFE Y AFC), se deben cumplir ciertos requerimientos teóricos que se exponen a continuación.

3.4.1 Requerimientos teóricos

Para poder aplicar el análisis factorial, ya sea el confirmatorio o el exploratorio, es necesario cumplir con ciertos requerimientos.⁶

- Poseer una matriz de correlaciones entre variables. Si se desea analizar la estructura de las relaciones entre las variables debe utilizarse la matriz de correlaciones entre variables; si lo que interesa es identificar sujetos similares, la matriz de correlaciones entre los encuestados es la adecuada para el análisis.
- Seleccionar las variables con escala de medición métrica. El análisis factorial está diseñado para aplicarse en variables métricas, aunque para el caso del AFC se pueden incluir variables ordinales o nominales. Si se está diseñando un estudio para valorar la estructura propuesta deben incluirse tres o más variables por factor propuesto. Si el estudio es exploratorio, debe conservarse un número razonable de variables por factor.

⁶ Hair, J. et al. *Análisis multivariante*. Págs. 86-88.

- Conservar un tamaño muestral adecuado. Algunos autores proponen por lo menos 20 observaciones por cada variable para el caso del AFE, mientras que para el análisis de ecuaciones estructurales se propone una muestra de más de 100 observaciones.

3.4.2 Supuestos

El análisis factorial es una técnica estadística que tiene la capacidad de reducir datos a través de la interrelación de las variables observadas, por lo que se requiere que dichas variables cumplan con determinados supuestos, mismos que de alguna forma determinan las altas correlaciones entre las variables, tales supuestos son: **normalidad, linealidad y homocedasticidad**. De no cumplirse dichos supuestos, las correlaciones que se observen en las variables serán muy bajas.

- **Normalidad.** Cada una de las variables incluidas debe de seguir una distribución normal. Este supuesto se puede verificar por medio de un histograma o de los tests de simetría y de curtosis. Si no se cumple dicho supuesto, se puede aplicar algunas transformaciones como la raíz cuadrada, logaritmo natural, entre otras, que puede permitir que se logre la normalidad.
- **Linealidad.** Se refiere a la dependencia lineal entre las variables medidas por las correlaciones. La linealidad indica que el modelo a contrastar predice los valores de la variable(s) dependiente(s) siempre que se produzca una modificación en las variables independientes. Para comprobar que este supuesto se cumpla basta con analizar los gráficos de dispersión de las variables e intentar identificar si los datos siguen un trazado lineal.
- **Homocedasticidad.** El objetivo que se persigue es conseguir una dispersión por igual de la varianza de la variable dependiente a lo largo de los valores de las variables independientes. Se puede comprobar éste supuesto mediante los gráficos de dispersión, observando que no haya ningún punto que se aleje mucho del resto, otra forma

complementaria, son el test de Levene y el test M de Box, ambos son para observar si la varianza de la variable dependiente permanece constante.

Estos son los supuestos que deben de cumplir tanto el análisis factorial exploratorio (AFE), como el confirmatorio (AFC). Para ambos análisis, la técnica de solución es distinta.

La técnica utilizada para AFE parte de supuestos de la alta correlación entre variables, que indican la adecuación de la aplicación del método. Posteriormente se extraen los factores, por métodos como el de máxima verosimilitud o componentes principales, conjuntamente, se analizan criterios para el número de factores a extraer, finalizando con la aplicación de métodos de rotación de factores, con el fin de hacer más fácil la interpretación de la solución.

Para el Análisis Factorial Confirmatorio (AFC), el método a aplicar es el de ecuaciones estructurales, para el cual se ha dedicado un capítulo de esta investigación. Sin embargo, antes de abordar el capítulo, es necesario tener en cuenta que el AFC y los Modelos de Ecuaciones Estructurales (SEM, Structural Equation Modeling) no son los mismo, los SEM, son modelos complejos que representan múltiples relaciones entre variables dependientes e independientes, latentes y observables, que son analizadas de cualquiera de las siguientes tres estrategias de modelización: *Confirmatoria*, el AFC cae dentro de esta categoría, cuyo propósito es la de confirmar el modelo a través de su significación estadística. *Modelos rivales*, se proponen varios modelos y se contrastan entre si para saber cual de ellos tiene mejor significación estadística. *Desarrollo del modelo*, se propone un modelo y el propósito es mejorarlo a través de modificaciones.

3.5 ANÁLISIS PATH O DE SENDEROS

El análisis de senderos es una técnica estadística multivariada que nos permite representar las relaciones causales entre variables por medio de senderos (grafos), en los cuales a partir de una matriz de correlaciones entre las variables observadas, se pueden estudiar los efectos directos e indirectos de unas variables consideradas como causas sobre otras a las cuales se les consideran efectos.

Para observar un ejemplo simple acerca de la notación básica para la construcción de un diagrama de senderos, observemos la figura 3.5, en donde las variables A y B representan la inteligencia del padre y la madre como causa de la inteligencia del hijo C, además la variable X que representa factores independientes de la inteligencia de los padres que afectan a la inteligencia del hijo. En la figura se aprecian tres flechas con una sola cabeza las cuales representan una relación causal entre dos variables, la variable que se escribe en la cola de la de la flecha es la causa de la variable que se escribe en la cabeza, mientras que la flecha curva bidireccional, se refiere a la correlación simple entre dos variables, generalmente las variables se representan con mayúsculas

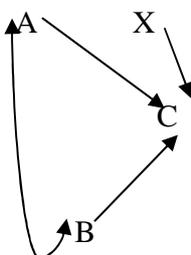


Fig. 3.5. Ejemplo práctico de diagrama de senderos.

Las variables A, X, B, son también variables exógenas (independientes), pues sus fuentes causales son externas al diagrama de senderos, mientras que la variable C es variable endógena (dependiente) ya que tiene sus fuentes causales en el diagrama de senderos por medio de las variables exógenas.

A la hora de construir un diagrama de senderos, se deben de tener en cuenta los siguientes supuestos⁷:

1. Las relaciones son lineales (qué causa qué).
2. Las variables fuente (exógenas) pueden estar conectadas por flechas curvas bidireccionales, pero las descendentes (endógenas), por lo que las variables descendentes no están correlacionadas.
3. El flujo causal es unidireccional.

⁷ Loehlin, Jhon C., *Latent Variable Models. An introduction to factor, Path, and Structural Analysis*. Págs. 5-7.

4. Las variables descendentes pueden tener una flecha residual la cual indica que las causas de variación de esas variables no están completamente explicadas por las variables fuente.
5. Las flechas residuales apuntan a las variables descendentes.
6. Las variables se miden en escalas de intervalo o superiores.

En el diagrama de senderos, también es posible encontrar con efectos directos o indirectos entre sus variables. El efecto directo se representa por una simple flecha causal entre dos variables, mientras que el efecto indirecto se produce cuando por medio de una variable se crea un efecto en dos variables que no están unidas por una flecha causal. En la siguiente Figura 3.6 A es causa indirecta de C por que pasa por D, mientras que B es causa directa de C.

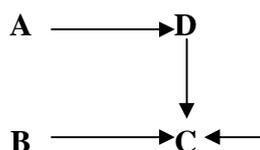


Fig. 3.6. Muestra los efectos directos e indirectos.

El propósito del análisis de senderos es proporcionar una representación gráfica de las relaciones entre variables latentes endógenas a través de sus diferentes indicadores o variables observadas para explicar algún fenómeno en estudio.

La construcción del análisis de senderos suele utilizar los siguientes símbolos⁸:

1. Las variables observadas (X,Y) van dentro de cuadros.
2. Las variables latentes η (eta) y ξ (xi) van dentro de círculos.
3. Los errores de medida δ (delta) y ϵ (epsilon), así como los 'elementos de perturbación' ζ (dseta) se incluyen en el diagrama sin enmarcarlos.
4. Una flecha unidireccional entre dos variables indica una influencia directa de una variable sobre la otra.

⁸ Bisquerra, R. *Introducción conceptual al análisis multivariado. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD. Vol I.* Págs. 490,491.

5. Una flecha bidireccional indica una correlación entre las dos variables sin interpretación causal.
6. Los parámetros asociados a cada flecha se simbolizan de la forma en que se especifica en la tabla.
7. Cada parámetro lleva dos subíndices, el primero corresponde a la variable de llegada de la flecha (efecto) y el segundo a la variable de donde sale (causa). En las flechas bidireccionales los subíndices pueden intercambiarse.
8. Todas las influencias directas de una variable sobre otra deben incluirse en el diagrama de paso. Por tanto, la inexistencia de la flecha entre dos variables significa que se supone que estas dos variables no están en relación directa.

Tabla 3.1 Símbolos de variables y parámetros

Desde	Hasta	Parámetro
ξ (xi)	ξ (psi)	ϕ (phi)
ξ (xi)	η (eta)	γ (gamma)
η (eta)	η (eta)	β (beta)
ξ (xi)	X	λ (lambda)
η (eta)	Y	λ (lambda)
X	X	ϕ (phi)
X	Y	γ (gamma)
Y	Y	β (beta)
ζ (dseta)	ζ (dseta)	ψ (psi)
Errores de medida de las variables exógenas observadas	δ (delta)	$\Theta\delta$ (zeta-delta)
Errores de medida de las variables endógenas observadas	ϵ (epsilon)	$\Theta\epsilon$ (zeta-epsilon)

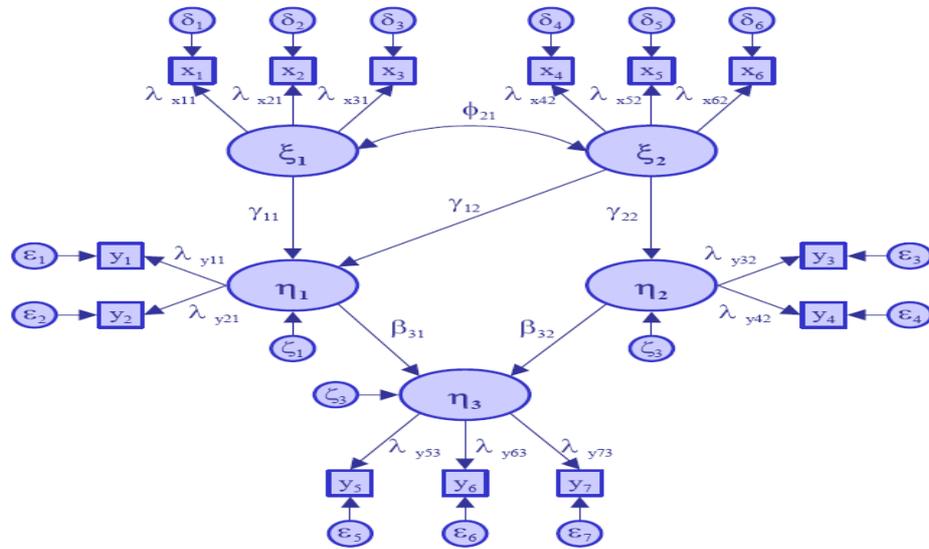


Fig. 3.7 Diagrama de senderos de acuerdo a la notación Lisrel de modelos de ecuaciones estructurales

3.6 MODELO DE ECUACIONES ESTRUCTURALES

El modelo de ecuaciones estructurales nos permitirá examinar un conjunto de relaciones entre una o más variables independientes, sean estas continuas o discretas, y una o más variables dependientes, continuas o discretas". Las variables pueden ser a su vez *variables medidas* o *variables latentes*. Las variables medidas, también llamadas variables observadas o indicadores, son variables que pueden ser observadas y medidas directamente. Las variables latentes son variables que no pueden ser observadas directamente, y tienen que ser inferidas a partir sus efectos en las variables observadas. Las variables latentes reciben también el nombre de constructos, factores o variables no observadas.

El modelamiento de ecuaciones estructurales representa la síntesis de dos aproximaciones estadísticas: Por una parte, el análisis factorial que permite la estimación de constructos o factores, y por otra parte, el modelo de regresión lineal múltiple que mide el grado de relación lineal existente entre variables observadas o manifiestas; esta segunda técnica estadística es la base teórica del modelo de ecuaciones estructurales. Esta ecuación, en el contexto de los modelos causales recibe el nombre de ecuación estructural y sus elementos

se denominan parámetros estructurales, donde estos últimos, pueden ser estimados y su resultado toma el valor de un coeficiente que representa la relación entre las variables. En el caso de una sola variable dependiente, el parámetro estructural tendrá el mismo valor que el coeficiente de correlación. En caso de más de una variable dependiente, los coeficientes de regresión parcial determinan el valor de los parámetros⁹. Así, la formulación básica para el modelo de ecuaciones estructurales es:¹⁰

$$Y_1 = X_{11} + X_{12} + X_{13} + \dots + X_{1n}$$

$$Y_2 = X_{21} + X_{22} + X_{23} + \dots + X_{2n}$$

$$Y_m = X_{m1} + X_{m2} + X_{m3} + \dots + X_{mn}$$

Los modelos de ecuaciones estructurales son una metodología óptima para el análisis de validez de constructo, sobre todo, como en esta investigación, donde el interés es de tipo explicativo.

3.6.1 ETAPAS EN EL MODELAMIENTO DE ECUACIONES ESTRUCTURALES

El objetivo de los modelos de ecuaciones estructurales es encontrar un modelo que se ajuste a los datos empíricos lo suficientemente bien como para servir como una representación útil de la realidad. De esta forma, el comportamiento de las variables observadas podría ser explicado a partir de las relaciones causa-efecto estimadas en el modelo.

Un análisis mediante modelos de ecuaciones estructurales se desarrolla en las siguientes etapas:¹¹ 1) Especificación del modelo, 2) Identificación del modelo, 3) Estimación de parámetros, 4) Evaluación del modelo, 5) Interpretación del modelo.

⁹ Bisquerra, R. *Introducción conceptual al análisis multivariado* Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD. Vol. II. Pág 495.

¹⁰ Hair, J. et al. *Análisis multivariante*. Pág 612.

¹¹ Mangin Lévy, Jean-Pierre., et al. *Análisis multivariable para las ciencias sociales*. Pág. 781

3.6.1.1 Especificación del modelo

Esta etapa corresponde a la formulación de las hipótesis que explicarían el comportamiento de las variables observadas. De acuerdo a la teoría, el investigador identificará las variables latentes del modelo y establecerá las relaciones causa-efecto entre las variables latentes y sus indicadores. La especificación del modelo es la descripción formal de estas relaciones, y su eventual representación en un diagrama. Esta descripción formal incluye la determinación de qué parámetros del modelo estarán fijados a priori, y qué parámetros serán estimados a partir de los datos empíricos. Normalmente, se fijan a cero los coeficientes para establecer la hipótesis de que dos variables no están relacionadas, mientras que los coeficientes entre variables que el investigador supone relacionadas se dejan libres para ser estimados.

Por otra parte, las variables latentes no tienen una métrica definida, por lo que es necesario establecer una escala de medida para cada variable latente. Este aspecto se resuelve usualmente asignando un valor de 1 a la relación entre la variable latente y uno de sus indicadores. Este indicador será el indicador de referencia, y proporciona una escala interpretable para la variable latente con la que está relacionado.

Una vez que se tiene el modelo hipotético representado mediante un diagrama Path, se especifica el modelo matemático de ecuaciones estructurales, el cual, como se dijo anteriormente, considera variables latentes cuantificadas a través de variables observables y relaciones de dependencia con variables de doble función. Para poder construir el sistema de ecuaciones con estas características es necesario especificar en un primer sistema las relaciones de dependencia y en otro los indicadores de las variables latentes. El primer sistema se conoce como **modelo estructural** y el segundo, **modelo de medida**.

La ecuación del modelo estructural es:

$$\eta = \Gamma\xi + \beta\eta + \zeta$$

Donde:

η (eta): es un vector columna del orden $n \times 1$ con las variables latentes endógenas

ζ (zeta): representa el vector de errores de las ecuaciones estructurales

ξ (xi): representa el vector de las variables latentes exógenas de orden $m \times 1$

β (beta): Matriz de $n \times n$ que representa las relaciones de las variables latentes endógenas a variables latentes endógenas. El número de variables latentes endógenas se representa con la letra n .

Γ (gama): Representa las relaciones de las variables latentes exógenas sobre las variables latentes endógenas. El número de las variables latentes exógenas está dado por m . Además de estos elementos, el modelo estructural, también emplea las siguientes matrices para su estimación:

Φ (phi): Matriz cuadrada que representa las correlaciones o covarianzas entre las variables latentes exógenas.

Ψ (psi): correlaciones o correlaciones entre las variables latentes endógenas.

La forma matricial de β , Γ , Φ y ψ , así como la forma vectorial de ξ , η y ζ se representa en la tabla 3.2

Tabla3.2. Matrices y vectores para el modelo estructural

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{pmatrix} \quad \Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nm} \end{pmatrix}$$

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1m} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1} & \phi_{m2} & \dots & \phi_{mm} \end{pmatrix} \quad \Psi = \begin{pmatrix} \psi_{11} & \psi_{12} & \dots & \psi_{1n} \\ \psi_{21} & \psi_{22} & \dots & \psi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{n1} & \psi_{n2} & \dots & \psi_{nn} \end{pmatrix}$$

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix} \quad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \quad \zeta = \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix}$$

Para el modelo de medida se tienen las siguientes ecuaciones:

Constructos exógenos

$$X = \Lambda_x \xi + \delta$$

Constructos endógenos

$$Y = \Lambda_y \eta + \epsilon$$

Donde:

X: representa el vector columna de orden $p \times 1$ de las variables observables independientes

Y: es el vector de variables observables dependientes de orden $q \times 1$

ξ : representa el vector de variables latentes exógenas

η : es el vector de variables latentes endógenas

Λ_x (lambda): es una matriz $p \times m$ de pesos factoriales (o factores de carga λ_{pm}^x) que relacionan las variables manifiestas X con las latentes ξ . Los coeficientes lambda representan las influencias lineales de las variables latentes en los indicadores

Λ_y : es una matriz $q \times n$ de coeficientes factoriales de Y en las variables endógenas latentes η . n identifica el número de η . Esta matriz muestra las relaciones de η a Y, y cada uno de los elementos de la matriz se representa por λ_{qn}^y .

δ y ϵ : son los elementos de las matrices diagonales Θ_δ y Θ_ϵ .

En donde la matriz Θ_δ representa la covarianza de los errores de medida de los indicadores de las variables latentes exógenas y la matriz Θ_ϵ representa la covarianza de los errores de medida de los indicadores de las variables latentes endógenas.

La siguiente tabla (tabla 3.3) muestra los vectores y matrices necesarias para el modelo de medida

Tabla 3.3. Matrices y vectores para el modelo de medida

$$\Lambda_x = \begin{pmatrix} \lambda^x_{11} & \lambda^x_{12} & \dots & \lambda^x_{1m} \\ \lambda^x_{21} & \lambda^x_{22} & \dots & \lambda^x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^x_{p1} & \lambda^x_{p2} & \dots & \lambda^x_{pm} \end{pmatrix} \quad \Lambda_y = \begin{pmatrix} \Lambda^y_{11} & \lambda^y_{12} & \dots & \lambda^y_{1n} \\ \Lambda^y_{21} & \lambda^y_{22} & \dots & \lambda^y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda^y_{p1} & \lambda^y_{p2} & \dots & \lambda^y_{qn} \end{pmatrix}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{pmatrix} \quad \Theta_\delta = \text{diag}(\delta_1, \delta_2, \delta_3, \dots, \delta_p)$$

$$\Theta_\epsilon = \text{diag}(\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_p)$$

3.6.1.2 Identificación del modelo

Si el modelo teórico creado es correcto, se procede a la identificación del modelo, donde es necesario verificar que los parámetros del modelo puedan ser estimados. El modelo está identificado si todos los parámetros lo están¹², es decir, si existe una solución única para cada uno de los parámetros estimados. Determinar si un modelo está identificado debe estudiarse antes de la recolección de datos, en donde se debe comprobar que al menos se dispone para cada parámetro de una expresión algebraica que lo exprese en función de las varianzas y covarianzas muestrales

La identificación del modelo permite preguntarse si existen valores únicos para cada parámetro; para lo cual se calcularán los grados de libertad de cada modelo. Los grados de libertad deben de ser superiores o iguales a cero. La diferencia entre el número de

¹² BATISTA, J.M. y COENDERS, G. (2000). *Modelos de ecuaciones estructurales. Cuadernos de estadística*, 6, La Muralla, Madrid, pág. 68

correlaciones o covarianzas y el número de parámetros a estimar se denomina **grados de libertad**. Los cuales se calculan de la siguiente forma:

$$gl = (p+q)(p+q+1)/2 - t$$

Donde p es el número de variables observables independientes, q el número de variables observable dependientes y t el número de parámetros a estimar.

El modelo se dirá *sobreidentificado* si los grados de libertad son superiores a cero, lo cual sería lo ideal, ya que existirán más ecuaciones que incógnitas y el modelo se puede generalizar, por otro parte, el modelo es *identificado* cuando los grados de libertad son iguales a cero, lo que implica que existe una solución única por que hay igual número de incógnitas y ecuaciones, sin embargo, la solución no tiene gran relevancia, ya que el modelo no se puede generalizar, por último, se dirá que se tiene un modelo *infraidentificado* cuando los grados de libertad son negativos, por lo que el sistema de ecuaciones no tiene solución y en este caso no se puede estimar el modelo.

Por otra parte, cada ecuación del modelo debe de ser diferente de las demás o de cualquier combinación lineal de ellas, ya que se podría caer en un modelo infraidentificado, para este caso, lo que se hace es dar valores de partida a los parámetros a estimar.

3.6.1.3 Estimación de parámetros

Para la estimación de parámetros, necesitamos tener los datos de entrada, los cuales pueden ser: los datos brutos (datos originales) que consisten en una matriz de variables en columna y observaciones en línea; también podemos utilizar la matriz de correlaciones o covarianzas de las distintas variables observables.

Para la estimación de parámetros, el método más usual es el de **máxima verosimilitud**. La estimación mediante el método de máxima verosimilitud requiere que las variables observadas mantengan una distribución normal multivariante, es decir, que se cumpla el

supuesto de normalidad. Por otra parte, cuando no se cumple con dicho supuesto, se puede utilizar el método de **mínimos cuadrados no estandarizados**, en el cual también es posible utilizar mínimos cuadrados ponderados y asintóticamente de distribución libre, esté último quiere decir que no se necesita que la población de interés siga una distribución de probabilidad específica.

3.6.1.4 Evaluación del modelo

Es la etapa en la cual realizamos las pruebas de bondad y ajuste de nuestro modelo, es decir, consiste en comprobar hasta que punto el modelo se ajusta a los datos observados, para lo cual, hay que eliminar en un principio a las estimaciones infractoras¹³, que no son más que coeficientes estimados en los modelos de medida y estructurales que exceden los límites aceptables, una vez eliminadas estas estimaciones, hay que evaluar la bondad de ajuste del modelo en forma global (conjunta) y por separado, por medio de los modelos de medida y estructurales.

Es muy importante considerar a las estimaciones infractoras, pues de ellas depende en gran parte la bondad de ajuste de nuestro modelo. Los casos más comunes de estimaciones infractoras son:¹⁴ 1) varianzas de error negativas o varianzas de error no significativas para cualquier constructo, para este caso, el problema se soluciona fijando las varianzas de los errores infractores en un valor positivo muy pequeño, por ejemplo, 0.005; 2) coeficientes estandarizados que sobrepasan o están muy cerca de 1.0, en este caso la solución es considerar la eliminación de una variable, o bien asegurarnos que las variables latentes se han establecido con una verdadera validez discriminante. 3) errores estándar muy elevados asociados con cualquier coeficiente estimado, podríamos seguir la misma solución del punto dos para tratar este problema.

¹³ Hair, J. et al. *Análisis multivariante*. Pág 637

¹⁴ Ibid, Pág 637

Para las medidas de **ajuste global** tenemos las siguientes Pruebas

⇒ **Ji-Cuadrado χ^2** consiste en probar la hipótesis nula H_0 : el modelo se ajusta contra la hipótesis alternativa H_1 : se rechaza el modelo comparado con el nivel de significancia α con la probabilidad p asociada a un valor de χ^2 . Cuanto más bajo seas el valor absoluto de la ji-cuadrada en función de los grados de libertad, mejor será el ajuste.

⇒ **Índice de bondad del ajuste** (Goodness of Fit Index GFI) es un índice de variabilidad y representa el grado general de ajuste del modelo. Los valores de este índice cercanos a cero indican un ajuste muy pobre, por el contrario, valores muy cercanos a uno indican un ajuste perfecto. El índice viene explicado de la siguiente forma:

$$GFI=1-(\text{tr}(\Sigma^{-1}S-I)^2/\text{tr}(\Sigma^{-1}S))^2$$

Donde:

Σ^{-1} es la inversa de la matriz de varianzas reproducida a partir de los valores estimados.

S la matriz de covarianzas entre las variables observadas

tr (traza) es la suma de los elementos de la diagonal de la matriz que va entre paréntesis

I es la matriz identidad

⇒ **Índice residual de la raíz cuadrada media (RMSR)** este índice representa la raíz cuadrada de la media de los residuos cuadrados, es decir, una media de los residuos entre las matrices observadas y las reproducidas (estimadas). Si este índice fuera próximo a cero los errores entre la matriz observada y la matriz reproducida serían muy bajos, lo que significa que el ajuste es muy bueno.

$$RMSR=[\sum_{i=1}^k \sum_{j=1}^k (S_{ij}-\sigma_{ij})^2/p(p+1)]^{-1/2}$$

Donde

K es el número de variables observables X y Y

S la matriz de covarianzas entre las variables observadas

σ la desviación estándar entre las variables observadas

p es el número de variables observables independiente

⇒ **Gráfica de residuales normalizados** es un buen indicador, en resumen, del ajuste de nuestro modelo, en el eje de las abscisas están los residuales normalizados¹⁵, en el eje de las ordenadas los cuantiles. En la gráfica hay una línea de puntos que está ubicada a 45 grados de los ejes. Los puntos (X) representan los valores estimados, los cuales tienden a formar una línea. Cuando la línea de X tiene una pendiente superior a uno, es decir, es más vertical que la línea de 45 grados, significa que el ajuste es bueno. Si las pendientes se acercan ligeramente a uno, de modo que se superponen a la línea de 45 grados, se tiene un ajuste moderado. Si la pendiente es inferior a la de 45 grados, se tiene un ajuste muy pobre. Por otra parte, si los puntos (X) no representan una línea recta se interpreta como errores de especificación o desviaciones de normalidad.

La fiabilidad y varianza extraída de un constructo deben de calcularse por separado para cada constructo de acuerdo a los indicadores especificados que representan dicho constructo. Así pues, la fiabilidad de constructo se refiere a que dichos indicadores son suficientes para representar al constructo explicando un porcentaje suficiente de varianza. La fiabilidad de un constructo se calcula como:¹⁶

$$\text{Fiabilidad de constructo} = \frac{\left(\sum \text{ponderaciones estandarizadas}\right)^2}{\left(\sum \text{ponderaciones estandarizadas}\right)^2 + \sum \xi_j}$$

Donde la suma de las ponderaciones estandarizadas, son estimadas por los programas LISREL y AMOs; y ξ_j es el error de medida de cada indicador. Valores próximos a 0.7 para este indicador se consideran aceptables.

¹⁵ Los residuales normalizados representan la diferencia entre la correlación original y la estimada a partir de los parámetros del modelo.

¹⁶ Hair, J. et al. *Análisis multivariante*. Pág 637

La varianza extraída se calcula de la siguiente forma:

$$\text{Varianza extraída} = \frac{\left(\sum \text{ponderaciones estandarizadas}^2\right)}{\left(\sum \text{ponderaciones estandarizadas}\right)^2 + \sum \xi_j}$$

En general se sugiere que el valor del indicador de varianza extraída debería de exceder a 0.5 para cada constructo.

3.6.1.5 Interpretación del modelo

Los modelos deben de ser estudiados en función de cierto número de criterios. Una vez diseñado, conceptualizado y ejecutado el modelo, el primer análisis que hay que llevar a cabo es el del ajuste, si el ajuste es aceptable de acuerdo a los índices proporcionados por el programa que se utilice en su momento para la estimación de estos parámetros, entonces no hay diferencia estadísticamente significativa entre el modelo conceptualizado y el teórico. Luego, habrá que identificar si todos los estimadores del modelo son significativos, si alguno de ellos no lo fueran, se podría suprimir la relación entre variables latentes teniendo cuidado de que no se distorsione la teoría. Por ello es importante tener en cuenta tres características¹⁷ :

- 1) *La plausibilidad de un modelo*, se refiere a que el modelo debe de ser congruente con la teoría;
- 2) *La proporción de varianza explicada de las variables endógenas*, ésta debe de ser alta, en caso contrario se puede deber a que se han ignorado errores de medida grandes, se han olvidado variables importantes, la forma de las relaciones funcionales son diferentes de cómo se habían supuesto;
- 3) *La replicabilidad de los resultados*, los efectos causales deben de ser los mismos en distintas muestras y en diferentes espacios de tiempo.

¹⁷ Bisquerra, R. *Introducción conceptual al análisis multivariado* Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD. Vol. II. Pág 547

CAPÍTULO 4

EVALUACIÓN Y VALIDACIÓN DEL
EXÁMEN DE COLOCACIÓN DE UN
CURSO EN LÍNEA PARA LA
COMPRENSIÓN DE LECTURA EN INGLÉS

4.1 PLANTEAMIENTO DEL PROBLEMA

En materia educativa, existe una gran cantidad de técnicas que permiten evaluar los conocimientos que los alumnos tienen antes, durante y al final de un curso, sin embargo una de las técnicas de evaluación más usadas son los exámenes, los cuales deben reflejar en cada uno de sus reactivos y en conjunto, los elementos necesarios para emitir un juicio acerca del nivel de conocimientos del sustentante. En otras palabras, los exámenes deben de cumplir con ciertas características que permitan asegurar la fiabilidad y validez del examen. A manera de ejemplo, entiéndase por validez y fiabilidad al hecho encontrar una respuesta a las siguientes preguntas: ¿el examen mide realmente lo que pretende medir?, ¿el examen tiene precisión?, ¿el examen conduce a mediciones independientes del evaluador o del sujeto?, ¿cuál es el poder discriminativo de la prueba?, ¿cuál es la probabilidad que un sujeto conteste correctamente una pregunta al azar?, y ¿qué tan difícil es el examen?. De esta serie de cuestiones surge la necesidad de esta investigación.

El presente trabajo está integrado a un proyecto PAPIME de la FES Acatlán, el cual tiene como objetivo desarrollar un curso en línea para la comprensión de lectura en inglés e integra exámenes de comprensión de lectura en inglés como lengua extranjera, para medir la habilidad de los estudiantes al elegir y aplicar las estrategias de lectura adecuadas para comprender de manera eficiente textos académicos escritos en esa lengua.

De esta forma, la presente investigación aborda el problema de realizar la validación estadística del examen de ingreso al curso en línea para la comprensión de lectura en inglés, presentando un modelo causal que busca explicar las habilidades que tiene un apersona para elegir y aplicar las estrategias de lectura adecuadas en la comprensión eficiente de textos académicos escritos en inglés, a partir de un conjunto de variables que a priori se asumen como indicadores de dichas habilidades.

Las relaciones hipotetizadas se presentan en un diagrama de senderos, evaluando posteriormente el modelo a través de un análisis de ecuaciones estructurales. Pero antes de abordar el modelo se realiza el análisis de fiabilidad del examen para verificar la consistencia interna, conocer el poder discriminativo de la misma y el nivel de dificultad, que en algún

momento serán de interés y utilidad para hacer inferencias acerca del contenido y diseño del examen.

4.1.1 Objetivo general

- Evaluar y validar el examen de ingreso al curso en línea de comprensión de lectura en inglés.

4.1.2 Objetivos particulares

- Realizar un análisis de fiabilidad
- Trazar un diagrama de senderos para representar las relaciones causales
- Hacer el análisis de ecuaciones estructurales para efectuar la validez de constructo
- Obtener resultados e inferir a cerca del modelo y del análisis de fiabilidad

4.1.3 Marco Contextual

En la Facultad de Estudios Superiores Acatlán (FESA) se imparten 18 licenciaturas, de las cuales sólo 3 son impartidas en el sistema de universidad abierta. El plan de estudios de cada una de las carreras prevé cuando menos el estudio de algún idioma diferente al español, para lo cual el Centro de Enseñanza de Idiomas (CEI) de la misma facultad, ofrece 15 cursos en dos modalidades, comprensión de lectura (CL) y plan global (PG), entre los cuales destaca el idioma inglés, pues para la mayoría de las carreras es requisito de titulación acreditar el idioma en alguna de sus dos modalidades (PG o CL). Sin embargo, dada la saturación de los cursos de comprensión de lectura y la incapacidad para satisfacer la demanda de los alumnos en el CEI, surge la necesidad de la elaboración de un curso en línea de comprensión de lectura en inglés que les permita a los alumnos prepararse para acreditar el examen de requisito, de modo que aparejada a esta necesidad, surge la de evaluar el examen de colocación a esos cursos.

4.1.4 Selección y Descripción de la Muestra

Para este estudio la población se limita a todos aquellos alumnos que estén inscritos en el curso en línea de comprensión de lectura en inglés y que realizaron el examen de colocación.

Para obtener la muestra se publicó a través de la página Web de la FES Acatlán, un comunicado que invitaba a los alumnos de la facultad a participar en el curso en línea para la comprensión de lectura en inglés. En el comunicado se les pedía a los alumnos registrarse, solicitando su nombre, carrera y número de cuenta, también se les daba la dirección del sitio web donde tenía que realizar el examen de colocación. Al final del periodo de inscripción, se consiguió una población de 213 alumnos, muestra sobre la que se hará el análisis estadístico.

Los 213 estudiantes de la FES Acatlán fueron clasificados de la siguiente manera: 99 estudiantes (45.5%) fueron colocados como principiantes, 93 estudiantes (43.7%) colocaron en el nivel intermedio y 21 estudiantes (9.9%) quedaron en el nivel avanzado. Con los datos resultantes del examen se llevó a cabo el análisis de fiabilidad y el modelamiento de ecuaciones estructurales para la validación de constructo.

4.1.5 Instrumento para la recolección de datos

Para realizar este estudio se debe de contar con un instrumento que nos permita obtener información acerca de tres dimensiones: microestructura, macroestructura y procesamiento mental, que representan la habilidad de los estudiantes para comprender textos en lengua inglesa¹

Para poder medir dichas dimensiones se diseña un examen de colocación², el cual es publicado a través de una página web³ desarrollada mediante el software libre moodle. Dicha página consta de un tutorial, donde el usuario es entrenado para hacer uso de la página, en caso de que éste no este familiarizado con el internet.

¹ García Jurado, R. (2003). Factores instruccionales que inciden en el aprendizaje de la comprensión de lectura en inglés como lengua extranjera en alumnos de la UNAM. Pag.151

² Realizado por Raquel García Jurado Valverde.

³ <https://www.andromeda.acatlan.unam.mx>

Para acceder al examen, el alumno debe estar previamente registrado y contar nombre de usuario y clave, que le son asignados cuando éste se registra o inscribe para el examen. Al alumno se le pide la clave y el nombre de usuario para poder contestar el examen; la página valida que el alumno no conteste dos veces el examen.

El examen de colocación consta de un Texto descriptivo en lengua inglesa (ver anexo A) y de un examen de diagnóstico de estrategias de lectura en inglés.

El texto de tipo descriptivo, en lengua inglesa llamado “21st Century Slaves”, explora el tema de las diferentes formas de esclavitud en el siglo XXI que existen en el mundo. Es un texto narrativo que consta de 650 palabras y forma parte del artículo de portada de la Revista National Geographic en inglés, del mes de Septiembre del 2003.

Por otra parte el examen de colocación consta de 15 reactivos de opción múltiple que comprenden tres dimensiones: *microestructura*, *macroestructura* y *procesamiento Mental*; los distractores se cosntruyen con información contenida en el texto, pero no directamente relacionada con la información preguntada y sólo una opción es correcta.

El componente microestructura se conforma por dos estrategias de lectura de bajo nivel: la decodificación de la lengua y el reconocimiento de los rasgos estructurales del texto; el constructo macroestructura se conformó por tres estrategias de lectura de alto nivel, inferencia de vocabulario, identificación de ideas principales y de relaciones causales. Finalmente el constructo procesamiento cognitivo se integró por demandas cognitivas como: generalización, categorización, clasificación, y síntesis de la información.

4.1.6 Selección de Variables

La relación entre constructos y variables observables fue la siguiente:

La dimensión microestructura esta conformada por estrategias de bajo nivel, midiendo las siguientes variables:

El reconocimiento de la distribución oracional (pregunta uno), comprueba si el estudiante es capaz de de reconocer los formantes de una palabra; la afijación (pregunta dos),

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea
para la comprensión de lectura en inglés.

la localización de palabras clave, la capacidad de unir los referentes y sus correspondientes referidos, la habilidad de comprender la función de los conectores, la capacidad de discriminar la oración tópica de los detalles de apoyo, se evalúan en los reactivos, tres, cuatro, cinco, seis y siete respectivamente.

La dimensión macroestructura está conformada de la siguiente manera: discriminación de la idea principal, síntesis de la información, ordenamiento de la información, reconocimiento de la organización textual y reconocimiento del tipo de texto, evaluados por los reactivos ocho, nueve, doce, trece y quince respectivamente.

Finalmente los reactivos diez, once y catorce, son indicadores de la dimensión procesamiento mental conformada por: reconocimiento de relaciones causales, reconocimiento de la información factual y generalización de la información.

Tabla 4.1 Indicadores de los constructos (dimensiones) del examen de colocación

<i>Dimensión</i>	<i>Variables</i>	<i>Número de ítem que mide la variable</i>
Microestructura	X_1, X_2, X_3, X_4, X_5 <i>Distribución oracional</i>	1.1-1.5
	X_6, X_{14} <i>Cohesión</i>	2 y 5
	X_7 <i>Palabra clave</i>	3
	$X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}$ <i>Referencia Contextual</i>	4.1-4.6
	X_{15} <i>Coherencia</i>	6

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea
para la comprensión de lectura en inglés.

	<i>X₁₆ Oración Tópico</i>	7
Macroestructura	<i>X₁₇ Idea Principal</i>	8
	<i>X₁₈ Síntesis</i>	9
	<i>X₂₁, X₂₂, X₂₃, X₂₄, X₂₅ Ordenamiento de la información</i>	12.1-12.5
	<i>X₂₆ Organización Textual</i>	13
	<i>X₂₈ Tipo Texto</i>	15
Procesamiento Mental	<i>X₁₉ Causalidad</i>	10
	<i>X₂₀ Información Factual</i>	11
	<i>X₂₇ Generalización</i>	14

En la tabla 4.1 se presenta un resumen de las variables consideradas en el modelo, mostrando la relación que existe entre los constructos (dimensiones) y las variables observadas (cursivas).

Cabe señalar que la puntuación total del examen, X_{29} , es una variable que evidentemente está relacionada con las tres dimensiones.

4.2 DESARROLLO DE LA SOLUCIÓN

El análisis de fiabilidad de los tests o exámenes generalmente se realiza en base a la teoría de respuesta al ítem (TRI), pues ésta viene a suplir las desventajas de *invarianza* del test e *independencia* con respecto a la muestra que presenta la TCT. Sin embargo, existen algunos elementos de la teoría clásica (TCT) que son de interés para el presente estudio, en específico: el cálculo del coeficiente de fiabilidad de Cronbach, pues es de gran utilidad para verificar la consistencia interna de las tres dimensiones que mide el examen.

De este modo se analizan algunos coeficientes propuestos por la TCT y posteriormente se desarrolla el análisis de un modelo de tres parámetros propuesto por la TRI. Pero antes hay que verificar los supuestos necesarios.

4.2.1 Evaluación de los supuestos de la TRI

Los supuestos que se tiene que cumplir son prácticamente los siguientes:

- **Normalidad multivariable:** Puesto que para la estimación de los tres parámetros que definen la curva característica de información (CCI) en el modelo de tres parámetros de la TRI se emplea el método de máxima verosimilitud, es necesario suponer que los datos siguen una distribución normal multivariable.

Para evaluar la normalidad de cada variable, se aplican los test de simetría y curtosis, así como la prueba no paramétrica de Kolmogorov-Smirnov. La simetría hace referencia al reflejo de los datos a partir de la media muestral, mientras que la curtosis indica el apuntamiento o llanura de la distribución, para evaluar la simetría y curtosis se emplean los siguientes estadísticos.

$$Z_{simetría} = \frac{simetría}{\sqrt{6/N}} \quad y \quad Z_{curtosis} = \frac{curtosis}{\sqrt{6/N}}$$

Donde N es el tamaño de la muestra, *simetría* y *curtosis* son los coeficientes de asimetría (*as*) y curtosis (*k*), que proporcionan la mayoría de los paquetes estadísticos.

Por otra parte, el estadístico de Kolmogorov-Smirnov (K-S) es una prueba de bondad de ajuste que contrasta la hipótesis nula de que la distribución de una variable se ajusta a una determinada distribución teórica de probabilidad. Para contrastar la hipótesis nula, la prueba se basa en la comparación de una función de distribución empírica $F(X_i)$ y una función de distribución teórica $F_0(X_i)$, el estadístico K-S se calcula como sigue

$$Z_{K-S} = \text{Máx } |D_i|/\sqrt{n}$$

Donde $D_i = F(X_i) - F_0(X_i)$.

Una vez calculados los valores de $Z_{\text{simetría}}$, Z_{curtosis} y Z_{K-S} , se comparan con un valor crítico basado en un nivel de significación α para aceptar o rechazar el incumplimiento de normalidad.

Para que una variable cumpla con el supuesto de normalidad, los valores $Z_{\text{simetría}}$, Z_{curtosis} y Z_{K-S} deben ser inferiores al valor definido por α . En el caso de Z_{K-S} , algunos paquetes estadísticos como SPSS, proporcionan un nivel de significación, el cual para que la variable se ajuste a la distribución, este deberá de ser mayor que el nivel α .

La siguiente tabla 4.2 el análisis de normalidad de las variables, en la que se pueden observar los valores de $Z_{\text{simetría}}$, Z_{curtosis} y Z_{K-S} (y el nivel de significancia), así como la decisión de rechazar o aceptar la hipótesis nula H_0 : la variable en cuestión sigue una distribución normal.

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Tabla 4.2. Normalidad de las variables del examen de comprensión de lectura en inglés

Variable	Califica (X ₂₉)	realizad(X ₁)	acción(X ₂)	Receptor (X ₃)	tiempo(X ₄)
Asimetría	-0,145	-1,207	-0,49	0,204	0,539
error	0,167	0,167	0,167	0,167	0,167
Zsim	-0,86	-7,22	-2,9	1,22	3,22
Curtosis	-0,376	0,734	-1,152	,382	1,538
error	,332	,332	,332	,332	,332
Zcurto	-1,132	2,21	-3,46	1,15	4,63
Z de Kolmogorov-Smirnov	,728	2,356	3,835	1,184	2,592
Sig. asintót. (bilateral)	,664	,000	,000	,121	,000
Decisión	Aceptar	Rechazar	Rechazar	Aceptar	Rechazar

Variable	lugar(X ₅)	formantes(X ₆)	clave(X ₇)	her(X ₈)	he(X ₉)	me(X ₁₀)
asimetría	0,442	-0,424	0,042	-0,748	0,812	0,409
error	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	2,64	-2,53	0,25	-4,22	4,86	2,44
curtosis	1,164	-1,091	-0,106	-0,541	0,767	0,43
error	0,332	0,332	0,332	0,332	0,332	0,332
Zsim	3,5	-3,28	-0,31	-1,62	2,31	1,29
Z de Kolmogorov-Smirnov	6,515	3,625	0,669	4,207	5,188	6,096
Sig. asintót. (bilateral)	,000	,000	,762	,000	,000	,000
Decisión	Rechazar	Rechazar	Aceptar	Rechazar	Rechazar	Rechazar

Variable	we(X ₁₁)	their(X ₁₂)	his(X ₁₃)	Concausa (X ₁₄)	Concontr (X ₁₅)	Tópico (X ₁₆)	Principa (X ₁₇)
asimetría	-1,741	0,31	0,08	0,615	0,09	0,872	0,113
error	0,167	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	-10,42	1,85	0,47	3,68	0,538	5,22	0,67
curtosis	3,004	-1,388	0,602	-0,726	-1,789	0,356	-1,69
error	0,332	0,332	0,332	0,332	0,332	0,332	0,332
Zcurt	9,04	-4,18	1,81	-2,18	-5,38	1,07	-5,09
Z de Kolmogorov-Smirnov	7,237	4,730	1,088	4,923	3,987	3,607	4,068
Sig. asintót. (bilateral)	,000	,000	,187	,000	,000	,000	,000
Decisión	Rechazar	Rechazar	Aceptar	Rechazar	Rechazar	Rechazar	Rechazar

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Variable	resumen(X_{18})	causalid(X_{19})	Factual (X_{20})	ordenam1 (X_{22})	ordenam2 (X_{23})	ordenam3 (X_{24})
asimetría	-0,952	-0,522	0,62	1,32	1,08	-0,273
error	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	-5,7	-3,125	3,71	7,9	6,46	-1,63
curtosis	-0,619	-1,474	-0,241	1,912	,422	0,221
error	0,332	0,332	0,332	0,332	0,332	0,332
Zcurt	-1,86	-4,43	-0,72	5,75	1,27	0,66
Z de Kolmogorov-Smirnov	5,504	4,997	4,971	4,653	4,227	1,243
Sig. asintót. (bilateral)	,000	,000	,000	,000	,000	,107
Decisión	Rechazar	Rechazar	Rechazar	Rechazar	Rechazar	Aceptar

Variable	ordenam4 (X_{25})	ordenam5(X_{26})	orgtextu(X_{27})	generali(X_{21})	tipotext(X_{28})
Asimetría	0,339	-,403	-0,224	1,041	0,371
error	0,167	0,167	0,167	0,167	0,167
Zsim	2,02	-2,41	-1,34	6,23	2,22
curtosis	-1,303	-,654	-,302	-0,556	-0,248
error	0,332	0,332	0,332	0,332	0,332
	-3,92	1,96	0,9	1,67	0,74
Z de Kolmogorov-Smirnov	4,388	4,435	1,123	5,832	4,107
Sig. asintót. (bilateral)	,000	,000	,165	,000	,000
Decisión	Rechazar	Rechazar	Aceptar	Rechazar	Rechazar

Al observar los resultados de la tabla anterior se tiene que tan solo las variables califica, receptor, clave, his, ordenam3 y orgtextu, cumplen con los supuestos de normalidad. Para las variables cumplan con el supuesto de normalidad, es necesario aplicar algunas transformaciones tales como: raíz cuadrada, inversa, logaritmo, entre otras, que nos ayuden a conseguir el cumplimiento de este supuesto.

La tabla 4.3 muestra las transformaciones que permiten mejorar los estadísticos.

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Tabla 4. 3 Muestra las transformaciones realizadas para cumplir el supuesto de normalidad

	Califica (X ₂₉)	realizad(X ₁)	acción(X ₂)	Receptor (X ₃)	tiempo(X ₄)
Transformación	Ninguna	cúbica	A.C	Ninguna	Raiz 2
Asimetría	-0,145	-0,157	-,254	0,204	0,314
error	0,167	0,167	0,167	0,167	0,167
Zasim	-0,86	-0,94	-1,52	1,22	1,88
Curtosis	-0,376	0,131	-,702	,382	,428
error	0,332	0,332	,332	,332	,332
Zcurto	-1,132	0,39	-2,11	1,15	1,28
Z de Kolmogorov-Smirnov	,728	1,146	2,210	1,184	1,227
Sig. asintót. (bilateral)	,664	,145	,000	,121	,098
Decisión	Aceptar	Aceptar	Rechazar	Aceptar	Aceptar

Variable	lugar(X ₅)	formantes(X ₆)	clave(X ₇)	her(X ₈)	he(X ₉)	me(X ₁₀)
Transformación	inversa	A.C	Ninguna	Potencia 2	Raiz 2	Raiz 2
asimetría	0,124	-0,329	0,042	-0,448	0,412	0,105
error	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	0,74	-1,97	0,25	-2,68	2,46	0,62
cusrtosis	,352	-,291	-0,106	-0,441	0,376	0,109
error	0,332	0,332	0,332	0,332	0,332	0,332
Zsim	1,06	-0,87	-0,31	-1,328	1,132	0,32
Z de Kolmogorov-Smirnov	1,101	2,135	,669	2,507	2,188	0,824
Sig. asintót. (bilateral)	,177	,000	,762	,000	,000	,506
Decisión	Aceptar	Rechazar	Aceptar	Rechazar	Rechazar	Aceptar

Variable	we(X ₁₁)	their(X ₁₂)	his(X ₁₃)	Concausa (X ₁₄)	Concontr (X ₁₅)	Tópico (X ₁₆)	Principa (X ₁₇)
Transformación	A.C	Cubica	Ning.	Cubica	Raiz 2	A.C	Raiz 2
asimetría	-0,251	0,287	0,08	0,231	0,123	0,028	0,243
error	0,167	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	-1,502	1,71	0,47	1,38	0,736	0,16	1,45
curtosis	0,127	-0,387	0,602	-0,485	-,319	0,237	-,501
error	0,332	0,332	0,332	0,332	0,332	0,332	0,332
Zcurt	0,38	-1,16	1,81	-1,46	-0,96	0,71	-1,5
Z de Kolmogorov-Smirnov	,928	1,131	1,088	1,116	2,965	,938	2,278
Sig. asintót. (bilateral)	,372	,147	,187	,166	,000	,343	,000
Decisión	Aceptar	Aceptar	Aceptar	Aceptar	Rechazar	Aceptar	Rechazar

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Variable	resumen(X_{18})	Causalid (X_{19})	Factual (X_{20})	ordenam1 (X_{22})	ordenam2 (X_{23})	ordenam3 (X_{24})
Transformación	Cubica	Raiz 2	Inversa	Potencia 2	A.C	Ninguna
asimetría	-0,3 15	-0,062	0,284	0,253	0,345	-0,273
error	0,167	0,167	0,167	0,167	0,167	0,167
Zsim	-1,88	-0,37	1,7	1,51	-2,06	-1,63
curtosis	-0,271	-0,231	-0,319	,217	0,253	0,221
error	0,332	0,332	0,332	0,332	0,132	0,332
Zcurt	-0,81	-0,69	-0,96	0,65	1,91	0,66
Z de Kolmogorov-Smirnov	2,504	0,697	0,871	0,753	2,027	1,143
Sig. asintót. (bilateral)	,000	,683	,472	,587	,000	,107
Decisión	Rechazar	Aceptar	Aceptar	Aceptar	Rechazar	Aceptar

Variable	ordenam4 (X_{25})	ordenam5 (X_{26})	Orgtextu (X_{27})	Generali (X_{21})	Tipotext (X_{28})
Transformación	Raiz 2	Inversa	Ninguna	Cubica	A.C
Asimetría	0,127	-0,393	-0,224	,305	0,171
error	0,167	0,167	0,167	0,167	0,167
Zsim	0,76	-2,35	-1,34	1,82	1,02
curtosis	-,231	-,638	-,302	-0,256	-0,379
error	0,332	0,332	0,332	0,332	0,332
	-0,69	1,92	0,9	0,77	-1,14
Z de Kolmogorov-Smirnov	0,878	2,435	1,123	2,832	2,107
Sig. asintót. (bilateral)	,457	,000	,165	,000	,000
Decisión	Aceptar	Rechazar	Aceptar	Rechazar	Rechazar

A partir de los resultados mostrados en la tabla 4.3, se puede concluir lo siguiente:

- $X_1, X_4, X_5, X_{10}, X_{11}, X_{12}, X_{14}, X_{16}, X_{19}, X_{20}, X_{21}, X_{23}, X_{24}, X_{25}$. Cumplen con el supuesto de normalidad después de haber aplicado una transformación.
- $X_6, X_8, X_9, X_{15}, X_{17}, X_{18}, X_{26}, X_{28}, X_2, X_{22}$. No cumplieron el supuesto de normalidad después de haberles aplicado transformaciones.
- **Unidimensionalidad:** Se refiere a que el ítem este midiendo una sola dimensión, lo cual, para el examen de comprensión de lectura, se puede observar que los ítems están diseñados para medir una única dimensión, como se puede apreciar en el apartado 4.1.6 de este capítulo.

- **Independencia local:** Al observar los ítems del examen (ver apartado 4.1.6) podemos deducir que no hay ítems encadenados en los que las respuesta de uno de ellos dependan de la respuesta de otros, y tampoco existen ítems en los que influya el orden de presentación.

4.2.2 Modelo TRI de tres parámetros

Este modelo asume que la Curva Característica de los ítems (CCI) esta dada por una función logística con tres parámetros: b, índice de dificultad del ítem; a, índice de discriminación del ítem; c, probabilidad de que un individuo acierte al azar. La función viene dada por:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

Donde **a,b,c**, son los parámetros a ser estimados. Para ellos algunos programas como XCALIBRE, permiten realizar la estimación de dichos parámetros a través del método de Máxima Verosimilitud.

La estimación de parámetros se realizó a través del software XCALIBRE, con las siguientes especificaciones (ver anexo B parte I).

- El Archivo de datos, que contiene todas las respuestas de los alumnos en el examen de colocación, se especifico de la siguiente forma:

Número de items: 28

La secuencia de las respuestas correctas fue:
5621432531443211134223514412

El número de alternativas de cada ítem fue:
7777743595556444434466666434

Todas las variables fueron incluidas en las especificaciones
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY

- Las especificaciones del Programa fueron las siguientes:

Item Parameter Priors: COMMON

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Allow Priors to Float: YES

IRT Model Used: 3-parameter

Maximum Number of Loops: 12

Starting Prior Distribution Moments:

	Mean	SD
a	0.7500	0.1200
b	0.0000	1.0000
c	0.2500	0.0250

Una vez especificados los datos de entrada y las opciones con las que XCALIBRE inicia la estimación, se ejecuto el programa obteniendo una solución en la iteración 4, la cual se muestra en al anexo B parte I y tabla 4.4.

Tabla 4.4 muestra los resultados obtenidos por xcalibre

Item	Lnk Fgl	a	b	c	PC	a error	b error	c error	Resid
1		1.01	-0.52	0.23	0.71	0.170	0.128	***	0.44
2		1.05	0.94	0.21	0.37	0.234	0.147	***	0.87
3		1.12	-0.67	0.23	0.75	0.161	0.123	***	0.79
4		1.18	-1.66	0.23	0.91	0.157	0.163	***	0.79
5		1.12	-1.51	0.23	0.89	0.151	0.157	***	0.62
6		0.99	1.22	0.22	0.37	0.250	0.176	***	0.49
7		0.73	-0.10	0.24	0.66	0.219	0.163	***	0.67
8		0.94	0.69	0.23	0.47	0.237	0.149	***	0.32
9		1.02	-0.37	0.21	0.69	0.175	0.124	***	0.38
10		0.90	-0.38	0.23	0.69	0.183	0.138	***	0.36
11		0.89	-1.40	0.24	0.86	0.145	0.168	***	0.16
12		0.86	1.64	0.22	0.32	0.257	0.232	***	0.31
13		0.90	0.26	0.23	0.56	0.225	0.141	***	0.32
14		0.75	0.61	0.24	0.53	0.249	0.174	***	1.07
15		0.89	1.09	0.23	0.42	0.247	0.179	***	0.20
16		0.84	1.30	0.24	0.39	0.249	0.202	***	0.33
17		0.80	1.32	0.25	0.42	0.253	0.212	***	0.46
18		0.81	0.04	0.23	0.61	0.218	0.150	***	0.37
19		0.96	0.34	0.23	0.54	0.224	0.135	***	0.32
20		0.81	0.32	0.24	0.57	0.237	0.156	***	0.26
21		0.90	0.49	0.23	0.52	0.235	0.148	***	0.15
22		0.94	0.27	0.23	0.56	0.223	0.137	***	0.25
23		0.73	-0.35	0.24	0.69	0.198	0.163	***	0.67
24		0.91	0.79	0.23	0.46	0.241	0.158	***	0.23

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

25		0.96	1.10	0.23	0.40	0.246	0.17	***	0.45
26		0.97	1.74	0.21	0.27	0.273	0.231	***	0.68
27		0.79	-0.09	0.24	0.65	0.213	0.153	***	0.37
28	K	0.71	1.47	0.29	0.50	0.267	0.255	***	1.49

En la tabla 4.4 se observan, aparte de los parámetros estimados (a,b,c) para cada ítem, los residuales (Resid), que deben de estar en un rango de 0.2 y 1.0 para aceptar la hipótesis nula H_0 : Los datos se ajustan al modelo de tres parámetros, a un nivel de significación $\alpha = 0.005$. PC es el índice de dificultad convencional de la TCT, también se muestran los errores estándar de cada ítem, en el caso del parámetros c *** significa que son mayores que 0.1.

Con lo anterior, podemos deducir que el ítem 28 asociado a la variable Tipotext presenta problemas de ajuste al modelo de tres parámetros, pues su residual (1.49) es mayor a 1.0 lo que indica un desajuste en el modelo de tres parámetros, además de presentar una bandrea (Lnk Flg, K) que se refiere a que existe una alta correlación de una opción incorrecta del ítem con la puntuación total de las respuestas correctas.

Para corregir este problema, se elimina el ítem 28 y se vuelve a correr el modelo de tres parámetros en XCALIBRE. Los resultados obtenidos de XCALIBRE se muestran en el anexo B parte II.

La tabla 4.5 muestra un resumen de los resultados obtenidos con xcalibre

Item	a	b	c	PC	a(error)	b(error)	c(error)	Resid
1	1.05	-0.52	0.22	0.71	0.168	0.124	***	0.39
2	1.09	0.92	0.20	0.37	0.231	0.142	***	0.84
3	1.15	-0.67	0.22	0.75	0.160	0.120	***	0.78
4	1.22	-1.65	0.22	0.91	0.158	0.22	***	0.76
5	1.16	-1.50	0.22	0.89	0.152	0.153	***	0.60
6	1.03	1.19	0.22	0.37	0.249	0.168	***	0.46
7	0.75	-0.12	0.24	0.66	0.214	0.159	***	0.73
8	0.97	0.67	0.22	0.47	0.233	0.144	***	0.30
9	1.05	-0.38	0.21	0.69	0.174	0.121	***	0.34
10	0.94	-0.39	0.23	0.69	0.180	0.133	***	0.32
11	0.91	-1.38	0.23	0.86	0.145	0.164	***	0.14
12	0.89	1.59	0.22	0.32	0.257	0.221	***	0.31
13	0.93	0.25	0.22	0.56	0.220	0.136	***	0.36
14	0.77	0.58	0.24	0.53	0.245	0.169	***	1.10
15	0.91	1.07	0.23	0.42	0.244	0.173	***	0.21

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea
para la comprensión de lectura en inglés.

16	0.87	1.27	0.23	0.39	0.248	0.194	***	0.37
17	0.83	1.28	0.24	0.42	0.251	0.204	***	0.51
18	0.84	0.02	0.22	0.61	0.213	0.145	***	0.39
19	0.99	0.32	0.22	0.54	0.220	0.131	***	0.29
20	0.83	0.31	0.23	0.57	0.232	0.152	***	0.29
21	0.93	0.47	0.23	0.52	0.229	0.142	***	0.16
22	0.98	0.26	0.22	0.56	0.218	0.131	***	0.22
23	0.75	-0.37	0.23	0.69	0.195	0.158	***	0.71
24	0.94	0.76	0.23	0.46	0.236	0.151	***	0.21
25	0.99	1.07	0.22	0.40	0.244	0.164	***	0.42
26	1.00	1.71	0.21	0.27	0.274	0.222	***	0.64
27	0.81	-0.10	0.23	0.65	0.208	0.149	***	0.43

Como se observa en la tabla 4.5 y anexo B parte II, no existen residuales superiores a 1, de igual forma, no hay bandera alguna que indique la existencia de un posible desajuste del modelo.

A parte de la estimación de parámetros XCALIBRE ofrece el índice de fiabilidad K- R 21 (Kuder-Richardson) de todo el examen, este índice esta dado por la siguiente expresión:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{j=1}^n p_j q_j}{\sigma_x^2} \right)$$

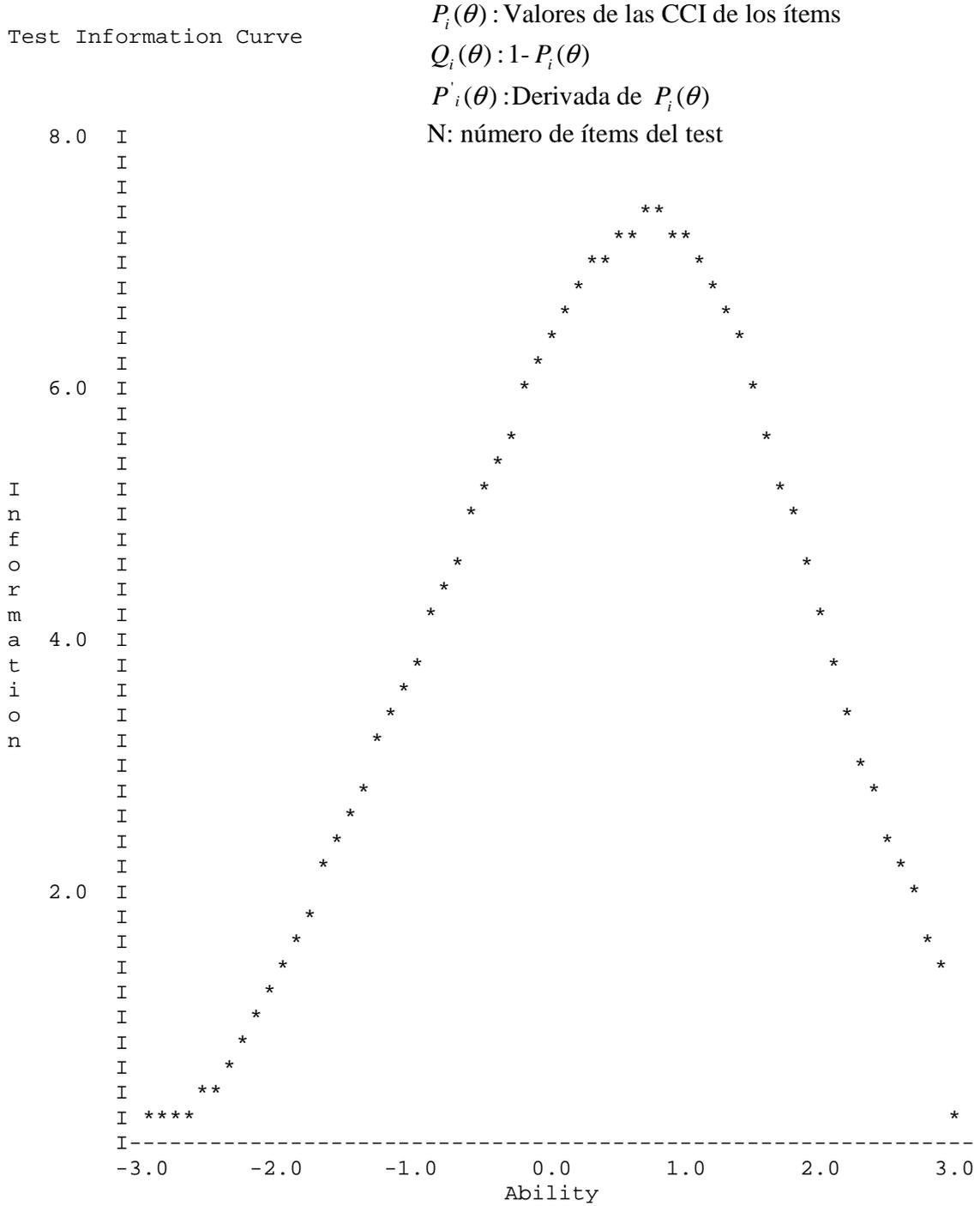
Donde:

n : Número de ítems del test

p_j : Proporción de sujetos que aciertan el ítem j

q_j : Proporción de sujetos que fallan el ítem j

σ_x^2 : Varianza de las puntuaciones del test



Test characteristics		
K-R 21 Reliability	Expected Information	Average Information
0.834	5.531	3.899

Figura 4.1 Muestra la función de información del examen de colocación y las características del test.

La fiabilidad del examen en total es de 0.823, de lo que se deduce que el examen de colocación mide de forma consistente y precisa lo que pretende medir. La figuras 4.1 y 4.2 muestran la curva característica del test y la función de información.

$$I(\theta) = \sum_{i=1}^n \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

De la figura 4.1 se observa que la función aporta mayor información para valores de θ entorno a 0.8, y aparentemente el examen resulta ser no muy difícil.

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

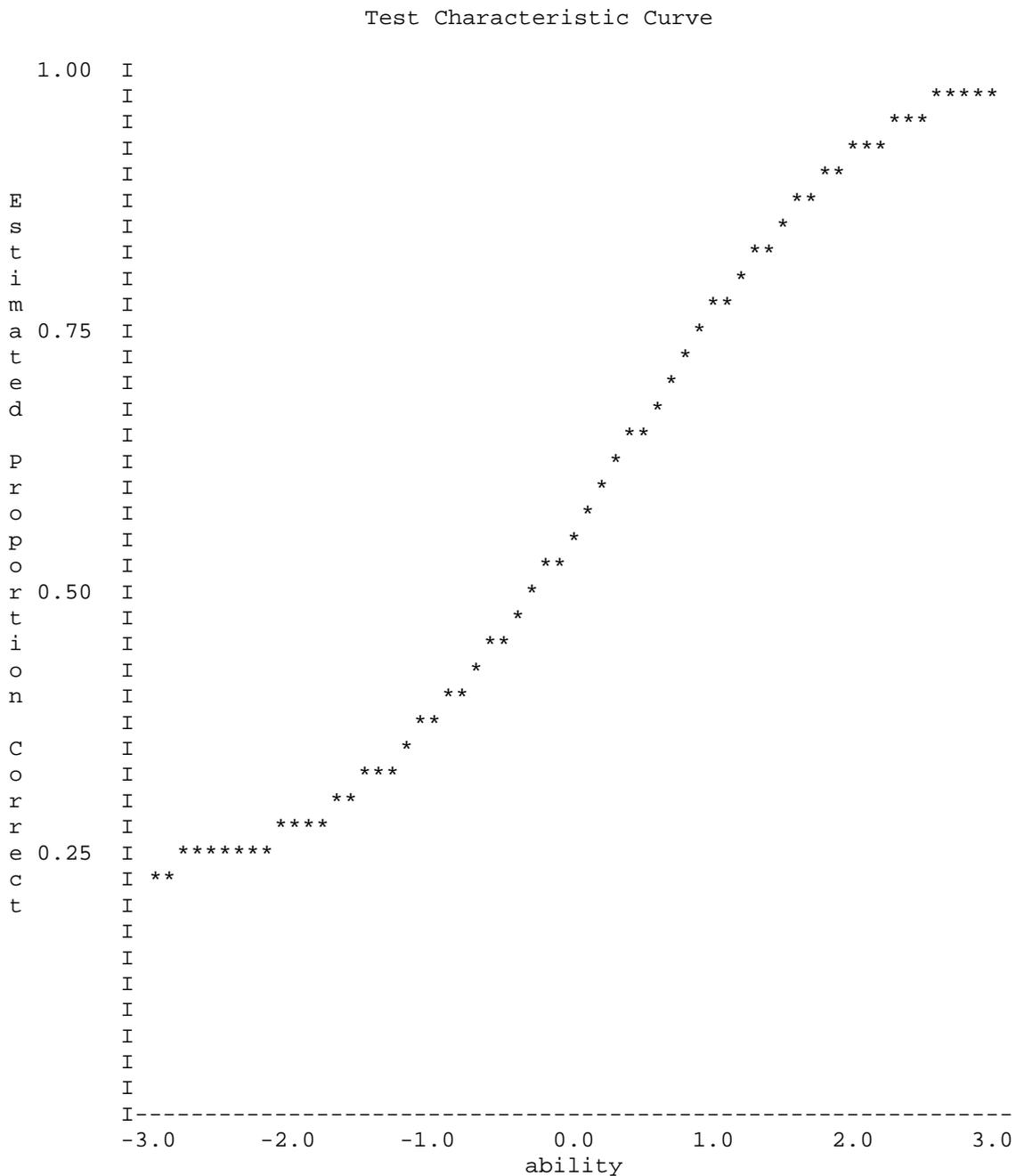


Figura 4.2 Muestra la curva característica del examen de colocación.

4.2.3 Consistencia Interna

El método de consistencia interna más utilizado para estimar la fiabilidad de una prueba cuando se espera que conjuntos de reactivos midan el mismo atributo, es Alpha de Cronbach, que es un coeficiente de correlación entre los reactivos y todo el examen. Se considera que la consistencia interna de una prueba es aceptable cuando el coeficiente α es igual o mayor que 0.7.

Podemos observar que el índice de consistencia interna (α) de los constructos medidos es ligeramente mayor que 0.7 (tabla 4.6), por lo que se considera que los componentes del examen poseen consistencia interna.

Tabla 4.6. Coeficiente alfa estimado Para el examen de colocación

Dimensión	Alfa de Cronbach
Microestructura	.702
Macroestructura	.721
Procesamiento mental	.745

Por otra parte algunos paquetes estadísticos como el SPSS nos permiten hacer un análisis de la variación que tendría el coeficiente alfa si es que se eliminase un ítem determinado, lo cual nos permite, en un momento dado, eliminar un ítem con la finalidad de que la consistencia interna se mejore; pues recuérdese que la consistencia interna no asegura validez de constructo, pero sí ayuda a su convergencia. En la tabla 4.7 se muestra este análisis.

Como se puede observar en la tabla 4.7, la variable que tiene más importancia para la dimensión microestructura es la de *lugar*, pues si se eliminase esta variable el coeficiente de fiabilidad se reduciría en 0.075, lo cual no sería bueno. Por otra parte, si el ítem *concausa* se eliminase, el coeficiente de alfa aumenta a 0.729, que es mejor valor que el 0.702 estimado, además de que la diferencia entre eliminar la variable *lugar* y *concausa* es de .102. Mientras

que para la dimensión Macroestructura la variable más relevante es *ordenam1* y la menos relevante es *tipotext*. Por otra parte para la dimensión de procesamiento mental no hay ninguna exclusión de variable, que mejore el coeficiente alfa, es decir, todas son altamente significativas, y si se eliminase alguna, el coeficiente alfa se reduciría en lugar de aumentar, por lo cual no se puede eliminar ninguna de estas variables.

Tabla 4.7. Muestra el valor del coeficiente alfa si el ítem en cuestión se elimina.

Dimensión	Ítem (variable)	Alfa de Cornbach si el ítem se elimina
Microestructura	Realizad (X ₁)	.708
	Acción (X ₂)	.695
	Receptor (X ₃)	.727
	Tiempo (X ₄)	.711
	Lugar (X ₅)	.627
	Formantes (X ₆)	.714
	Clave (X ₇)	.716
	Her (X ₈)	.689
	He (X ₉)	.724
	Me (X ₁₀)	.701
	We (X ₁₁)	.698
	Their (X ₁₂)	.713
	His (X ₁₃)	.689
	Concausa (X ₁₄)	.729
	Concontr (X ₁₅)	.702
	Tópico (X ₁₆)	.727
	Principa (X ₁₇)	.716
	Resumen (X ₁₈)	.706
Procesamiento Mental	Causalid (X ₁₉)	.713
	Factual (X ₂₀)	.736
	Generali (X ₂₁)	.707
Macroestructura	ordenam1 (X ₂₂)	.744
	ordenam2 (X ₂₃)	.738
	ordenam3 (X ₂₄)	.746
	ordenam4 (X ₂₅)	.717
	ordenam5 (X ₂₆)	.709
	Orgtextu (X ₂₇)	.697
	Tipotext (X ₂₈)	.732

El hecho de que el coeficiente de fiabilidad aumente al eliminar un ítem del test, indica que probablemente ese elemento no está midiendo la misma dimensión por lo que no debería de formar parte del test.

En la tabla 4.8, se muestra la eliminación de aquellas variables, que parecen mejorar el valor estimado del coeficiente alpha de Cronbach.

Nota.

- La eliminación se hizo mediante un método iterativo: se eliminaba a la variable que mejora el coeficiente alfa de Cronbach, y luego se volvía a correr el análisis de fiabilidad en el paquete estadístico SPSS, ya que el eliminar una variable, modifica la validez completa de la escala.
- Se repitió el paso anterior, hasta que la diferencia entre la eliminación de un ítem y otro era muy poco significativa (se estableció a 0.05 como diferencia máxima) entre una variable y otra.

Tabla 4.8. Eliminación de las variables que mejoran el alfa de cronbach.

Dimensión	Ítem	Cronbach's Alpha if Item Deleted
Microestructura	He (X ₉)	.731
	Me (X ₁₀)	.735
	Her (X ₈)	.739
	Realizad (X ₁)	.741
	Formantes (X ₆)	.744
	Clave (X ₇)	.747
	His (X ₁₃)	.749
	We (X ₁₁)	.750
	Their (X ₁₂)	.754
	Principa (X ₁₇)	.757
	Concontr (X ₁₅)	.761
	Resumen (X ₁₈)	.764
Macroestructura	tipotext (X ₂₈)	.732
	Ordenam5 (X ₂₅)	.741
	Orgtextu (X ₂₆)	.759

Como puede observarse, al eliminar las variables poco significativas en cada dimensión, se logró aumentar la consistencia interna de cada una de ellas, quedando de la siguiente forma (ver tabla 4.9).

Tabla 4.9 Consistencia interna de las dimensiones si se eliminan o no las variables

Dimensión	Alfa de Cronbach (sin eliminar variables)	Alfa de Cronbach (al eliminar variables)
Microestructura	.702	.764
Macroestructura	.721	.759
Procesamiento mental	.745	.745

4.3 ANÁLISIS DE ECUACIONES ESTRUCTURALES.

La construcción del modelo teórico, se hará a partir de aquellas variables que aseguran consistencia interna en cada una de las dimensiones (constructos, para asegurar que el modelo converja a una solución factible. En la tabla 4.10 se muestran las variables que serán analizadas en el modelo, así como la dimensión a la que están midiendo.

Tabla 4.10 Indicadores de las variables latentes de las dimensiones que componen el modelo de las habilidades para la comprensión de textos en lengua inglesa.

<i>Dimensión (factores Exógenos)</i>	<i>Variables(observadas)</i>
	<i>Acción X₂, Receptor X₃, Tiempo X₄, Lugar X₅. (Distribución oracional)</i>
	<i>Concausa X₁₄. (Cohesión)</i>

ξ_1 : Microestructura	<i>Tópico X₁₆</i> . (Oración Tópico)
ξ_2 : Macroestructura	<i>Ordenam1 X₂₁, Ordenam2 X₂₂, Ordenam3 X₂₃, Ordenam4 X₂₄</i> . (Ordenamiento de la información)
ξ_3 : Procesamiento Mental	<i>Causalid X₁₉</i> . (Causalidad)
	<i>Factual X₂₀</i> . (Información Factual)
	<i>Generali X₂₇</i> . (Generalización)

Por otra parte, también se considera a la variable *Califica X₂₉* que representa la puntuación total del sujeto obtenida en el examen de colocación, y es un indicador común para las tres dimensiones.

4.3.1 Requerimientos Teóricos

Para lograr el objetivo del análisis factorial confirmatorio que consiste en rechazar o confirmar la estructura factorial propuesta (que supone la estrecha interrelación entre variables), debe de evaluarse el supuesto de *normalidad*, *homocedasticidad* y *linealidad*, que garantizan la convergencia hacia una solución factible.

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

Tabla 4.11 Resumen de las variables

	Califica (X ₂₉)	acción(X ₂)	Receptor (X ₃)	tiempo(X ₄)	lugar(X ₅)
Asimetria	-0,145	-,174	0,204	0,314	0,124
error	0,167	0,167	0,167	0,167	0,167
Zasim	-1,04	1,22	1,88	0,74	-0,86
Curtosis	-,302	,382	,428	,352	-0,376
error	,332	,332	,332	0,332	0,332
Zcurto	-0,90	1,15	1,28	1,06	-1,132
Z de Kolmogorov-Smirnov	,728	1,210	1,184	1,227	1,101
Sig. asintót. (bilateral)	,664	,107	,121	,098	,177
	Aceptar	Aceptar	Aceptar	Aceptar	Aceptar

El supuesto de normalidad ya fue analizado, en la tabla 4.11 se muestra el resumen de las variables a ser analizadas en el modelo y la evaluación del supuesto de normalidad.

	Concausa (X ₁₄)	Tópico (X ₁₆)	Causalid (X ₁₉)	Factual (X ₂₀)	ordenam1 (X ₂₂)
Asimetria	0,231	0,028	-0,062	0,284	0,253
error	0,167	0,167	0,167	0,167	0,167
Zasim	0,16	-0,37	1,7	1,51	1,88
Curtosis	0,356	-0,231	-0,319	,217	,428
error	0,332	0,332	0,332	0,332	,332
Zcurto	1,07	-0,69	-0,96	0,65	1,28
Z de Kolmogorov-Smirnov	1,116	,938	0,697	0,871	0,753
Sig. asintót. (bilateral)	,166	,343	,683	,472	,587
	Aceptar	Aceptar	Aceptar	Aceptar	Aceptar

	ordenam2 (X ₂₃)	ordenam3 (X ₂₄)	ordenam4 (X ₂₅)	Generali (X ₂₇)
Asimetria	0,145	-0,273	0,127	,305
0,167	0,167	0,167	0,167	0,167
-1,63	-1,22	0,76	1,82	-0,94
0,221	0,249	-1,303	-0,256	0,131
0,332	0,332	0,332	0,332	0,332
0,66	0,75	-3,92	0,77	0,39
Z de Kolmogorov-Smirnov	1,027	1,143	0,878	2,832
Sig. asintót. (bilateral)	,197	,107	,457	,000
	Aceptar	Aceptar	Aceptar	Rechazar

A partir de los resultados mostrados en la tabla 4.11, se concluye que tan solo la variable Generali (X_{27}), no cumple con el supuesto de normalidad. Por otra parte, recuerdese que la normalidad soluciona muchas de las veces los problemas de heterocedasticidad.

Para evaluar la linealidad, se debe de tener en cuenta que esta representa la correlación entre cada par de variables, de modo que si la correlación es baja entonces no existe asociación lineal y las variables o variables salen del análisis. La evaluación de este supuesto se refleja en los índices de modificación⁴. Por lo que este supuesto va implícito en la estimación del modelo.

Para efectuar el análisis factorial confirmatorio, se siguen las etapas del modelado de ecuaciones estructurales que se exponen a continuación

4.3.2 Especificación del Modelo

Utilizando la notación LISREL para la especificación del modelo, las relaciones causales que se observan en la tabla 4.10 quedan representadas en el diagrama de senderos de la figura 4.3.

A partir de la figura 4.3 se puede observar que todos los constructos del diagrama de relaciones son exógenos, por lo que sólo necesitamos considerar el modelo de medida y su representación matricial, que se aprecian en las figuras 4.4 y 4.5 respectivamente.

⁴ Son aquellos que indican que la variable puede explicar mejor a otros constructos.

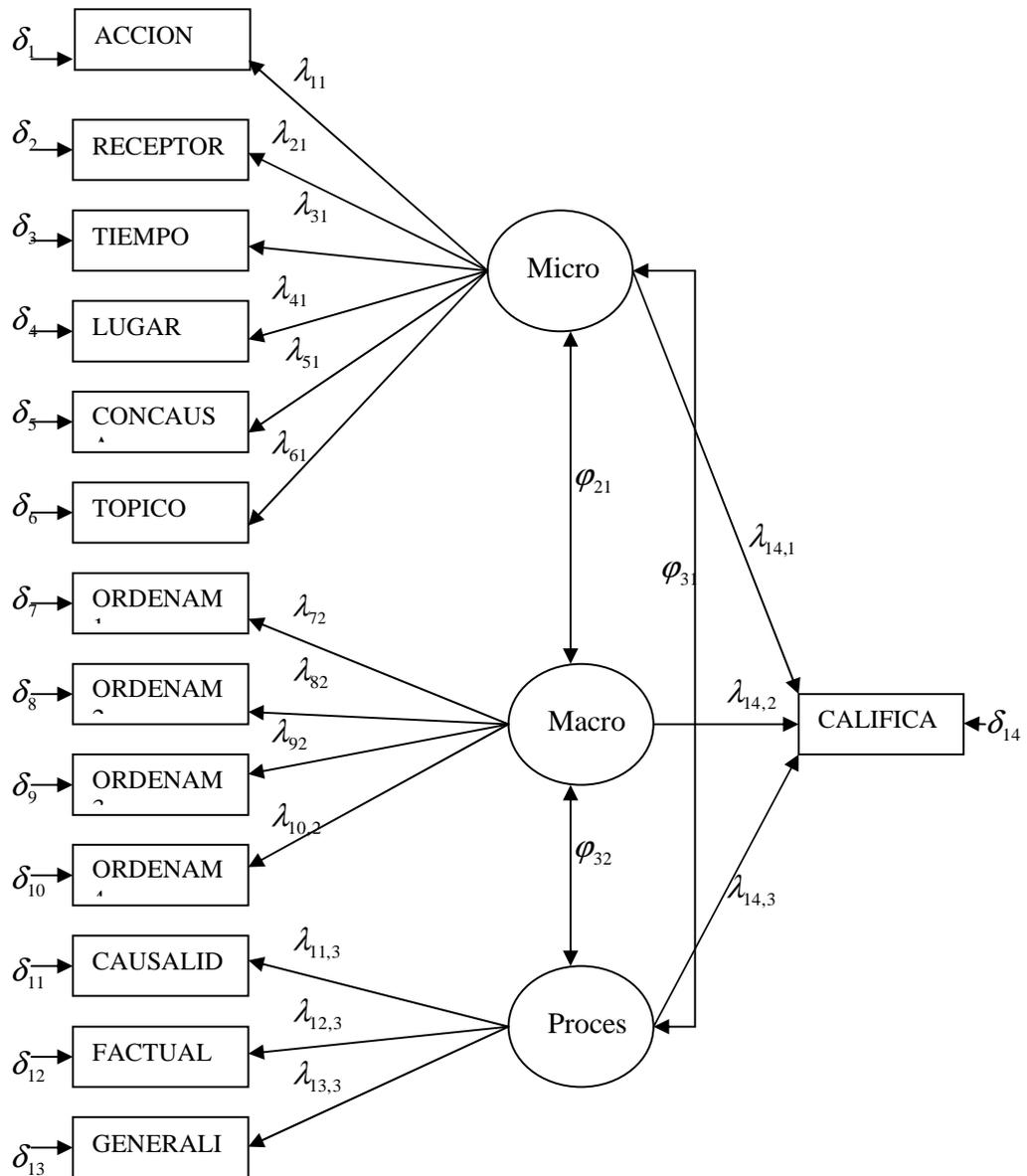


Figura 4.3. Diagrama de senderos que representa el modelo teórico de las habilidades de los estudiantes para la comprensión de textos en inglés

$$\begin{aligned}
 \text{CALIFICA} &= \lambda_{14,1}^x \text{Micro} + \lambda_{14,2}^x \text{Macro} + \lambda_{14,3}^x \text{Proces} + \delta_{14} \\
 \text{ACCION} &= \lambda_{11}^x \text{Micro} + \delta_1 \\
 \text{RECEPTOR} &= \lambda_{21}^x \text{Micro} + \delta_2 \\
 \text{TIEMPO} &= \lambda_{31}^x \text{Micro} + \delta_3 \\
 \text{LUGAR} &= \lambda_{41}^x \text{Micro} + \delta_4 \\
 \text{CONCAUSA} &= \lambda_{51}^x \text{Micro} + \delta_5 \\
 \text{TOPICO} &= \lambda_{61}^x \text{Micro} + \delta_6 \\
 \text{CAUSALID} &= \lambda_{11,1}^x \text{Proces} + \delta_7 \\
 \text{FACTUAL} &= \lambda_{12,1}^x \text{Proces} + \delta_8 \\
 \text{GENERALI} &= \lambda_{13,1}^x \text{Proces} + \delta_9 \\
 \text{ORDENAM1} &= \lambda_{71}^x \text{Macro} + \delta_{10} \\
 \text{ORDENAM2} &= \lambda_{81}^x \text{Macro} + \delta_{11} \\
 \text{ORDENAM3} &= \lambda_{91}^x \text{Macro} + \delta_{12} \\
 \text{ORDENAM4} &= \lambda_{10,1}^x \text{Macro} + \delta_{13}
 \end{aligned}$$

Figura 4.4. Modelo de Medida para el AFC

$$\Lambda_x = \begin{array}{c} \left| \begin{array}{c} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \\ \lambda_{41} \\ \lambda_{51} \\ \lambda_{61} \end{array} \right. \\ \left| \begin{array}{c} \lambda_{72} \\ \lambda_{82} \\ \lambda_{92} \\ \lambda_{10,2} \end{array} \right. \\ \left| \begin{array}{c} \lambda_{11,3} \\ \lambda_{12,3} \\ \lambda_{13,3} \end{array} \right. \\ \left| \begin{array}{ccc} \lambda_{14,1} & \lambda_{14,2} & \lambda_{14,3} \end{array} \right. \end{array} \quad \Phi = \begin{array}{c} \left| \begin{array}{cc} 0 & \\ \varphi_{21} & 0 \\ \varphi_{31} & \varphi_{32} & 0 \end{array} \right. \end{array}$$

$$\Theta = \text{diag}(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \delta_{10}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14})$$

Figura 4.5 Forma matricial del diagrama de senderos para el AFC.

4.3.3 Identificación del Modelo

Para asegurar que los parámetros del modelo puedan ser estimados y encontrar un valor a cada uno de ellos, una de las primeras condiciones que se deben de cumplir es la de obtener grados de libertad positivos mediante la siguiente expresión.

$$gl = ((p+q)(p+q+1)/2) - t$$

Donde p es el número de variables observables independientes, que para el presente modelo son 14, q el número de variables observable dependientes, que en este caso son cero, y t el número de parámetros a estimar, que para este caso son 33. Sustituyendo estos valores en la expresión anterior se obtiene:

$$gl = ((14+0)(0+14+1)/2) - 33 = 105 - 33 = 72 \geq 0$$

El resultado anterior implica que el modelo es *sobreidentificado*, existiendo más ecuaciones que incógnitas, lo cual es lo deseable, pues el modelo tiene varias soluciones.

Una segunda condición, es que cada ecuación del modelo debe de ser diferente de las demás o de cualquier combinación lineal de ellas, ya que de no ser así, el modelo no tendría solución factible, lo que implica que la matriz de correlaciones o covarianzas no son inversibles (definida no positiva) y la solución del sistema de ecuaciones no tendría lugar. Esta condición no se puede comprobar con lo que hasta ahora se ha expuesto, sin embargo, se adelanta que el modelo está especificado, pues en el proceso de solución, ninguna ecuación es combinación lineal de otra ya que se encontró un valor para cada parámetro del diagrama de senderos (figura 4.3). Además, de acuerdo con Hair (1999), los grados de libertad no aseguran para todos los casos que el modelo quede especificado, pero si resulta ser en la mayoría de ellos un fuerte indicio de que el modelo ha quedado especificado.

4.3.4 Estimación de Parámetros

Existe software informático que permiten la estimación de parámetros de los modelos de ecuaciones estructurales, destacan entre este tipo de programas: LISREL, AMOS y EQS. Para la presente investigación la estimación se hace mediante LISREL 8.7 (Linear Structural Relationship by Method of Likelihood) versión del estudiante. El software pide como requisito el cumplimiento de normalidad multivariable de las variables que forman parte de la matriz de correlaciones o covarianzas con la que el programa inicia la estimación. Para el presente estudio, la matriz de entrada fue la de correlaciones, por los siguientes motivos:

- La matriz de correlaciones proporciona coeficientes en valores estandarizados, es decir, el valor de los parámetros oscila entre -1 y 1, lo que facilita la interpretación de los mismos, mientras que la interpretación de los resultados es más difícil cuando se utilizan covarianzas por que los coeficientes deben interpretarse en las unidades de medida del constructos.
- El empleo de la matriz de correlaciones es apropiado cuando el objetivo del investigador es entender el patrón de las relaciones entre constructos sin explicar la varianza total de los mismos.
- Se ha demostrado que la matriz de correlación ofrece estimaciones más conservadoras de la significación de los coeficientes y no están sesgadas al alza.

Una vez que se ha elegido la matriz de entrada, para estimar el valor de los 33 parámetros del diagrama de senderos, se debe de especificar toda la estructura del modelo en el programa informático a través de su lenguaje de compilación, que para el caso de LISREL se denomina *lenguaje SIMPLIS*.

De acuerdo con el lenguaje *SIMPLIS* las instrucciones requeridas para la estimación de los parámetros del presente modelo (ver figura 4.3) son:

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea
para la comprensión de lectura en inglés.

AFC para el modelo que mide las habilidades de los estudiantes para la comprensión de textos en Inglés

Observed Variables

*CALIFICA ACCION RECEPTOR TIEMPO LUGAR CONCAUSA TOPICO CAUSALID
FACTUAL GENERALI ORDENAM1 ORDENAM3 ORDENAM2 ORDENAM4*

Correlation Matrix from file 'C:\program files\lisrel87s\FINAL2\corre.cor'

Sample Size = 213

Latent Variables Micro Macro Proces

Relationships

CALIFICA = Micro Macro Proces

RECEPTOR TIEMPO LUGAR CONCAUSA TOPICO ACCION = Micro

ORDENAM1 ORDENAM2 ORDENAM3 ORDENAM4 = Macro

CAUSALID FACTUAL GENERALI = Proces

Path Diagram

Options: ND=3 ME=ML IT=250

LISREL Output: MI RS SC EF

End of Problem

Los resultados obtenidos al ejecutar el código anterior se muestran en el anexo C parte I, donde se observan la estimación de los valores de los parámetros del modelo, mismos que se resumen en la tabla 4.12.

Tabla 4.12 Valor de los parámetros estimados para el AFC.

$\lambda_{11}^x = 0.194$	$\lambda_{12,3}^x = 1.655$	$\delta_4 = -0.969$
$\lambda_{21}^x = 0.452$	$\lambda_{13,3}^x = 0.449$	$\delta_5 = 0.768$
$\lambda_{31}^x = 0.566$	$\lambda_{14,1}^x = 0.783$	$\delta_6 = 0.781$
$\lambda_{41}^x = 1.677$	$\lambda_{14,2}^x = 0.564$	$\delta_7 = 0.437$
$\lambda_{51}^x = 0.532$	$\lambda_{14,3}^x = 0.484$	$\delta_8 = 0.634$
$\lambda_{61}^x = 0.348$	$\varphi_{21} = 0.737$	$\delta_9 = 0.512$
$\lambda_{72}^x = 0.736$	$\varphi_{31} = 0.495$	$\delta_{10} = 0.895$
$\lambda_{82}^x = 0.426$	$\varphi_{32} = 0.573$	$\delta_{11} = 0.287$
$\lambda_{92}^x = 0.670$	$\delta_1 = 0.912$	$\delta_{12} = -0.997$
$\lambda_{10,2}^x = 0.305$	$\delta_2 = 0.748$	$\delta_{13} = 0.754$
$\lambda_{11,3}^x = 0.872$	$\delta_3 = 0.632$	$\delta_{14} = 0.317$

De la tabla 4.12 podemos observar que el parámetro λ_{31}^x que es correspondiente a la relación entre la variable *TIEMPO*, resulta ser la variable más significativa para medir la variable latente *Micro*, así como lo son las variables *ORDENAMI* y *CAUSALID* para los constructos *Macro* y *Proces* respectivamente. Sin embargo, antes de abundar sobre los parámetros, es preciso efectuar la evaluación del modelo, pues en la tabla 4.12 se observan dos valores superiores a uno, lo que indica la presencia de estimaciones infractoras.

4.3.5 Evaluación del Modelo

Una vez que se han estimado los parámetros se procede a la evaluación del modelo. Que consiste en comprobar si el modelo se ajusta a los datos observados en forma conjunta y, a través del modelo estructural y de medida, pero antes de evaluar dicha calidad, es necesario poner atención en la presencia de estimaciones infractoras, que comúnmente se presentan de dos formas: estimaciones superiores a 1 y términos de error negativos en las variables observadas⁵.

Como se puede observar en la tabla 4.12, los parámetros λ_{41}^x y $\lambda_{12,3}^x$, correspondientes a la relación existente entre las variables (LUGAR, Micro) y (FACTUAL, Proces) equivale a 1.677 y 1.655 respectivamente. Por otra parte sus términos de error δ_4 (LUGAR) y δ_{12} (FACTUAL) resultan ser negativos -0.969 y -0.997 (ver anexo C parte I o tabla 4.12). Lo que resultan ser estimaciones infractoras en las variables LUGAR y FACTUAL.

Una forma de corregir las estimaciones infractoras es fijar la varianza de cada una de las variables a un valor muy pequeño de 0.005. De este modo se fija la varianza del error de las variables LUGAR y FACTUAL a 0.005 y se vuelve a ejecutar la especificación del modelo en el programa LISREL.

La reespecificación del modelo y código, así como los resultados, se muestran en el anexo C parte II. y en la tabla 4.13, en los que se puede observar la ausencia de estimaciones infractoras, por lo que se procede a evaluar la calidad del ajuste del modelo.

La calidad del ajuste mide la correspondencia entre la matriz de entrada real, que en este caso es la de correlaciones, y con la que se predice mediante el modelo propuesto.

⁵ Hair, J. et al. *Análisis multivariante*. Pág 637.

Tabla 4.13 Estimación de los parámetros con varianza fija para LUGAR y FACTUAL

$\lambda_{11}^x = 0.382$	$\lambda_{12,3}^x = 0.978$	$\delta_4 = 0.005$
$\lambda_{21}^x = 0.468$	$\lambda_{13,3}^x = 0.426$	$\delta_5 = 0.668$
$\lambda_{31}^x = 0.593$	$\lambda_{14,1}^x = 0.783$	$\delta_6 = 0.854$
$\lambda_{41}^x = 0.975$	$\lambda_{14,2}^x = 0.564$	$\delta_7 = 0.376$
$\lambda_{51}^x = 0.532$	$\lambda_{14,3}^x = 0.484$	$\delta_8 = 0.671$
$\lambda_{61}^x = 0.346$	$\varphi_{21} = 0.723$	$\delta_9 = 0.524$
$\lambda_{72}^x = 0.724$	$\varphi_{31} = 0.472$	$\delta_{10} = 0.965$
$\lambda_{82}^x = 0.429$	$\varphi_{32} = 0.386$	$\delta_{11} = 0.312$
$\lambda_{92}^x = 0.670$	$\delta_1 = 0.768$	$\delta_{12} = 0.005$
$\lambda_{10,2}^x = 0.187$	$\delta_2 = 0.682$	$\delta_{13} = 0.687$
$\lambda_{11,3}^x = 0.888$	$\delta_3 = 0.557$	$\delta_{14} = 0.362$

Las medidas de ajuste global o medidas de evaluación conjunta, se muestran en el anexo C parte II. Las medidas de ajuste global presentaron los siguientes valores: El valor de **chi-cuadrada** χ^2 es de 48.431 con 72 grados de libertad, y con una significación p de 0.985. Recuerdese que para que la hipótesis nula H_0 : El modelo se ajusta a los datos, sea aceptada, el valor p debe de ser mayor al valor α , que para este caso es $\alpha=0.005$, de este modo se acepta H_0 , es decir, no existen diferencias significativas entre la matriz de correlaciones producida y la observada o real. El **Índice de bondad de ajuste GFI**, tiene un valor de 0.968, lo que implica un ajuste muy bueno en cuanto a la variabilidad de los datos. De la misma forma el **índice ajustado de bondad y ajuste AGFI** tiene un valor de 0.954, indicando un ajuste muy alto. La **raíz cuadrada media de los residuos RMR** es 0.0467 que es muy próximo a cero, lo

que indica que los errores entre la matriz observada y la matriz reproducida son muy bajos. De acuerdo a estos resultados se infiere que el ajuste es muy bueno.

Para las medidas de ajuste incrementales, que comparan el modelo analizado con un modelo nulo (sin parámetros a estimar), tenemos: el **índice de ajuste no normalizado NFI**, aunque no es un valor superior a 0.9, si es muy próximo a él, por lo que se acepta el ajuste, por otra parte, se tiene un valor de 0.918 para el índice **NNFI** y un valor de 1 para el **índice de ajuste comparativo CFI** lo que indica un modelo bien ajustado.

Por otra parte el índice de parsimonia, **PNFI** (0.251), no es cercano a uno, pero tampoco es muy cercano a cero, lo que indica la posible existencia de otro modelo que mejore este índice, sin embargo, de acuerdo con los valores de los índices anteriores, en general el modelo presenta un muy buen ajuste, por lo que este índice puede no tener relevancia.

También se pueden observar las tablas de los residuales estandarizados que se muestran en el anexo C parte II. Donde podemos observar solo existe un residual que supera el valor de 1.96 para un $\alpha = 0.005$ que esta relacionada con la variable ORDENAM3, que puede indicar un desajuste en el modelo lo que implica una posible reespecificación, sin embargo, al observar la gráfica de estos residuales (ver anexo C parte II), se puede apreciar un ajuste bueno, pues la línea que se traza con los puntos X es más vertical que la trazada de 45°. De esta forma podemos concluir que las medidas de ajuste global (conjunto) del modelo proporcionan evidencias suficientes para decir que el modelo es ajustable y que se tiene una buena representación del modelo teórico propuesto.

Una vez que evaluamos la calidad del ajuste del modelo en su conjunto, se procede con la evaluación del modelo de medida. Recuérdese que el modelo de medida representa la relación entre variables observadas exógenas y constructos exógenos. Así, al efectuar el análisis sobre cada una de las variables observables tenemos: Las todas las variables que explican el constructo microestructura (Micro), están significativamente relacionadas con dicho constructo, pues sus valores t (3.329, 2.742, 2.569, 3.238, 4.532, 3.268) (ver anexo C parte II) superan el valor crítico de 1.96 a un nivel de significación de $\alpha = 0.005$; de igual

forma, las variables CAUSALID, FACTUAL Y GENERALI, con valores t de 2.979, 15.278 y 2.623 respectivamente, son significativas con el constructo procesamiento mental (Proces). Sin embargo, para el constructo Macroestructura (Macro), todos los indicadores son significativo pues los valores t superan el valor de 1.96, excepto uno, el de ORDENAM3, cuyo valor t es de 1.834. De igual forma, para la variable CALIFICA, que esta relacionada con todos los constructos, podemos observar que sólo es significativa para el constructo Micro, pues el valor t (3.329), supera el de 1.96, sin embargo, para los constructos Macro y Proces presenta valores t de 1.835 y 1.937 respectivamente.

Le relación existente entre los constructos también debe de ser significativa. La significación se muestra en la tabla PHI del anexo C parte II, en el que se observa que los constructos Macro y Micro están significativamente relacionados, con un valor t de 2.025, de igual forma los constructos Macro y Proces, están significativamente relacionados con un valor t de 2.131. La única relaciona no significativa es la existente entre Proces y Micro, pues se presenta un valor t de 1.768.

Otro aspecto importante al hacer la evaluación del modelo, es conocer la fiabilidad y la varianza extraída para cada constructo, pues conocer estos dos aspectos, nos permite verificar que los indicadores especificados son suficientes en la representación de los constructos. Los cálculos se realizan a través de las siguientes formulas.

$$Fiabilidad\ de\ constructo = \frac{\left(\sum ponderaciones\ estandarizadas\right)^2}{\left(\sum ponderaciones\ estandarizadas\right)^2 + \sum \xi_j}$$

$$Varianza\ extraída = \frac{\left(\sum ponderaciones\ estandarizadas^2\right)}{\left(\sum ponderaciones\ estandarizadas\right)^2 + \sum \xi_j}$$

Los resultados al aplicar las fórmulas a cada constructo se muestran a continuación.

➤ **Fiabilidad.**

Suma de las ponderaciones estandarizadas

$$Micro = 0.382 + 0.468 + 0.593 + 0.975 + 0.532 + 0.346 + 0.788 = 4.03$$

$$Macro = 0.724 + 0.429 + 0.670 + 0.187 + 0.612 = 2.622$$

$$Proces = 0.888 + 0.978 + 0.426 + 0.379 = 2.671$$

Suma del error de medida

$$Micro = 0.362 + 0.768 + 0.682 + 0.557 + 0.005 + 0.668 + 0.854 = 3.896$$

$$Macro = 0.362 + 0.376 + 0.671 + 0.524 + 0.965 = 2.898$$

$$Proces = 0.362 + 0.312 + 0.005 + 0.687 = 1.366$$

Cálculo de la fiabilidad

$$Micro = (4.03)^2 / (4.03^2 + 3.896) = 0.814$$

$$Macro = (2.622)^2 / (2.622^2 + 2.898) = 0.703$$

$$Proces = (2.671)^2 / (2.671^2 + 1.366) = 0.839$$

➤ **Varianza extraída**

Suma de los cuadrados de las ponderaciones estandarizadas

$$Micro = 0.382^2 + (-0.468)^2 + 0.593^2 + 0.975^2 + 0.532^2 + 0.346^2 + 0.788^2 = 2.69$$

$$Macro = 0.724^2 + 0.429^2 + 0.670^2 + 0.187^2 + 0.612^2 = 1.566$$

$$Proces = 0.888^2 + 0.978^2 + 0.426^2 + 0.379^2 = 2.07$$

Cálculo de la varianza extraída

$$Micro = (2.69) / (2.69 + 3.896) = 0.408$$

$$Macro = (1.566) / (1.566 + 2.898) = 0.350$$

$$Proces = (2.07) / (2.07 + 1.366) = 0.602$$

En términos de fiabilidad de las variables latentes Micro, Macro y Proces, todas cumplen con el nivel sugerido de fiabilidad (mayor que 0.7), de esta forma se dice que los

indicadores individuales de cada constructo son consistentes en sus medidas, además de ser suficientes para estos constructos en el modelo de medida.

Por otra parte, las medidas de varianza extraída indican que sólo el constructo Proces está sobre el nivel sugerido de 0.5 o más. El valor de la varianza extraída para el constructo Micro esta muy cerca de éste valor, sin embargo, el valor de Macro no lo esta tanto, lo que significa que gran parte de la varianza de los indicadores no se toma en cuenta en éste constructo.

Hasta ahora, se puede decir que las medidas de ajuste son muy buenas, y que el modelo teórico queda muy bien representado por los indicadores propuestos, no obstante, existe un indicador que no cumple con algunas de las condiciones de calidad de ajuste, y que debe de ser revisado.

El indicador ORDENAM3, presenta un nivel de significación t de 1.834, que no es superior a 1.96, lo cual indica que, individualmente no es un indicador significativo para el constructo Macroestructura. Por otra parte, al observar los residuales estandarizados para este indicador (ver anexo C parte II), se aprecia que el valor asociado a este indicador supera el valor 1.96 con significación de $\alpha = 0.005$ que indica una posible reespecificación del modelo. Pero no se puede afirmar este hecho hasta revisar los índices de modificación.

Los índices de modificación corresponden aproximadamente a la reducción de la chi-cuadrada que se produciría si el parámetro fuera estimado. Un valor de cinco o superior sugiere que se obtiene una reducción estadísticamente significativa en la chi-cuadrada cuando se estima el parámetro.

Al revisar los índices de modificación que se encuentran en el anexo D parte II se presenta un índice superior a 5, siendo éste índice el más alto de todos y encontrado en la matriz LAMBDA X entre la variable ORDENAM3 y el constructo Micro, con un valor de 8.125, lo que sugiere una reespecificación del modelo que produce un mejor ajuste.

Una posible reespecificación es relacionar el indicador ORDENAM3 con los constructos Micro y Macro, entonces se tendría que reescribir el código en el lenguaje SIMPLIS y ejecutar el programa nuevamente, para posteriormente hacer la evaluación del modelo, sin embargo, para poder reespecificar el modelo, las modificaciones tiene que estar apoyadas teóricamente, para la presente investigación no lo es así, ya que cada ítem esta diseñado para medir una única dimensión, tal y como se ve en el análisis de fiabilidad de éste capítulo. De modo que no se efectúa ninguna reespecificación en el modelo.

4.4 ANÁLISIS DE RESULTADOS

4.4.1 Interpretación del modelo

En base a la calidad de bondad y ajuste, podemos decir que no existen evidencias suficientes para afirmar que el modelo no se ajusta a los datos, aun cuando los índices de modificación indican la reespecificación del modelo, pues recuérdese que la cualquier modificación del modelo debe de tener una justificación teórica antes de ser considerada⁶. De esta forma, el modelo que representa las habilidades necesarias que tiene los alumnos para comprender textos en lengua inglesa, se presenta en la figura 4.6.

La validez de constructo se compone de validez convergente y divergente. La validez convergente se establece a partir de correlaciones altas y significativas entre las variables manifiestas y las variables latentes; la validez divergente se encuentra cuando se producen correlaciones bajas y no significativas entre los factores medidos⁷.

La validez convergente se establece a partir de las relaciones altas y significativas entre los indicadores y las variables latentes microestructura, macroestructura y procesamiento cognitivo. A partir de los resultados obtenidos (ver la figura 4.6 y anexo C parte II) se puede establecer que los factores mencionados poseen validez convergente.

⁶ Hair, et al. Idem

⁷ (Corral, 1995; Hair, Anderson, Tatham & Black, 1999).

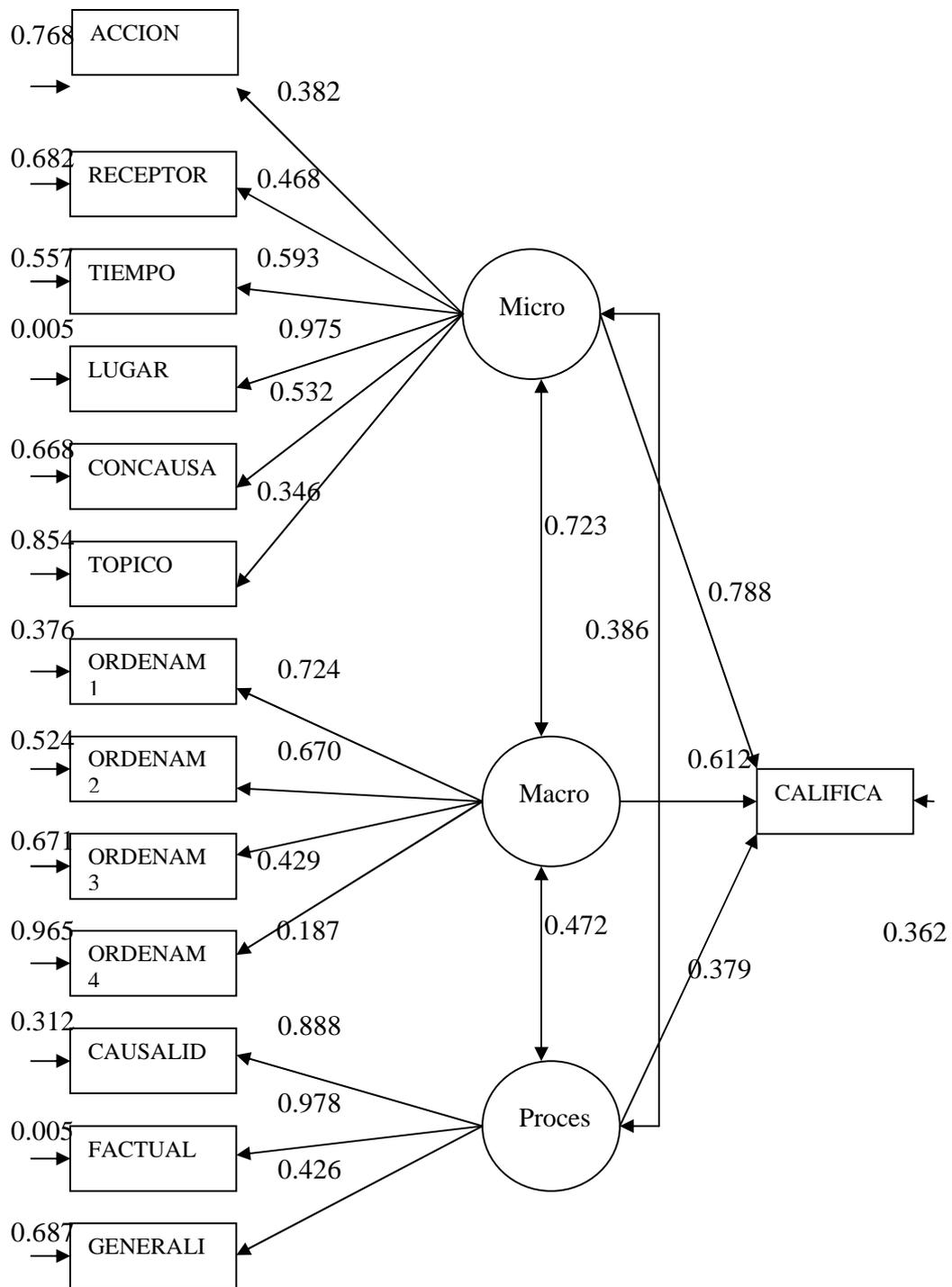
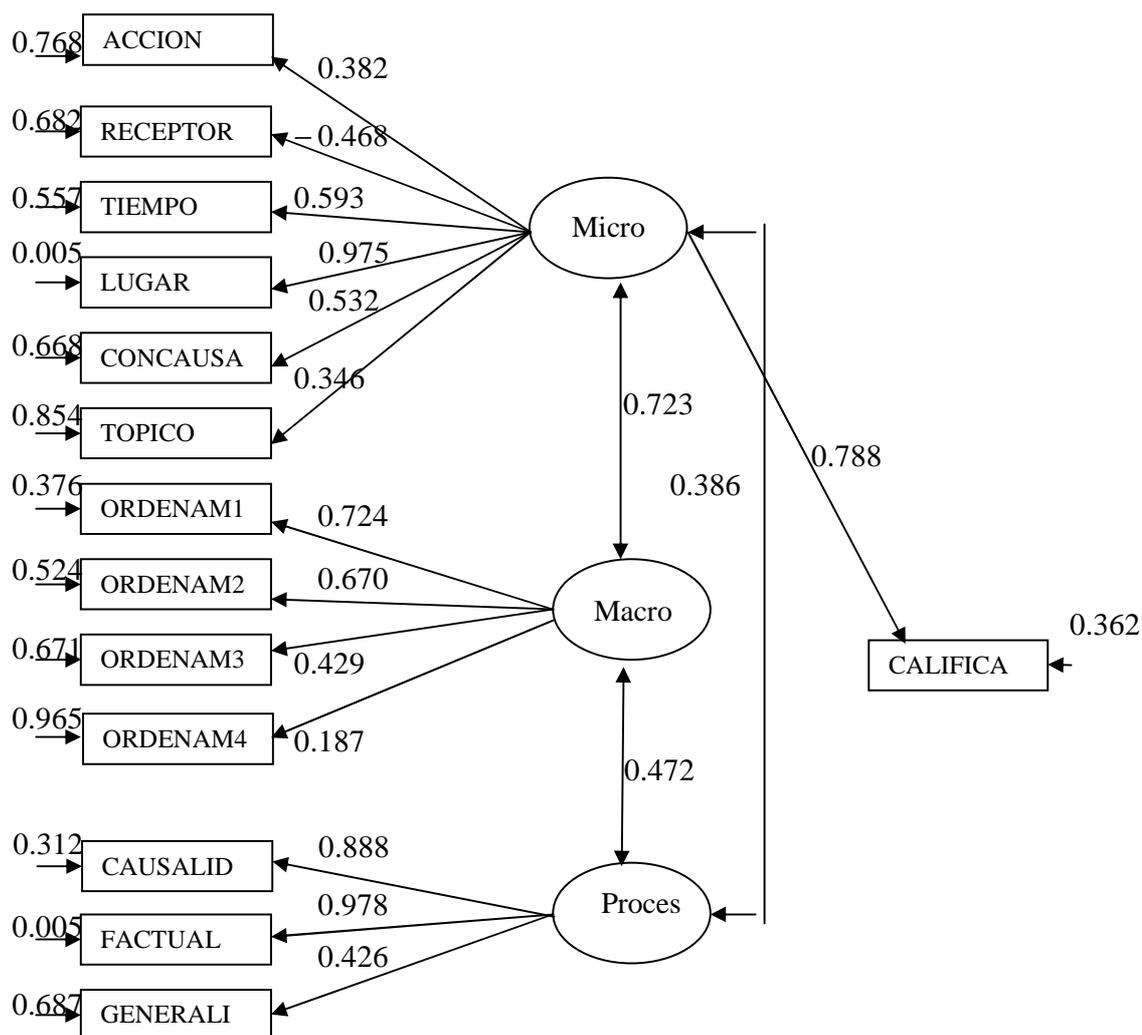


Figura 4.6. Modelo con los parámetros estimados

Capítulo 4: Evaluación y validación del examen de colocación al curso en línea para la comprensión de lectura en inglés.

También se observó en la matriz LAMBDA X y PHI (ver anexo C parte II), que los factores macroestructura y procesamiento cognitivo presentan correlaciones pequeñas y no significativas; esto es, que no están asociados entre sí, lo que muestra que la ejecución de los examinandos fue capaz de establecer una diferencia entre estos constructos; lo anterior indica que el modelo posee validez divergente.



Los resultados mostraron que los indicadores observados de la variable latente microestructura tuvieron un peso estructural de .788 con respecto al puntaje total del examen (variable CALIFICA), mientras que los indicadores de las variables latentes macroestructura y procesamiento cognitivo tuvieron un peso estructural de 0.612 y 0.379 respectivamente, pero

sus valores t de significación son menores a 1.96 (a un nivel de $\alpha = 0.005$), lo que indica que no tuvieron un peso estructural directo con respecto a la variable CALIFICA (puntaje total); no obstante, la variable latente macroestructura mostró un peso estructural de .723 con referencia a la microestructura y el procesamiento cognitivo tuvo un peso estructural de .472 con respecto a la macroestructura y de .386 en relación con la microestructura.

4.4.2 Limitaciones

Una de las primeras limitaciones presentadas en la presente investigación, es la impuesta por el software LISREL 8.7 versión del estudiante utilizado en la estimación de parámetros del modelo causal planteado, ya que tan solo permite trabajar con un máximo de quince variables observadas, lo cual impidió que se analizara el modelo con un mayor número de indicadores de cada constructo.

Lo anterior, obligó a buscar una forma de reducir el número de 28 variables (indicadores) inicialmente propuestas a un número de no más de 15, por lo que en el análisis de fiabilidad, en la parte de evaluación de la consistencia interna de cada dimensión, se fue más estricto tratando de aumentar el coeficiente de correlación alfa de Cronbach, de manera que se tuvo que construir un método iterativo para aumentar dicho coeficiente e ir eliminando variables hasta que la diferencia entre la eliminación de una variable y otra fuese muy reducida. Sin embargo esto permitió resolver el problema presentado por el software LISREL 8.7 (versión del estudiante).

Por otra parte, en cuanto a la reespecificación del modelo que es sugerida por un índice de modificación, esta no se realiza, debido a que en esta investigación se limita únicamente a la evaluación del examen, de manera que los resultados presentados se analizan de forma general y desde un punto de vista estadístico, con el fin de que al presentar los resultados a quienes elaboraron el modelo, les sea de utilidad y les permita inferir y tomar decisiones acerca del mismo.

ANEXOS

ANEXO A

Texto utilizado en el examen de colocación

There are more slaves today than were seized from Africa in four centuries of the trans-Atlantic slave trade. The modern commerce in humans rivals illegal drug trafficking in its global reach—and in the destruction of lives.

Sherwood Castle, headquarters to Milorad Milakovic, the former railway official who rose to become a notorious slave trafficker in Bosnia, looms beside the main road just outside the northwest Bosnian town of Prijedor. Under stucco battlements, the entrance is guarded by well-muscled, heavily tattooed young men, while off to one side Milakovic's trio of pet Siberian tigers prowls their caged compound.

I arrived there alone one gray spring morning—alone because no local guide or translator dared accompany me—and found my burly 54-year-old host waiting for me at a table set for lunch beside a glassed-in aquamarine swimming pool.

The master of Sherwood has never been shy about his business. He once asked a dauntless human rights activist who has publicly detailed his record of buying women for his brothels in Prijedor: "Is it a crime to sell women? They sell footballers, don't they?"

Milakovic threatened to kill the activist for her outspokenness, but to me he sang a softer tune. Over a poolside luncheon of seafood salad and steak, we discussed the stream of young women fleeing the shattered economies of their home countries in the former Soviet bloc. Milakovic said he was eager to promote his scheme to legalize prostitution in Bosnia—"to stop the selling of people, because each of those girls is someone's child."

One such child is a nearsighted, chain-smoking blonde named Victoria, at 20 a veteran of the international slave trade. For three years of her life she was among the estimated 27 million men, women, and children in the world who are enslaved—physically confined or restrained and forced to work, or controlled through violence, or in some way treated as property.

Victoria's odyssey began when she was 17, fresh out of school in Chisinau, the decayed capital of the former Soviet republic of Moldova. "There was no work, no money," she explained simply. So when a friend—"at least I thought he was a friend"—suggested he could help her get a job in a factory in Turkey, she jumped at the idea and took up his offer to drive her there, through Romania. "But when I realized we had driven west, to the border with Serbia, I knew something was wrong."

It was too late. At the border she was handed over to a group of Serb men, who produced a new passport saying she was 18. They led her on foot into Serbia and raped her, telling her that she would be killed if she resisted. Then they sent her under guard to Bosnia, the Balkan republic being rebuilt under a torrent of international aid after its years of genocidal civil war.

Victoria was now a piece of property and, as such, was bought and sold by different brothel owners ten times over the next two years for an average price of \$1,500. Finally, four months pregnant and fearful of a forced abortion, she escaped. I found her hiding in the Bosnian city of Mostar, sheltered by a group of Bosnian women.

In a soft monotone she recited the names of clubs and bars in various towns where she had to dance seminaked, look cheerful, and have sex with any customer who wanted her for the price of a few packs of cigarettes. "The clubs were all awful, although the Artemdia, in Banja Luka, was the worst—all the customers were cops," she recalled.

Victoria was a debt slave. Payment for her services went straight to her owner of the moment to cover her "debt"—the amount he had paid to buy her from her previous owner. She was held in servitude unless or until the money she owed to whomever controlled her had been recovered, at which point she would be sold again and would begin to work off the purchase price paid by her new owner. Although slavery in its traditional form survives in many parts of the world, debt slavery of this kind, with variations, is the most common form of servitude today.

ANEXO B (parte I)

Análisis de fiabilidad del examen de colocación

XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 1
Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT
Date: may 29, 2007 Time: 14,11

***** ANALYSIS SUMMARY INFORMATION *****

Data (Input) File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT
Analysis Output File: D:\tesina\revisiones_de
tesis\datos\salida_xcalibre2.doc
Score Output File: NONE
Item Name File: NONE
Statistics Output File: NONE

***** CONFIGURATION INFORMATION *****

Item Parameter Priors: COMMON

Allow Priors to Float: YES

IRT Model Used: 3-parameter

Maximum Number of Loops: 12

Starting Prior Distribution Moments:

	Mean	SD
a	0.7500	0.1200
b	0.0000	1.0000
c	0.2500	0.0250

NOTE: *** will be printed when the c standard error value > 0.10

XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 2
Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT

Date: may 29, 2007

Time: 14,11

The input was from file: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT

The number of items was: 28

There was no item linkage

The key was:
5621432531443211134223514412

The numbers of alternatives were:
777774359556444434466666434

The inclusion specifications were:
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY

The maximum parameter change on loop 1 was 1.413
The maximum parameter change on loop 2 was 0.166
The maximum parameter change on loop 3 was 0.078
The maximum parameter change on loop 4 was 0.048

Mean Number-Correct Score = 15.808
Number-Correct Standard Deviation = 5.432
K-R 21 Reliability = 0.823

The number of examinees was 213

Final Parameter Summary Information:

	Mean	SD
Theta	0.00	1.00
a	0.91	0.12
b	0.31	0.92
c	0.23	0.01

XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 3
Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT

Date: may 29, 2007

Time: 14,11

FINAL ITEM PARAMETER ESTIMATES

Item	Lnk	Flg	a	b	c	Resid	PC	PBs	PBt	N	Item name
1			1.01	-0.52	0.23	0.44	0.71	0.56	0.61	213	
2			1.05	0.94	0.21	0.87	0.37	0.55	0.56	213	
3			1.12	-0.67	0.23	0.79	0.75	0.60	0.65	213	
4			1.18	-1.66	0.23	0.79	0.91	0.52	0.60	213	

5		1.12	-1.51	0.23	0.62	0.89	0.49	0.59	213
6		0.99	1.22	0.22	0.49	0.37	0.40	0.36	213
7		0.73	-0.10	0.24	0.67	0.66	0.39	0.33	213
8		0.94	0.69	0.23	0.32	0.47	0.48	0.44	213
9		1.02	-0.37	0.21	0.38	0.69	0.54	0.57	213
10		0.90	-0.38	0.23	0.36	0.69	0.47	0.51	213
11		0.89	-1.40	0.24	0.16	0.86	0.39	0.44	213
12		0.86	1.64	0.22	0.31	0.32	0.33	0.31	213
13		0.90	0.26	0.23	0.32	0.56	0.48	0.48	213
14		0.75	0.61	0.24	1.07	0.53	0.34	0.32	213
15		0.89	1.09	0.23	0.20	0.42	0.37	0.33	213
16		0.84	1.30	0.24	0.33	0.39	0.33	0.29	213
17		0.80	1.32	0.25	0.46	0.42	0.27	0.22	213
18		0.81	0.04	0.23	0.37	0.61	0.47	0.46	213
19		0.96	0.34	0.23	0.32	0.54	0.52	0.50	213
20		0.81	0.32	0.24	0.26	0.57	0.41	0.37	213
21		0.90	0.49	0.23	0.15	0.52	0.45	0.44	213
22		0.94	0.27	0.23	0.25	0.56	0.51	0.50	213
23		0.73	-0.35	0.24	0.67	0.69	0.35	0.36	213
24		0.91	0.79	0.23	0.23	0.46	0.42	0.40	213
25		0.96	1.10	0.23	0.45	0.40	0.40	0.36	213
26		0.97	1.74	0.21	0.68	0.27	0.36	0.32	213
27		0.79	-0.09	0.24	0.37	0.65	0.41	0.40	213
28	K	0.71	1.47	0.29	1.49	0.50	0.01	-0.05	213

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,11

ITEM PARAMETER ESTIMATES W/STANDARD ERRORS

Item	Lnk	Flg	a	a error	b	b error	c	c error	Resid	Item name
----	----	----	-----	-----	-----	-----	-----	-----	-----	-----
1			1.01	0.170	-0.52	0.128	0.23	***	0.44	
2			1.05	0.234	0.94	0.147	0.21	***	0.87	
3			1.12	0.161	-0.67	0.123	0.23	***	0.79	
4			1.18	0.157	-1.66	0.163	0.23	***	0.79	
5			1.12	0.151	-1.51	0.157	0.23	***	0.62	
6			0.99	0.250	1.22	0.176	0.22	***	0.49	
7			0.73	0.219	-0.10	0.163	0.24	***	0.67	
8			0.94	0.237	0.69	0.149	0.23	***	0.32	
9			1.02	0.175	-0.37	0.124	0.21	***	0.38	
10			0.90	0.183	-0.38	0.138	0.23	***	0.36	
11			0.89	0.145	-1.40	0.168	0.24	***	0.16	
12			0.86	0.257	1.64	0.232	0.22	***	0.31	
13			0.90	0.225	0.26	0.141	0.23	***	0.32	
14			0.75	0.249	0.61	0.174	0.24	***	1.07	
15			0.89	0.247	1.09	0.179	0.23	***	0.20	
16			0.84	0.249	1.30	0.202	0.24	***	0.33	
17			0.80	0.253	1.32	0.212	0.25	***	0.46	
18			0.81	0.218	0.04	0.150	0.23	***	0.37	
19			0.96	0.224	0.34	0.135	0.23	***	0.32	
20			0.81	0.237	0.32	0.156	0.24	***	0.26	
21			0.90	0.235	0.49	0.148	0.23	***	0.15	
22			0.94	0.223	0.27	0.137	0.23	***	0.25	
23			0.73	0.198	-0.35	0.163	0.24	***	0.67	
24			0.91	0.241	0.79	0.158	0.23	***	0.23	
25			0.96	0.246	1.10	0.171	0.23	***	0.45	
26			0.97	0.273	1.74	0.231	0.21	***	0.68	
27			0.79	0.213	-0.09	0.153	0.24	***	0.37	
28		K	0.71	0.267	1.47	0.255	0.29	***	1.49	

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,11

ITEM ANALYSIS

Item	Endorsement Rate										Item-Theta Corr.						
	1	2	3	4	5	6	7	8	9	Oth	1	2	3	4	5	6	7
8	9	Oth	-----														
1	5	13	2	1	71~	4	4				-40	-29	-25	-13	61~-13	-12	
2	0	5	26	1	15	37~	16				-12	-38	-7	-18	-28	56~	-8
3	2	75~	9	2	7	2	3				-28	65~-37	-17	-29	-11	-13	
4	91~	2	0	2	1	2	1				60~-21	0	-34	-22	-30	-21	
5	1	3	2	89~	2	1	2				-21	-29	-17	59~-31	-18	-19	
6	20	15	37~	29							-19	-16	36~	-9			
7	8	66~	26								-23	33~-22					
8	7	7	27	12	47				~		-31	1	-20	-17	44		
~																	
9	15	7	69~	2	7	0	0	0	0		-28	-36	57~-27	-12	0	0	
0	-9																
10	69~	5	11	3	12						51~-31	-27	-13	-19			
11	3	4	2	86~	4						-26	-21	-20	44~-17			
12	3	49	4	32~	11						-20	1	-10	31~-30			
13	9	11	56~	10	10	1				1	-28	-13	48~-17	-14	-25		
7																	
14	19	53~	7	21							-17	32~-33	-3				
15	42~	10	13	36							33~-11	-20	-14				
16	39~	44	12	5							29~	-4	-28	-15			
17	42~	20	38	0							22~-18	-6	-11				
18	16	23	61							~	-24	-32	46				
~																	
19	22	17	7	54						~	-28	-25	-17	50			
~																	
20	8	57~	21	14							-20	37~-19	-16				
21	16	52~	16	10	2	3					-21	44~-22	-18	-10	3		
22	2	35	56~	3	3	1					-17	-31	50~-16	-16	-17		
23	3	4	12	3	69~	8					-23	-13	-14	-13	36~-11		
24	46~	4	2	38	2	8					40~-25	-9	-19	-24	-6		
25	39	2	2	40~	15	1					-8	-21	-23	36~-18	-8		
26	23	19	31	27						~	-2	-26	-7	32			
~																	
27	65~	17	18								40~-20	-30					
28	18	50~	25	7							-11	-5~	15	-1			

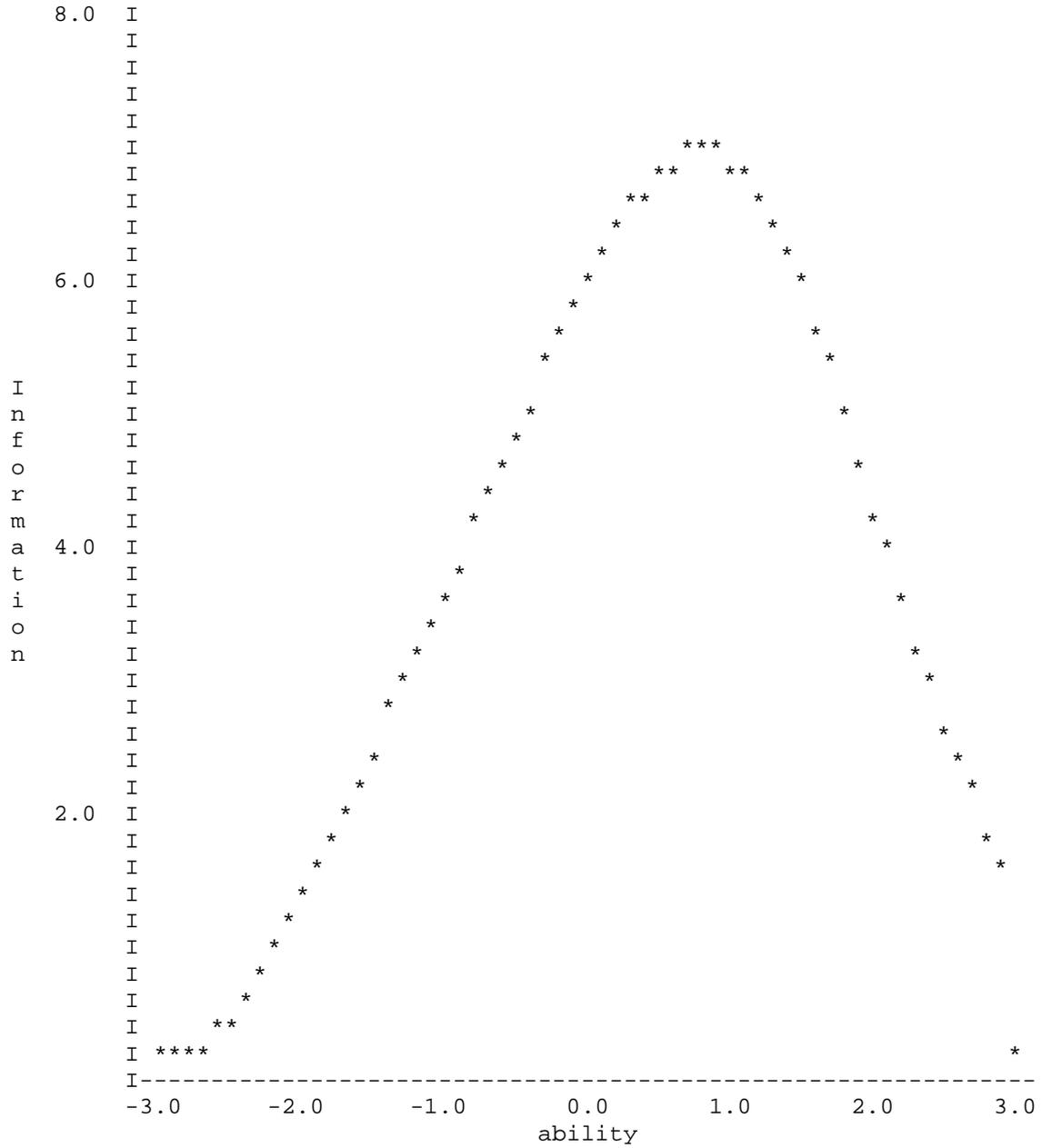
Test characteristics:

K-R 21
Reliability
0.823

Expected
Information
5.251

Average
Information
3.778

Test Information Curve

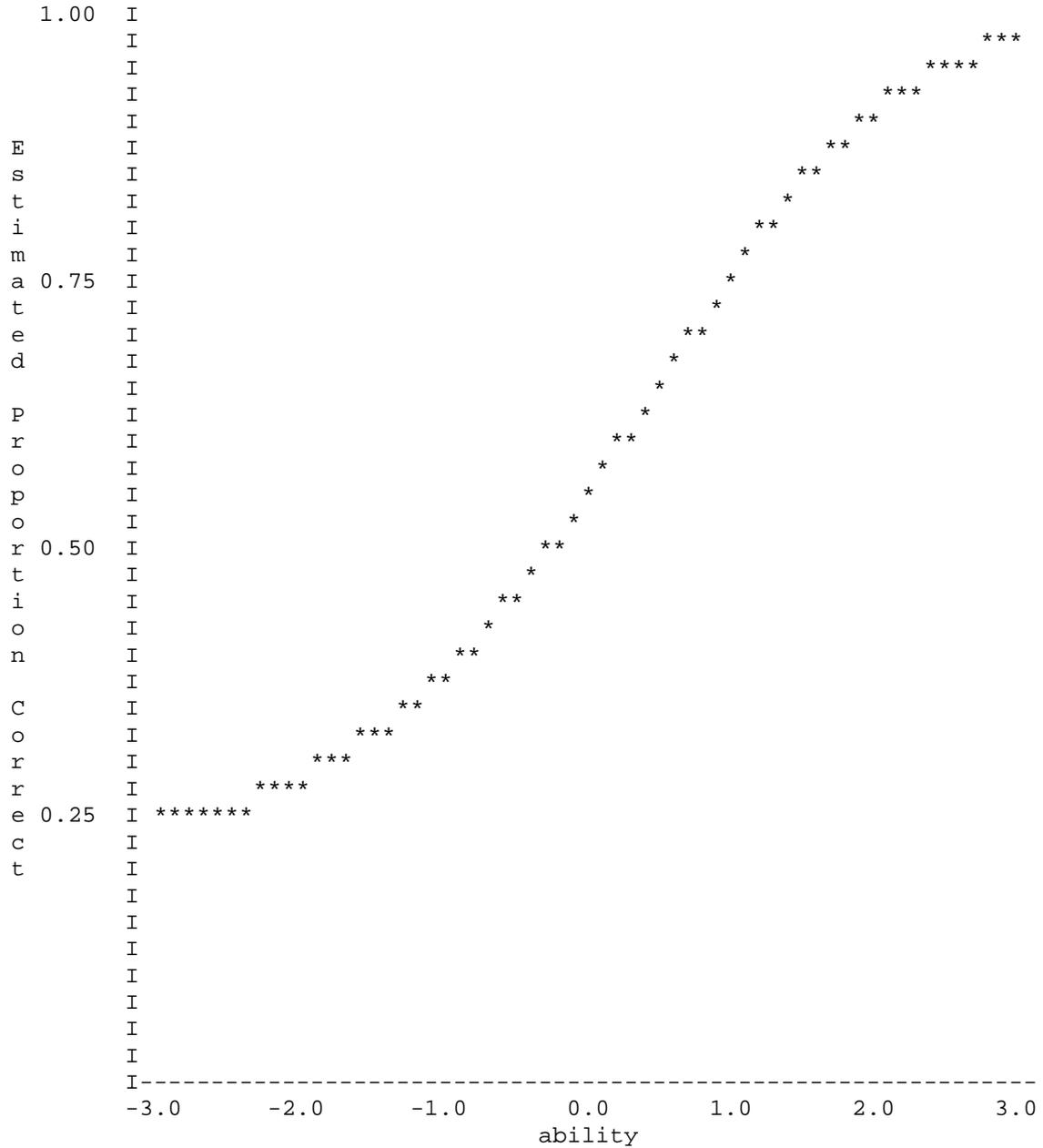


XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 7
 Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
 Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,11

Test Characteristic Curve



XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 1
 Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
 Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT

Date: may 29, 2007

Time: 14,03

***** ANALYSIS SUMMARY INFORMATION *****

Data (Input) File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT
Analysis Output File: D:\tesina\revisiones_de
tesis\datos\salida_xcalibre_ajustados.doc
Score Output File: NONE
Item Name File: NONE
Statistics Output File: NONE

***** CONFIGURATION INFORMATION *****

Item Parameter Priors: COMMON

Allow Priors to Float: YES

IRT Model Used: 3-parameter

Maximum Number of Loops: 12

Starting Prior Distribution Moments:

	Mean	SD
a	0.7500	0.1200
b	0.0000	1.0000
c	0.2500	0.0250

NOTE: *** will be printed when the c standard error value > 0.10

XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 2
Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT

Date: may 29, 2007

Time: 14,03

The input was from file: D:\tesina\revisiones_de
tesis\datos\datos_xcalibre2.DAT

The number of items was: 27

There was no item linkage

The key was:
562143253144321113422351441

The numbers of alternatives were:
777774359555644443446666643

The inclusion specifications were:
YYYYYYYYYYYYYYYYYYYYYYYYYYYY

The maximum parameter change on loop 1 was 1.492
The maximum parameter change on loop 2 was 0.180
The maximum parameter change on loop 3 was 0.086
The maximum parameter change on loop 4 was 0.058
The maximum parameter change on loop 5 was 0.043

Mean Number-Correct Score = 15.305
Number-Correct Standard Deviation = 5.455
K-R 21 Reliability = 0.834

The number of examinees was 213

Final Parameter Summary Information:

	Mean	SD
Theta	0.00	1.00
a	0.95	0.12
b	0.25	0.89
c	0.22	0.01

ANEXO B (parte II)

Análisis de fiabilidad del examen de colocación (al eliminar el ítem 28 del examen de colocación)

XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 3
 Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
 Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,03

FINAL ITEM PARAMETER ESTIMATES

Item	Lnk	Flg	a	b	c	Resid	PC	PBs	PBt	N	Item name
1			1.05	-0.52	0.22	0.39	0.71	0.56	0.61	213	
2			1.09	0.92	0.20	0.84	0.37	0.55	0.56	213	
3			1.15	-0.67	0.22	0.78	0.75	0.59	0.65	213	
4			1.22	-1.65	0.22	0.76	0.91	0.52	0.60	213	
5			1.16	-1.50	0.22	0.60	0.89	0.50	0.59	213	
6			1.03	1.19	0.22	0.46	0.37	0.40	0.36	213	
7			0.75	-0.12	0.24	0.73	0.66	0.38	0.33	213	
8			0.97	0.67	0.22	0.30	0.47	0.48	0.44	213	
9			1.05	-0.38	0.21	0.34	0.69	0.54	0.57	213	
10			0.94	-0.39	0.23	0.32	0.69	0.48	0.51	213	
11			0.91	-1.38	0.23	0.14	0.86	0.40	0.44	213	
12			0.89	1.59	0.22	0.31	0.32	0.33	0.31	213	
13			0.93	0.25	0.22	0.36	0.56	0.49	0.48	213	
14			0.77	0.58	0.24	1.10	0.53	0.33	0.32	213	
15			0.91	1.07	0.23	0.21	0.42	0.37	0.33	213	
16			0.87	1.27	0.23	0.37	0.39	0.34	0.29	213	
17			0.83	1.28	0.24	0.51	0.42	0.26	0.22	213	
18			0.84	0.02	0.22	0.39	0.61	0.48	0.46	213	
19			0.99	0.32	0.22	0.29	0.54	0.52	0.50	213	
20			0.83	0.31	0.23	0.29	0.57	0.41	0.37	213	
21			0.93	0.47	0.23	0.16	0.52	0.45	0.45	213	
22			0.98	0.26	0.22	0.22	0.56	0.51	0.50	213	
23			0.75	-0.37	0.23	0.71	0.69	0.35	0.36	213	
24			0.94	0.76	0.23	0.21	0.46	0.43	0.41	213	
25			0.99	1.07	0.22	0.42	0.40	0.40	0.36	213	
26			1.00	1.71	0.21	0.64	0.27	0.35	0.31	213	
27			0.81	-0.10	0.23	0.43	0.65	0.41	0.39	213	

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,03

ITEM PARAMETER ESTIMATES W/STANDARD ERRORS

Item	Lnk	Flg	a	a error	b	b error	c	c error	Resid	Item name
----	----	----	-----	-----	-----	-----	-----	-----	-----	-----
1			1.05	0.168	-0.52	0.124	0.22	***	0.39	
2			1.09	0.231	0.92	0.142	0.20	***	0.84	
3			1.15	0.160	-0.67	0.120	0.22	***	0.78	
4			1.22	0.158	-1.65	0.159	0.22	***	0.76	
5			1.16	0.152	-1.50	0.153	0.22	***	0.60	
6			1.03	0.249	1.19	0.168	0.22	***	0.46	
7			0.75	0.214	-0.12	0.159	0.24	***	0.73	
8			0.97	0.233	0.67	0.144	0.22	***	0.30	
9			1.05	0.174	-0.38	0.121	0.21	***	0.34	
10			0.94	0.180	-0.39	0.133	0.23	***	0.32	
11			0.91	0.145	-1.38	0.164	0.23	***	0.14	
12			0.89	0.257	1.59	0.221	0.22	***	0.31	
13			0.93	0.220	0.25	0.136	0.22	***	0.36	
14			0.77	0.245	0.58	0.169	0.24	***	1.10	
15			0.91	0.244	1.07	0.173	0.23	***	0.21	
16			0.87	0.248	1.27	0.194	0.23	***	0.37	
17			0.83	0.251	1.28	0.204	0.24	***	0.51	
18			0.84	0.213	0.02	0.145	0.22	***	0.39	
19			0.99	0.220	0.32	0.131	0.22	***	0.29	
20			0.83	0.232	0.31	0.152	0.23	***	0.29	
21			0.93	0.229	0.47	0.142	0.23	***	0.16	
22			0.98	0.218	0.26	0.131	0.22	***	0.22	
23			0.75	0.195	-0.37	0.158	0.23	***	0.71	
24			0.94	0.236	0.76	0.151	0.23	***	0.21	
25			0.99	0.244	1.07	0.164	0.22	***	0.42	
26			1.00	0.274	1.71	0.222	0.21	***	0.64	
27			0.81	0.208	-0.10	0.149	0.23	***	0.43	

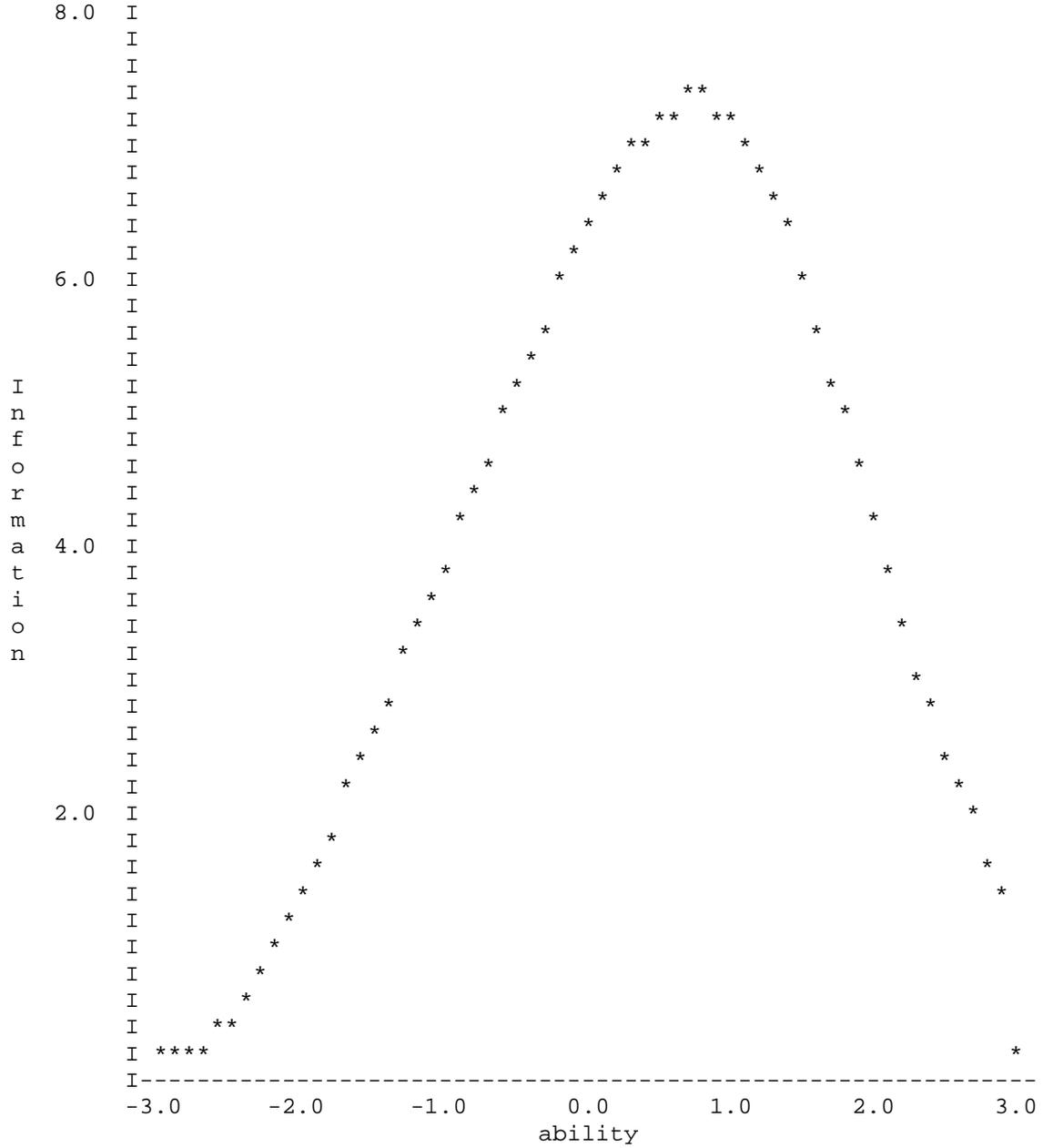
XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007 Time: 14,03

ITEM ANALYSIS

Item	Endorsement Rate										Item-Theta Corr.						
	1	2	3	4	5	6	7	8	9	Oth	1	2	3	4	5	6	7
8	9	Oth	-----														
1	5	13	2	1	71~	4	4				-39	-29	-25	-13	61~-13	-12	
2	0	5	26	1	15	37~	16				-12	-38	-7	-18	-28	56~	-8
3	2	75~	9	2	7	2	3				-28	65~-37	-17	-29	-11	-13	
4	91~	2	0	2	1	2	1				60~-21	0	-34	-22	-30	-21	
5	1	3	2	89~	2	1	2				-21	-30	-17	59~-31	-18	-20	
6	20	15	37~	29							-19	-15	36~-10				
7	8	66~	26								-23	33~-22					
8	7	7	27	12	47				~		-31	1	-20	-17	44		
~																	
9	15	7	69~	2	7	0	0	0	0		-28	-37	57~-27	-12	0	0	
0	-9																
10	69~	5	11	3	12						51~-31	-28	-13	-20			
11	3	4	2	86~	4						-26	-21	-20	44~-17			
12	3	49	4	32~	11						-20	1	-11	31~-30			
13	9	11	56~	10	10	1				1	-28	-13	48~-17	-14	-25		
8																	
14	19	53~	7	21							-17	32~-33	-3				
15	42~	10	13	36							33~-11	-20	-13				
16	39~	44	12	5							29~	-4	-28	-15			
17	42~	20	38	0							22~-18	-6	-11				
18	16	23	61							~	-24	-32	46				
~																	
19	22	17	7	54						~	-28	-25	-16	50			
~																	
20	8	57~	21	14							-19	37~-19	-16				
21	16	52~	16	10	2	3					-21	45~-22	-18	-10	3		
22	2	35	56~	3	3	1					-17	-31	50~-16	-16	-17		
23	3	4	12	3	69~	8					-23	-13	-14	-13	36~-11		
24	46~	4	2	38	2	8					41~-25	-9	-19	-24	-6		
25	39	2	2	40~	15	1					-8	-21	-23	36~-18	-8		
26	23	19	31	27						~	-2	-25	-7	31			
~																	
27	65~	17	18								39~-19	-30					

Test characteristics:	K-R 21	Expected	Average
	Reliability	Information	Information
	0.834	5.531	3.899

Test Information Curve

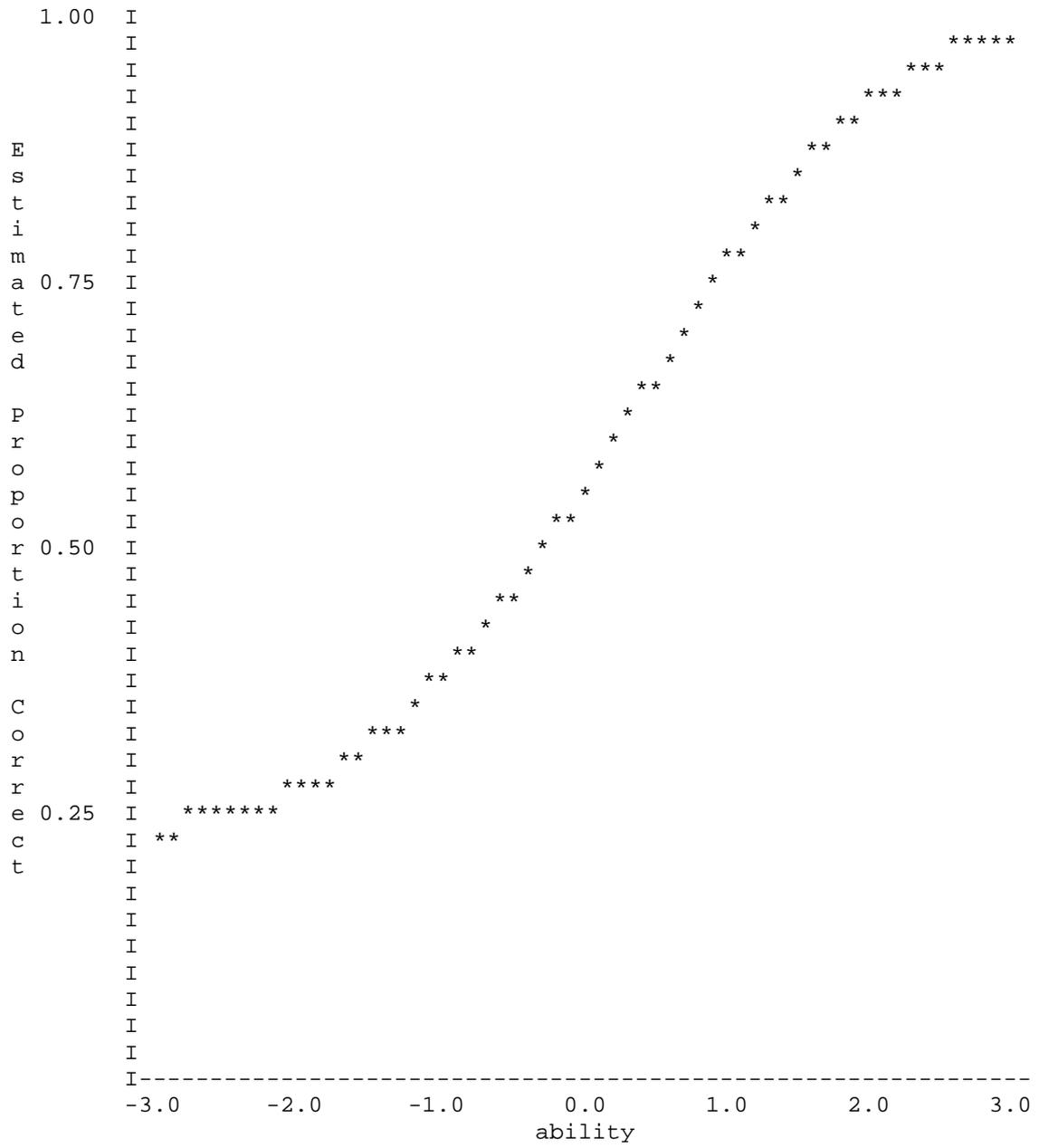


XCALIBRE (tm) for Windows95/NT -- Version 1.10 Page 7
 Copyright (c) 1995 by Assessment Systems Corporation, All Rights Reserved
 Marginal Maximum-Likelihood IRT Parameter Estimation Program

XCALIBRE Analysis from Data File: D:\tesina\revisiones_de
 tesis\datos\datos_xcalibre2.DAT
 Date: may 29, 2007

Time: 14,03

Test Characteristic Curve



ANEXO C (parte I)

Análisis Factorial confirmatorio (AFC) para el examen de colocación

Análisis factorial confirmatorio para el examen de colocación al curso de comprensión de lectura en Ingles. (Primera parte)

DATE: 5/6/2007

TIME: 21:55

L I S R E L 8.72

BY

Karl G. Jöreskog & Dag Sörbom

This program is published exclusively by
Scientific Software International, Inc.
7383 N. Lincoln Avenue, Suite 100
Lincolnwood, IL 60712, U.S.A.

Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140
Copyright by Scientific Software International, Inc., 1981-2005
Use of this program is subject to the terms specified in the
Universal Copyright Convention.
Website: www.ssicentral.com

The following lines were read from file C:\program
files\lisrel87s\FINAL2\simplis_final.spj:

AFC para el examen de comprensión de lectura

Observed Variables
CALIFICA ACCION RECEPTOR TIEMPO LUGAR CONCAUSA TOPICO CAUSALID
FACTUAL GENERALI ORDENAM1 ORDENAM3 ORDENAM2 ORDENAM4
Correlation Matrix from file 'C:\program
files\lisrel87s\FINAL2\corre.cor'
Sample Size = 213
Latent Variables Micro Macro Proces
Relationships
CALIFICA = Micro Macro Proces
RECEPTOR = Micro
TIEMPO = Micro
LUGAR = Micro
CONCAUSA = Micro
TOPICO = Micro
ACCION = Micro
CAUSALID = Proces
FACTUAL = Proces
GENERALI = Proces

ORDENAM1 = Macro
 ORDENAM2 = Macro
 ORDENAM3 = Macro
 ORDENAM4 = Macro
 Path Diagram
 Options: ND=3 ME=ML IT=250
 LISREL Output: MI RS SC EF
 End of Problem

AFC para el examen de comprensión de lectura

Correlation Matrix

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
CALIFICA	1.000					
ACCION	0.011	1.000				
RECEPTOR	-0.051	-0.255	1.000			
TIEMPO	0.050	0.206	0.189	1.000		
LUGAR	0.173	0.128	-0.282	-0.117	1.000	
CONCAUSA	0.139	-0.130	0.161	0.241	0.208	1.000
TOPICO	0.066	0.103	0.138	0.218	0.188	0.232
CAUSALID	0.022	0.082	0.055	0.032	0.043	0.020
FACTUAL	0.044	0.004	0.041	0.013	0.053	0.011
GENERALI	0.070	0.020	0.006	0.025	-0.083	0.073
ORDENAM1	0.068	-0.021	0.013	0.034	0.032	-0.005
ORDENAM3	0.084	0.059	0.052	0.107	0.027	0.066
ORDENAM2	0.035	0.004	0.024	-0.009	-0.050	0.006
ORDENAM4	-0.003	0.050	0.069	0.014	0.026	0.080

Correlation Matrix

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
TOPICO	1.000					
CAUSALID	0.104	1.000				
FACTUAL	0.117	0.217	1.000			
GENERALI	0.070	0.129	0.175	1.000		
ORDENAM1	-0.036	0.011	-0.022	-0.114	1.000	
ORDENAM3	0.107	0.061	0.027	0.024	0.114	1.000
ORDENAM2	0.031	0.001	0.082	0.011	-0.213	0.218
ORDENAM4	0.019	0.102	0.064	0.018	0.201	0.197

Correlation Matrix

	ORDENAM2	ORDENAM4
ORDENAM2	1.000	
ORDENAM4	0.237	1.000

AFC para el examen de comprensión de lectura

Parameter Specifications

LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	1	2	3
ACCION	0	4	0
RECEPTOR	5	0	0
TIEMPO	6	0	0
LUGAR	7	0	0
CONCAUSA	8	0	0
TOPICO	9	0	0
CAUSALID	0	0	10
FACTUAL	0	0	11
GENERALI	0	0	12
ORDENAM1	0	13	0
ORDENAM3	0	14	0
ORDENAM2	0	15	0
ORDENAM4	0	16	0

PHI

	Micro	Macro	Proces
	-----	-----	-----
Micro	0		
Macro	17	0	
Proces	18	19	0

THETA-DELTA

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
-----	-----	-----	-----	-----	-----
20	21	22	23	24	25

THETA-DELTA

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
-----	-----	-----	-----	-----	-----
26	27	28	29	30	31

THETA-DELTA

ORDENAM2	ORDENAM4
-----	-----
32	33

AFC para el examen de comprensión de lectura

Number of Iterations =43

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	0.783 (0.368) 2.123	0.564 (0.192) 3.928	0.484 (0.157) 3.076
ACCION	0.194 (0.140) 1.380	- -	- -
RECEPTOR	0.452 (0.178) 2.518	- -	- -
TIEMPO	0.566 (0.182) 3.094	- -	- -
LUGAR	1.677 (2.214) 0.738	- -	- -
CONCAUSA	0.532 (0.112) 4.738	- -	- -
TOPICO	0.348 (0.124) 2.795	- -	- -
CAUSALID	- -	- -	0.872 (0.214) 4.068
FACTUAL	- -	- -	1.655 (2.322) 0.732
GENERALI	- -	- -	0.449 (0.285) 1.572
ORDENAM1	- -	0.736 (0.194) 3.769	- -
ORDENAM3	- -	0.426 (0.242) 1.754	- -
ORDENAM2	- -	0.670 (0.226) 2.953	- -
ORDENAM4	- -	0.305	- -

(0.129)
2.372

PHI

	Micro	Macro	Proces
	-----	-----	-----
Micro	1.000		
Macro	0.737 (0.432) 1.706	1.000	
Proces	0.573 (0.421) 1.358	0.495 (0.733) 0.672	1.000

THETA-DELTA

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
-----	-----	-----	-----	-----	-----
0.317 (0.129) 3.231	0.912 (0.105) 9.173	0.748 (0.099) 7.556	0.632 (0.097) 6.517	-0.969 (0.100) -9.733	0.768 (0.097) 7.89

THETA-DELTA

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
-----	-----	-----	-----	-----	-----
0.781 (0.188) 4.155	0.287 (0.113) 2.528	-0.997 (0.103) -9.685	0.754 (0.101) 7.458	0.437 (0.097) 4.502	0.634 (0.135) 4.637

THETA-DELTA

ORDENAM2	ORDENAM4
-----	-----
0.512 (0.098) 5.216	0.895 (0.099) 9.04

Squared Multiple Correlations for X - Variables

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
-----	-----	-----	-----	-----	-----

0.683 0.088 0.252 0.368 1.969 0.232

Squared Multiple Correlations for X - Variables

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
-----	-----	-----	-----	-----	-----
0.219	0.713	1.997	0.246	0.563	0.366

Squared Multiple Correlations for X - Variables

ORDENAM2	ORDENAM4
-----	-----
0.488	0.105

Goodness of Fit Statistics

Degrees of Freedom = 72
Minimum Fit Function Chi-Square = 53.166 (P = 0.953)
Normal Theory Weighted Least Squares Chi-Square = 52.698 (P = 0.958)a
Estimated Non-centrality Parameter (NCP) = 0.0
90 Percent Confidence Interval for NCP = (0.0 ; 0.0)

Minimum Fit Function Value = 0.251
Population Discrepancy Function Value (F0) = 0.0
90 Percent Confidence Interval for F0 = (0.0 ; 0.0)
Root Mean Square Error of Approximation (RMSEA) = 0.0
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.0)
P-Value for Test of Close Fit (RMSEA < 0.05) = 1.00

Expected Cross-Validation Index (ECVI) = 0.651
90 Percent Confidence Interval for ECVI = (0.651 ; 0.651)
ECVI for Saturated Model = 0.991
ECVI for Independence Model = 0.468

Chi-Square for Independence Model with 91 Degrees of Freedom = 71.276
Independence AIC = 99.276
Model AIC = 118.698
Saturated AIC = 210.000
Independence CAIC = 160.334
Model CAIC = 262.621
Saturated CAIC = 667.936

Normed Fit Index (NFI) = 0.254
Non-Normed Fit Index (NNFI) = -0.207
Parsimony Normed Fit Index (PNFI) = 0.201
Comparative Fit Index (CFI) = 1.000
Incremental Fit Index (IFI) = -25.000
Relative Fit Index (RFI) = 0.0572

Critical N (CN) = 410.983

Root Mean Square Residual (RMR) = 0.0479 na
Standardized RMR = 0.0479

Goodness of Fit Index (GFI) = 0.966 a
Adjusted Goodness of Fit Index (AGFI) = 0.950 a
Parsimony Goodness of Fit Index (PGFI) = 0.662

Dado que existen estimaciones infractoras en las variables observadas LUGAR Y FACTUAL, se procede a fijar sus varianzas de error a 0.005.

Los resultados una vez fijadas las variables son los siguientes:

ANEXO D Parte II

DATE: 5/7/2007
TIME: 10:35

ANEXO C (parte II)

Análisis Factorial confirmatorio (AFC) para el examen de colocación

L I S R E L 8.72

BY

Karl G. Jöreskog & Dag Sörbom

This program is published exclusively by
Scientific Software International, Inc.
7383 N. Lincoln Avenue, Suite 100
Lincolnwood, IL 60712, U.S.A.
Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140
Copyright by Scientific Software International, Inc., 1981-2005
Use of this program is subject to the terms specified in the
Universal Copyright Convention.
Website: www.ssicentral.com

The following lines were read from file
C:\program files\lisrel87s\FINAL2\simplis_final_inf.spj:

AFC para el examen de comprensión de lectura con estimaciones infractoras corregidas

Observed Variables
CALIFICA ACCION RECEPTOR TIEMPO LUGAR CONCAUSA TOPICO CAUSALID
FACTUAL GENERALI ORDENAM1 ORDENAM3 ORDENAM2 ORDENAM4
Correlation Matrix from file 'C:\program
files\lisrel87s\FINAL2\corre.cor'

Sample Size = 213

Latent Variables Micro Macro Proces
Relationships

CALIFICA = Micro Macro Proces
 RECEPTOR = Micro
 TIEMPO = Micro
 LUGAR = Micro
 CONCAUSA = Micro
 TOPICO = Micro
 ACCION = Micro
 CAUSALID = Proces
 FACTUAL = Proces
 GENERALI = Proces
 ORDENAM1 = Macro
 ORDENAM2 = Macro
 ORDENAM3 = Macro
 ORDENAM4 = Macro

Path Diagram

Options: ND=3 ME=ML IT=250

LISREL Output: MI RS SC EF

Set the Error Variance of LUGAR to 0.005
 Set the Error Variance of FACTUAL to 0.005

End of Problem

AFC para el examen de comprensión de lectura

Correlation Matrix

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
CALIFICA	1.000					
ACCION	0.011	1.000				
RECEPTOR	-0.051	-0.255	1.000			
TIEMPO	0.050	0.206	0.189	1.000		
LUGAR	0.173	0.128	-0.282	-0.117	1.000	
CONCAUSA	0.139	-0.130	0.161	0.241	0.208	1.000
TOPICO	0.066	0.103	0.138	0.218	0.188	0.232
CAUSALID	0.022	0.082	0.055	0.032	0.043	0.020
FACTUAL	0.044	0.004	0.041	0.013	0.053	0.011
GENERALI	0.070	0.020	0.006	0.025	-0.083	0.073
ORDENAM1	0.068	-0.021	0.013	0.034	0.032	-0.005
ORDENAM3	0.084	0.059	0.052	0.107	0.027	0.066
ORDENAM2	0.035	0.004	0.024	-0.009	-0.050	0.006
ORDENAM4	-0.003	0.050	0.069	0.014	0.026	0.080

Correlation Matrix

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
TOPICO	1.000					
CAUSALID	0.104	1.000				
FACTUAL	0.117	0.217	1.000			
GENERALI	0.070	0.129	0.175	1.000		
ORDENAM1	-0.036	0.011	-0.022	-0.114	1.000	
ORDENAM3	0.107	0.061	0.027	0.024	0.114	1.000
ORDENAM2	0.031	0.001	0.082	0.011	-0.213	0.218

ORDENAM4 0.019 0.102 0.064 0.018 0.201 0.197

Correlation Matrix

	ORDENAM2	ORDENAM4
ORDENAM2	1.000	
ORDENAM4	0.237	1.000

AFC para el examen de comprensión de lectura

Parameter Specifications

LAMBDA-X

	Micro	Macro	Proces
CALIFICA	1	2	3
ACCION	0	4	0
RECEPTOR	5	0	0
TIEMPO	6	0	0
LUGAR	7	0	0
CONCAUSA	8	0	0
TOPICO	9	0	0
CAUSALID	0	0	10
FACTUAL	0	0	11
GENERALI	0	0	12
ORDENAM1	0	13	0
ORDENAM3	0	14	0
ORDENAM2	0	15	0
ORDENAM4	0	16	0

PHI

	Micro	Macro	Proces
Micro	0		
Macro	17	0	
Proces	18	19	0

THETA-DELTA

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
20	21	22	23	0	24

THETA-DELTA

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
25	26	0	27	28	29

THETA-DELTA

ORDENAM2	ORDENAM4
-----	-----
32	33

AFC para el examen de comprensión de lectura

Number of Iterations =35

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	0.788 (0.236) 3.329	0.612 (0.216) 1.835	0.379 (0.129) 1.937
ACCION	0.382 (0.139) 2.742	- -	- -
RECEPTOR	0.468 (0.183) 2.569	- -	- -
TIEMPO	0.593 (0.176) 3.238	- -	- -
LUGAR	0.975 (0.067) 14.532	- -	- -
CONCAUSA	0.532 (0.119) 4.469	- -	- -
TOPICO	0.346 (0.183) 3.268	- -	- -
CAUSALID	- -	- -	0.888 (0.298) 2.979
FACTUAL	- -	- -	0.978 (0.064) 15.278
GENERALI	- -	- -	0.426 (0.162)

			2.623
ORDENAM1	- -	0.724 (0.198) 3.656	- -
ORDENAM3	- -	0.429 (0.234) 1.834	- -
ORDENAM2	- -	0.670 (0.213) 3.143	- -
ORDENAM4	- -	0.187 (0.086) 2.163	- -

PHI

	Micro	Macro	Proces
	-----	-----	-----
Micro	1.000		
Macro	0.723 (0.357) 2.025	1.000	
Proces	0.386 (0.217) 1.768	0.472 (0.245) 2.131	1.000

THETA-DELTA

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
-----	-----	-----	-----	-----	-----
0.362 (0.118) 3.059	0.768 (0.127) 6.051	0.682 (0.119) 5.735	0.557 (0.103) 5.412	0.005	0.668 (0.127) 5.259

THETA-DELTA

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
-----	-----	-----	-----	-----	-----
0.854 (0.163) 5.228	0.312 (0.116) 2.689	0.005	0.687 (0.107) 6.458	0.376 (0.094) 4.001	0.671 (0.136) 4.943

THETA-DELTA

ORDENAM2	ORDENAM4
0.524	0.965
(0.098)	(0.139)
5.346	6.947

Squared Multiple Correlations for X - Variables

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
0.638	0.688	0.318	0.443	0.995	0.332

Squared Multiple Correlations for X - Variables

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
0.146	0.688	0.995	0.313	0.624	0.329

Squared Multiple Correlations for X - Variables

ORDENAM2	ORDENAM4
0.476	0.035

Goodness of Fit Statistics

Degrees of Freedom = 72

Minimum Fit Function Chi-Square = 48.642 (P = 0.984)

Normal Theory Weighted Least Squares Chi-Square = 48.431 (P = 0.985)

Estimated Non-centrality Parameter (NCP) = 0.0

90 Percent Confidence Interval for NCP = (0.0 ; 0.0)

Minimum Fit Function Value = 0.229

Population Discrepancy Function Value (F0) = 0.0

90 Percent Confidence Interval for F0 = (0.0 ; 0.0)

Root Mean Square Error of Approximation (RMSEA) = 0.0

90 Percent Confidence Interval for RMSEA = (0.0 ; 0.0)

P-Value for Test of Close Fit (RMSEA < 0.05) = 1.00

Expected Cross-Validation Index (ECVI) = 0.651

90 Percent Confidence Interval for ECVI = (0.651 ; 0.651)

ECVI for Saturated Model = 0.991

ECVI for Independence Model = 0.468

Chi-Square for Independence Model with 91 Degrees of Freedom = 71.276

Independence AIC = 99.276

Model AIC = 114.431

Saturated AIC = 210.000

Independence CAIC = 160.334

Model CAIC = 258.354
 Saturated CAIC = 667.936

Normed Fit Index (NFI) = 0.918
 Non-Normed Fit Index (NNFI) = 0.897
 Parsimony Normed Fit Index (PNFI) = 0.251
 Comparative Fit Index (CFI) = 1.000
 Incremental Fit Index (IFI) = 1.147
 Relative Fit Index (RFI) = 0.637

Critical N (CN) = 449.117

Root Mean Square Residual (RMR) = 0.0467
 Standardized RMR = 0.0467
 Goodness of Fit Index (GFI) = 0.968
 Adjusted Goodness of Fit Index (AGFI) = 0.954
 Parsimony Goodness of Fit Index (PGFI) = 0.664

AFC para el examen de comprensión de lectura

Fitted Covariance Matrix

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----
CALIFICA	1.000					
ACCION	0.051	1.000				
RECEPTOR	0.009	0.022	1.000			
TIEMPO	0.004	0.009	0.010	1.000		
LUGAR	0.011	0.025	0.027	0.012	1.000	
CONCAUSA	0.004	0.009	0.009	0.004	0.011	1.000
TOPICO	0.029	0.067	0.071	0.031	0.083	0.029
CAUSALID	0.017	0.021	0.039	0.017	0.045	0.016
FACTUAL	0.013	0.016	0.030	0.013	0.035	0.012
GENERALI	0.012	0.014	0.027	0.011	0.031	0.011
ORDENAM1	0.001	0.001	0.001	0.000	0.001	0.000
ORDENAM3	-0.078	-0.057	-0.033	-0.014	-0.038	-0.013
ORDENAM2	0.018	0.014	0.008	0.003	0.009	0.003
ORDENAM4	0.028	0.020	0.012	0.005	0.014	0.005

Fitted Covariance Matrix

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	1.000					
CAUSALID	0.120	1.000				
FACTUAL	0.093	0.004	1.000			
GENERALI	0.082	0.004	0.003	1.000		
ORDENAM1	0.002	0.001	0.000	0.000	1.000	
ORDENAM3	-0.102	-0.032	-0.025	-0.022	-0.002	1.000
ORDENAM2	0.024	0.007	0.006	0.005	0.000	0.021
ORDENAM4	0.036	0.011	0.009	0.008	0.001	0.031

Fitted Covariance Matrix

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	1.000	
ORDENAM4	0.007	1.000

Fitted Residuals

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----
CALIFICA	0.000					
ACCION	0.040	0.000				
RECEPTOR	-0.041	-0.033	0.000			
TIEMPO	0.054	0.003	0.079	0.000		
LUGAR	0.184	0.003	-0.055	-0.005	0.000	
CONCAUSA	0.135	0.121	0.052	0.037	0.019	0.000
TOPICO	-0.037	0.036	0.033	0.013	0.006	0.003
CAUSALID	0.005	-0.011	0.016	0.015	0.003	0.004
FACTUAL	0.058	-0.020	0.071	-0.100	0.018	0.001
GENERALI	0.058	0.034	-0.021	0.014	-0.052	-0.084
ORDENAM1	0.067	-0.020	0.012	0.034	0.033	-0.005
ORDENAM3	0.006	0.002	0.018	0.121	0.012	0.079
ORDENAM2	0.116	0.017	0.031	-0.013	0.040	0.002
ORDENAM4	0.025	-0.070	-0.057	0.019	-0.040	0.085

Fitted Residuals

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	0.000					
CAUSALID	0.016	0.000				
FACTUAL	-0.024	0.013	0.000			
GENERALI	0.012	0.032	-0.022	0.000		
ORDENAM1	-0.038	0.012	0.021	-0.114	0.000	
ORDENAM3	0.005	0.029	0.051	0.003	0.016	0.000
ORDENAM2	0.055	0.009	0.088	-0.016	-0.014	0.039
ORDENAM4	0.018	-0.091	0.055	0.010	0.011	0.066

Fitted Residuals

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	0.000	
ORDENAM4	-0.030	0.000

Summary Statistics for Fitted Residuals

Smallest Fitted Residual = -0.184
 Median Fitted Residual = 0.000
 Largest Fitted Residual = 0.121

Standardized Residuals

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----
CALIFICA	-					

ACCION	0.808	- -					
RECEPTOR	-0.636	-0.512	- -				
TIEMPO	0.791	0.044	1.200	- -			
LUGAR	1.836	0.040	-0.872	-0.080	- -		
CONCAUSA	1.323	0.728	0.779	0.539	0.295	- -	
TOPICO	-0.434	0.834	0.949	0.303	0.183	0.064	
CAUSALID	0.104	-0.173	0.280	0.242	0.046	0.062	
FACTUAL	0.304	-0.310	0.166	-1.151	0.305	0.019	
GENERALI	0.966	0.525	-0.335	0.211	-0.860	-1.287	
ORDENAM1	0.109	-0.304	0.179	0.498	0.488	-0.073	
ORDENAM3	1.973	0.058	0.305	1.105	0.200	1.244	
ORDENAM2	0.173	0.273	0.471	-0.184	0.610	0.036	
ORDENAM4	0.432	-0.134	-0.859	0.277	-0.602	1.248	

Standardized Residuals

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	- -					
CAUSALID	0.674	- -				
FACTUAL	-0.897	0.261	- -			
GENERALI	0.443	0.583	-0.364	- -		
ORDENAM1	-0.708	0.174	0.311	-1.682	- -	
ORDENAM3	0.157	0.531	1.945	0.044	0.274	- -
ORDENAM2	0.053	0.134	1.300	-0.233	-0.202	0.696
ORDENAM4	0.353	-1.384	0.821	0.155	0.158	1.265

Standardized Residuals

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	- -	
ORDENAM4	-0.448	- -

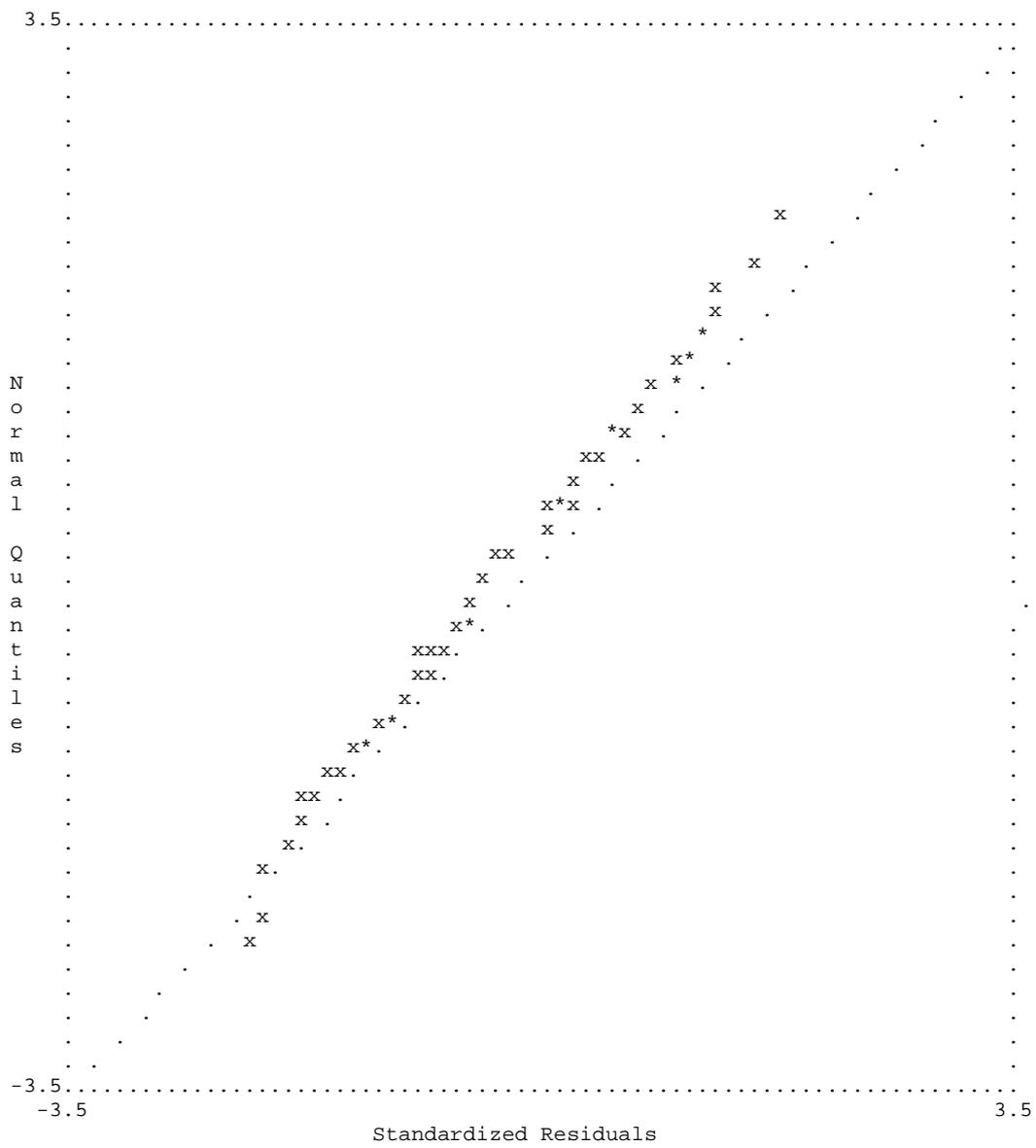
Summary Statistics for Standardized Residuals

Smallest Standardized Residual = -1.384
Median Standardized Residual = 0.000
Largest Standardized Residual = 1.973

Largest Positive Standardized Residuals
Residual for ORDENAM3 and CALIFICA 1.973

AFC para el examen de comprensión de lectura

Qplot of Standardized Residuals



AFC para el examen de comprensión de lectura

Modification Indices and Expected Change

Modification Indices for LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	- -	- -	- -
ACCION	0.427	- -	1.828
RECEPTOR	- -	0.005	0.008
TIEMPO	- -	0.104	0.021
LUGAR	- -	3.503	1.320
CONCAUSA	- -	2.450	0.022
TOPICO	- -	1.059	1.540
CAUSALID	0.074	0.593	- -
FACTUAL	0.258	1.282	- -
GENERALI	0.607	0.100	- -
ORDENAM1	1.627	- -	0.207
ORDENAM3	8.125	- -	0.202
ORDENAM2	2.582	- -	0.340
ORDENAM4	0.825	- -	0.860

Expected Change for LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	- -	- -	- -
ACCION	0.128	- -	-0.103
RECEPTOR	- -	0.014	0.006
TIEMPO	- -	0.062	0.009
LUGAR	- -	0.420	0.087
CONCAUSA	- -	0.300	0.009
TOPICO	- -	0.407	-0.283
CAUSALID	-0.108	-0.235	- -
FACTUAL	0.152	0.290	- -
GENERALI	0.208	-0.077	- -
ORDENAM1	0.218	- -	-0.028
ORDENAM3	-0.096	- -	0.049
ORDENAM2	0.329	- -	-0.036
ORDENAM4	0.161	- -	0.060

Standardized Expected Change for LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	- -	- -	- -
ACCION	0.128	- -	-0.103
RECEPTOR	- -	0.014	0.006
TIEMPO	- -	0.062	0.009
LUGAR	- -	0.420	0.087
CONCAUSA	- -	0.300	0.009
TOPICO	- -	0.407	-0.283

CAUSALID	-0.108	-0.235	- -
FACTUAL	0.152	0.290	- -
GENERALI	0.208	-0.077	- -
ORDENAM1	0.218	- -	-0.028
ORDENAM3	-0.096	- -	0.049
ORDENAM2	0.329	- -	-0.036
ORDENAM4	0.161	- -	0.060

Completely Standardized Expected Change for LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	- -	- -	- -
ACCION	0.128	- -	-0.103
RECEPTOR	- -	0.014	0.006
TIEMPO	- -	0.062	0.009
LUGAR	- -	0.420	0.087
CONCAUSA	- -	0.300	0.009
TOPICO	- -	0.407	-0.283
CAUSALID	-0.108	-0.235	- -
FACTUAL	0.152	0.290	- -
GENERALI	0.208	-0.077	- -
ORDENAM1	0.218	- -	-0.028
ORDENAM3	-0.096	- -	0.049
ORDENAM2	0.329	- -	-0.036
ORDENAM4	0.161	- -	0.060

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----
CALIFICA	- -					
ACCION	0.511	- -				
RECEPTOR	0.476	0.243	- -			
TIEMPO	0.863	0.001	1.440	- -		
LUGAR	2.612	0.024	0.760	0.006	- -	
CONCAUSA	2.798	3.851	0.607	0.291	0.087	- -
TOPICO	2.242	0.661	0.901	0.092	0.033	0.004
CAUSALID	0.000	0.002	0.064	0.058	0.019	0.005
FACTUAL	0.843	0.102	1.359	2.286	0.046	0.000
GENERALI	0.999	0.372	0.091	0.052	0.693	1.539
ORDENAM1	1.187	0.093	0.068	0.285	0.162	0.001
ORDENAM3	0.022	0.003	0.119	3.809	0.205	1.232
ORDENAM2	1.963	0.075	0.127	0.016	0.495	0.015
ORDENAM4	0.288	1.286	0.667	0.084	0.570	1.481

Modification Indices for THETA-DELTA

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	- -					
CAUSALID	0.439	- -				
FACTUAL	0.624	0.068	- -			
GENERALI	0.030	0.339	0.133	- -		

ORDENAM1	0.203	0.018	0.114	2.697	- -	
ORDENAM3	0.136	0.255	0.592	0.001	0.075	- -
ORDENAM2	0.430	0.009	1.515	0.043	0.041	0.485
ORDENAM4	0.397	2.006	0.817	0.037	0.025	1.600

Modification Indices for THETA-DELTA

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	- -	
ORDENAM4	0.200	- -

Expected Change for THETA-DELTA

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----
CALIFICA	- -					
ACCION	0.066	- -				
RECEPTOR	0.049	-0.035	- -			
TIEMPO	0.064	0.003	0.083	- -		
LUGAR	0.210	0.011	-0.061	-0.006	- -	
CONCAUSA	0.135	0.136	0.054	0.037	0.021	- -
TOPICO	-0.214	0.077	-0.098	-0.026	0.021	0.005
CAUSALID	0.000	0.003	0.019	0.017	0.011	0.005
FACTUAL	0.075	0.023	0.086	-0.107	0.016	0.001
GENERALI	0.078	0.043	-0.022	0.016	0.061	0.087
ORDENAM1	0.083	-0.022	0.018	0.037	0.028	0.003
ORDENAM3	-0.019	0.007	-0.025	0.138	-0.034	0.078
ORDENAM2	0.155	0.020	0.025	-0.009	0.049	0.008
ORDENAM4	0.043	-0.082	0.057	0.020	0.053	0.084

Expected Change for THETA-DELTA

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	- -					
CAUSALID	0.098	- -				
FACTUAL	0.108	0.024	- -			
GENERALI	0.023	0.049	0.028	- -		
ORDENAM1	0.035	0.009	0.023	0.113	- -	
ORDENAM3	0.043	0.041	0.058	-0.002	0.020	- -
ORDENAM2	-0.052	0.007	0.085	0.014	-0.014	0.054
ORDENAM4	0.052	-0.098	0.062	0.013	0.011	0.103

Expected Change for THETA-DELTA

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	- -	
ORDENAM4	0.031	- -

Completely Standardized Expected Change for THETA-DELTA

	CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
	-----	-----	-----	-----	-----	-----

CALIFICA	- -						
ACCION	0.066	- -					
RECEPTOR	0.049	-0.035	- -				
TIEMPO	0.064	0.003	0.083	- -			
LUGAR	0.210	0.011	-0.061	-0.006	- -		
CONCAUSA	0.135	0.136	0.054	0.037	0.021	- -	
TOPICO	-0.214	0.077	-0.098	-0.026	0.021	0.005	
CAUSALID	0.000	0.003	0.019	0.017	0.011	0.005	
FACTUAL	0.075	0.023	0.086	-0.107	0.016	0.001	
GENERALI	0.078	0.043	-0.022	0.016	0.061	0.087	
ORDENAM1	0.083	-0.022	0.018	0.037	0.028	0.003	
ORDENAM3	-0.019	0.007	-0.025	0.138	-0.034	0.078	
ORDENAM2	0.155	0.020	0.025	-0.009	0.049	0.008	
ORDENAM4	0.043	-0.082	0.057	0.020	0.053	0.084	

Completely Standardized Expected Change for THETA-DELTA

	TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
	-----	-----	-----	-----	-----	-----
TOPICO	- -					
CAUSALID	0.098	- -				
FACTUAL	0.108	0.024	- -			
GENERALI	0.023	0.049	0.028	- -		
ORDENAM1	0.035	0.009	0.023	0.113	- -	
ORDENAM3	0.043	0.041	0.058	-0.002	0.020	- -
ORDENAM2	-0.052	0.007	0.085	0.014	-0.014	0.054
ORDENAM4	0.052	-0.098	0.062	0.013	0.011	0.103

Completely Standardized Expected Change for THETA-DELTA

	ORDENAM2	ORDENAM4
	-----	-----
ORDENAM2	- -	
ORDENAM4	0.031	- -

Maximum Modification Index is 8.125 for Element (12, 1) of LAMBDA X

AFC para el examen de comprensión de lectura

Standardized Solution

LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	0.788	0.612	0.379
ACCION	0.382	- -	- -
RECEPTOR	0.468	- -	- -
TIEMPO	0.593	- -	- -
LUGAR	0.975	- -	- -
CONCAUSA	0.532	- -	- -
TOPICO	0.346	- -	- -
CAUSALID	- -	- -	0.888

FACTUAL	- -	- -	0.978
GENERALI	- -	- -	0.426
ORDENAM1	- -	0.724	- -
ORDENAM3	- -	0.429	- -
ORDENAM2	- -	0.670	- -
ORDENAM4	- -	0.187	- -

PHI

	Micro	Macro	Proces
	-----	-----	-----
Micro	1.000		
Macro	0.723	1.000	
Proces	0.386	0.472	1.000

AFC para el examen de comprensión de lectura

Completely Standardized Solution

LAMBDA-X

	Micro	Macro	Proces
	-----	-----	-----
CALIFICA	0.788	0.612	0.379
ACCION	0.382	- -	- -
RECEPTOR	0.468	- -	- -
TIEMPO	0.593	- -	- -
LUGAR	0.975	- -	- -
CONCAUSA	0.532	- -	- -
TOPICO	0.346	- -	- -
CAUSALID	- -	- -	0.888
FACTUAL	- -	- -	0.978
GENERALI	- -	- -	0.426
ORDENAM1	- -	0.724	- -
ORDENAM3	- -	0.429	- -
ORDENAM2	- -	0.670	- -
ORDENAM4	- -	0.187	- -

PHI

	Micro	Macro	Proces
	-----	-----	-----
Micro	1.000		
Macro	0.723	1.000	
Proces	0.386	0.472	1.000

THETA-DELTA

CALIFICA	ACCION	RECEPTOR	TIEMPO	LUGAR	CONCAUSA
-----	-----	-----	-----	-----	-----
0.362	0.768	0.682	0.557	0.005	0.668

THETA-DELTA

TOPICO	CAUSALID	FACTUAL	GENERALI	ORDENAM1	ORDENAM3
-----	-----	-----	-----	-----	-----

0.854 0.312 0.005 0.687 0.376 0.671

THETA-DELTA

ORDENAM2	ORDENAM4
-----	-----
0.524	0.965

Time used: 0.094 Seconds

BIBLIOGRAFÍA

1. Abad, F., et al (2006). **Introducción a la Psicometría: Teoría Clásica de los Tests y Teoría de Respuesta al Ítem**. Madrid: UAM.
2. Bisquerra, R. (1989). **Introducción conceptual al análisis multivariado. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD**. Vol. II, PPU, Barcelona.
3. Castañeda, S. et al (1993). **Un marco de trabajo neurocomputacional para el estudio de la estructuración del conocimiento a partir de lo leído**. Revista Latina de pensamiento y Lenguaje, vol. 1 núm 2.
4. Corral Verdugo, V., et al. (2001). **Análisis cuantitativo de variables latentes**. México: UniSon.
5. Fernández Ballesteros, R. (1997). **Evaluación psicológica y tests**. Madrid: TEA.
6. García Jurado, R. (2003). **Factores instruccionales que inciden en el aprendizaje de la comprensión de lectura en inglés como lengua extranjera en alumnos de la UNAM**. Campus Acatlán. Unpublished master's thesis, Facultad de Psicología, UNAM.
7. Hair, J., et al. (1999). **Análisis multivariante**. México: Prentice Hall.
8. Jöreskog, Kart G. y Dag Sörbon (1993). **Lisrel 8: Structural Equation Modeling with the Simplis Command Language**. U.S.A.: Lawrence Erlbaum Associates, Inc., Publishers.
9. Kleinbaum, David G (1998). **Applied regression analysis and other multivariate methods**. USA: PWS-KENT.
10. Lévy Manginm, J., et al. (2003). **Análisis multivariable para las ciencias sociales**. España: Prentice Hall.
11. Loehlin, C. (1992). **Latent Variable Models: An introduction to factor; path and structural analysis**. Hillsdale, N.J, U.S.A. Lawrence Erlbaum.
12. Martínez, R. y Moreno, R. (2002). **Integración de teoría sustantiva, diseños de pruebas y modelos de análisis en la medición psicológica**. En A. Bazán y A. J. Arce (eds.)Estrategias de evaluación y medición del comportamiento en psicología (87-119). Ciudad Obregón, Méx.: Instituto Tecnológico de Sonora.
13. Nunnally, Jum C. e Ira H. Bernstein. (1995). **Teoría psicométrica**. México: Mc Graw Hill. Segunda edición.

14. Reimers, F. (2003). **El contexto social de la evaluación educativa en América Latina**. Revista Latinoamericana de Estudios Educativos, vol. 33, núm. 3.
15. TenBrink, T. (2002). **Estrategias de enseñanza. Guía para una mejor instrucción**. México: Limusa.
16. Van Dijk, T. & Kintsch, W. (1983): **Strategies of Discourse Compreitension**. Londres. Academia Press.
17. Hambleton, R.K., Swaminathan, H., Rogers, H.J.(1991), **Fundamentals of Item Response Theory**. London, Sage Publications.
18. Murat F.(1973) **Evaluación Del Comportamiento Humano**. Tomo I. Capítulo I. Dirección Nacional de Publicaciones, Universidad Nacional de Córdoba.
19. Muñiz, J. (1998). **La teoría clásica de los tests**. Madrid: Pirámide.
20. Muñiz, J. (1997). **Introducción a la teoría de respuesta a los ítems**. Madrid: Pirámide.
21. Cohen, Ronald Jay; Swerdlik, Mark E: (2000). **Pruebas y Evaluación Psicológicas. Introducción a las Pruebas y a la Medición**. México, McGraw-Hill. Cuarta Edición.
22. Pardo, A. y Ruiz, M. A. (2002). **SPSS 11. Guía para el análisis de datos**. Madrid. McGraw-Hill.
23. Álvaro Page, Mariano(1990). **Hacia un modelo causal del rendimiento académico**, Madrid: Centro de Publicaciones del Ministerio de Educación y Ciencia (C.I.D.E).
24. BATISTA, J.M. y COENDERS, G. (2000). **Modelos de ecuaciones estructurales. Cuadernos de estadística 6**. La Muralla, Madrid