UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE EDUCACIÓN

UN MODELO DE MEDIDA CON INTERACCIÓN

TESIS DOCTORAL

DIRECTOR DE LA TESIS:
DR. JOSÉ LUIS GAVIRIA S.

PRESENTA:
MARIA TRIGUEROS GAISMAN

Madrid, España, 1999

AGRADECIMIENTOS

En el momento en que este trabajo llega a su fin y que la posibilidad de titularme se ve más cercana, caigo en la cuenta de que ha sido un gran privilegio para mi poder obtener un grado de doctor en la Universidad Complutense de Madrid, haciendo todo el trabajo desde México. Esto no hubiera sido posible de no haberse firmado el convenio de colaboración entre esta Universidad y la Universidad Anáhuac en México. Por mi parte quisiera agradecer a ambas instituciones la oportunidad y el enriquecimiento que ésta significa para quienes, como yo, no podemos viajar a otro país para realizar estudios de postgrado.

A través de este convenio tuve oportunidad de conocer y trabajar con un grupo selecto de profesores que siempre estuvieron dispuestos a brindarme ayuda. De entre ellos quisiera agradecer en forma especial la destacada labor que hicieron conmigo el Dr. Arturo de la Orden y el Dr. José Luis Gaviria. Los consejos de ambos, su apoyo incondicional y su paciencia y amistad contribuyeron a que la labor involucrada en el proyecto se hiciera menos pesada; la acertada y a la vez crítica dirección del Dr. Gaviria hizo, además, posible la culminación de este proyecto de doctorado.

Quisiera agradecer también el apoyo y dirección de la Dra. Guillermina Waldegg. Sus acertados comentarios tanto a la parte matemática, cuanto a la parte educativa orientaron siempre la realización de esta tesis.

Digno de especial mención es el apoyo con su amistad y todo tipo de sugerencias de mi compañera y amiga Ma. Dolores González y de la Dra. Sonia Ursini.

En la institución en la que trabajo, el Instituto Tecnológico Autónomo de México, conté asimismo, con apoyo material y del personal en la realización de este trabajo. Mis colegas del Departamento de Matemáticas contribuyeron y me apoyaron durante todos los años que tomó la realización de este proyecto con sugerencias y críticas constructivas. En particular quisiera agradecer a quienes tuvieron una relación más profunda con este proyecto, y sin quienes hubiera sido imposible terminarlo: A Marcela González y a Javier Alfaro por su enorme ayuda en todo lo que tiene que ver con el uso de la computadora; a José Luis Morales por su asesoría y contribución en materia de Métodos Numéricos; a Guillermo Pastor por su esmerada revisión de todos los aspectos matemáticos de este trabajo; a Malú, Trini y Graciela en la labor secretarial y a Paul Hernández en la labor editorial.

Este trabajo no existiría si no hubiera contado con el apoyo incondicional de todos los miembros de mi familia. A Jaime, mi marido y a Jaime, María y Francisco, mis hijos, muchas gracias por su paciencia y por su apoyo moral durante todos estos años.

ÍNDICE

INTRODUCCIÓN

El diagnóstico y la evaluación de habilidades, actitudes y conocimiento son problemas importantes en el ámbito de las ciencias de la educación. Existen una gran diversidad de métodos de diagnóstico y de evaluación. Muchos de estos métodos se basan en técnicas de análisis cualitativo, las cuales son muy útiles en el caso del estudio de grupos pequeños de sujetos, o cuando se ha hecho mucha investigación sobre algún fenómeno y sus resultados permiten hacer generalizaciones válidas. El diagnóstico y la evaluación se han ido extendiendo día con día. Ante el número cada vez mayor de educandos y de aspirantes a escuelas y puestos de trabajo, la búsqueda de modelos matemáticos eficientes y útiles para el diagnóstico efectivo se hace cada vez más necesaria.

Medir en educación no es una tarea fácil. Las variables a medir no se pueden controlar, no se pueden medir directamente e incluso, no se pueden definir con rigor y de manera objetiva. La teoría de la medida acepta este reto e intenta definir, de una manera confiable, las variables en las que centra la atención en cada caso particular, buscando el mayor grado de objetividad posible.

La teoría de medida es un área viva y dinámica dentro de las ciencias de la educación. El estudio de estos problemas y, sobre todo de los métodos para resolverlos, se conoce como teoría de los tests. Esta teoría trata principalmente con los métodos para establecer relaciones lógico matemáticas entre las definiciones expresamente construidas en forma teórica para dar cuenta de los fenómenos de la educación y de los datos que se obtienen de instrumentos diseñados para medirlos. Métodos que, se intenta sean además, útiles para la toma de decisiones.

Los modelos matemáticos de medida en educación no son nuevos. Surgen desde principios del siglo XX en la búsqueda de parámetros que permitan clasificar de alguna manera la inteligencia o la capacidad respecto a alguna habilidad de muestras específicas de sujetos. Desde su aparición, los métodos de medida en educación han recibido mucha atención por parte de la comunidad de investigadores en educación y, gracias a los avances en las técnicas matemáticas y en las herramientas computacionales, estos métodos se han diversificado tratando de responder, al mismo tiempo, a las necesidades sociales y a los nuevos adelantos en otras ramas de la investigación educativa.

Los métodos matemáticos que se han empleado tradicionalmente en la teoría de medida no dependen de una teoría psicológica o pedagógica particular. Pueden ser igualmente útiles para la medición de muy diversos atributos y se han derivado teniendo en cuenta las necesidades específicas de la investigación en educación y en las ciencias sociales. En la actualidad, sin embargo, se ha reconocido la importancia de que estos métodos puedan relacionarse de forma más directa con los resultados de las investigaciones en las teorías cognitivas y con los resultados obtenidos de la investigación en la didáctica de las áreas específicas del conocimiento. Esta necesidad plantea nuevos retos a la teoría de los tests.

A pesar de la diversificación en los métodos de medida, los modelos que se siguen empleando con mayor frecuencia en la práctica son la Teoría Clásica de los Tests y la Teoría de Respuesta a los Ítems.

La Teoría Clásica de los Tests se caracteriza por ser un modelo simple desde el punto de vista matemático que toma como unidad de análisis al test completo. Proporciona información útil, puede ser empleado en una gran diversidad de situaciones y no impone restricciones fuertes sobre el tamaño de la muestra que se requiere emplear. Sus mayores limitaciones radican en que no proporciona información acerca del comportamiento de los examinados de distintos niveles de habilidad y en que los parámetros de los ítems dependen fuertemente del grupo de examinados con los que la teoría se usa.

La Teoría de Respuesta a los Ítems intenta superar las limitaciones de la Teoría clásica de los Tests tomando como unidad de análisis al ítem. La teoría de respuesta a los ítems se desarrolla sobre un modelo matemático que intenta describir la forma en que los examinados de diferentes niveles de habilidad responderían a un ítem y permite hacer comparaciones entre los resultados de examinados que toman diferentes tests que contienen ítems equivalentes. Los modelos de respuesta a los ítems que se emplean con mayor frecuencia en la práctica son aquéllos que suponen para la curva característica de los ítems una distribución acumulativa logística y por ello se les conoce también como modelos logísticos. La estimación de los parámetros de estos modelos es compleja y requiere de la aplicación de los tests a un número grande de sujetos. Por otra parte, la información que se obtiene de los modelos logísticos es insuficiente para determinar las razones por las cuales un ítem particular puede ser fácil para sujetos que siguen una estrategia, pero difícil para sujetos que siguen otra, e incluso para determinar si el diseño mismo del ítem favorece cierto tipo de estrategias en la respuesta.

En contraste con los modelos clásicos y con los modelos de la Teoría de respuesta a los Ítems, los modelos que permiten la introducción de estrategias alternativas en la solución de un ítem, que permiten determinar las zonas en las que el sujeto entra en conflicto al resolver cierto tipo de problemas, o que permiten acercarse a las estructuras cognoscitivas, formadas o en formación, en los sujetos, debe comenzar por proponer explícitamente los procesos mediante los cuales los sujetos alcanzan cierto tipo de respuesta. La formulación de estos modelos requiere de la su conjugación de una teoría cognitiva o educativa sólida que dé cuenta de los procesos involucrados en la solución del ítem con un modelo matemático que permita caracterizar al ítem y obtener la información deseada.

El propósito del presente trabajo consiste en desarrollar un modelo de medida basado sobre un modelo logístico de dos parámetros que permita analizar las estrategias de solución de los ítems por parte de los sujetos. Este análisis se hace, por una parte, a través de la información que provee la elección que hacen los sujetos de las alternativas de respuesta a ítems de opción múltiple y, por otra, de la conjunción del modelo con una teoría específicamente diseñada para entender la forma en la que los estudiantes construyen conocimiento matemático.

La teoría de la medida en educación es un campo que se encuentra en pleno desarrollo. En él es posible explorar diversas posibilidades de análisis matemático. Por ello se consideró como un marco interesante para un trabajo de investigación. Esta posibilidad, aunada a los retos que representa el uso de modelos dinámicos en el ámbito de los fenómenos considerados hasta ahora como pertenecientes al área de la estadística y al interés personal por el estudio de los fenómenos ligados al aprendizaje de las matemáticas, dio origen a esta investigación.

El trabajo que aquí se presenta se inscribe dentro del marco de la teoría de los tests. Por ello, en la primera parte del mismo se revisan los métodos tradicionales de medida, sus posibilidades y sus limitaciones, así como algunos de los métodos que han surgido más recientemente, haciendo énfasis en sus aportaciones y sus limitaciones.

Una gran mayoría de los modelos de medida existentes se inscriben dentro del paradigma estadístico; es decir, las técnicas de medida utilizadas en ellos provienen del ámbito de la Estadística. La Estadística ha sido considerada, hasta recientemente, la técnica matemática adecuada para estudiar fenómenos no determinísticos. Sin embargo, los desarrollos en la Matemática de los últimos años han mostrado que es difícil distinguir entre fenómenos determinísticos y no determinísticos. Los métodos del Análisis Matemático, tradicionalmente empleados en el estudio de fenómenos determinísticos han cobrado relevancia en el estudio de fenómenos que no cumplen, aparentemente, con esta característica.

En este trabajo se utilizan métodos provenientes del análisis Matemático, en particular de las Ecuaciones Diferenciales, para desarrollar el modelo de medida. Se intenta mostrar cómo un cambio de paradigma y la consideración del problema de respuesta de ítems o de solución de tests como un problema dinámico, en el que las probabilidades de respuesta varían dependiendo de las habilidades de quienes responden y también de las situaciones de contexto en las que se responden los ítems de un test, puede aportar información diferente y nuevos resultados

En la construcción de un modelo de esta naturaleza se pueden tomar en consideración las diferentes estrategias de solución al ítem, los errores comunes que cometen los sujetos en la solución de un ítem dado, o alguna teoría acerca de la estructuración del conocimiento para diseñar los ítems. La inclusión de este tipo de consideraciones en el diseño de los tests y en los modelos de medida proporciona información útil para el diagnóstico, para la evaluación, para la clasificación y para el diseño de alternativas de respuesta adecuados para cada ítem. Las posibilidades de diseño de tests con base en este tipo de modelos son de enorme interés en el campo de la educación. A partir de las respuestas a los ítems así diseñados es posible obtener información acerca de la forma en la que diferentes sujetos resuelven un ítem particular y que resulta más valiosa que el simple conocimiento del número de ítems que los sujetos pueden resolver correctamente.

La formulación de un modelo de esta naturaleza requiere que se hagan explícitos los procesos por los cuales los sujetos eligen diferentes respuestas. Por ello, además del modelo matemático, en este trabajo se presenta un modelo educativo que sustenta las posibilidades de elección entre las alternativas de respuesta al ítem.

En el modelo que se presenta en este trabajo, la selección de las distintas opciones de respuesta al ítem está ligada al tipo de razonamiento que siguen los sujetos al resolverlo y a un modelo de desarrollo cognitivo sustentado en las ideas de Piaget. El modelo, por otra parte, toma en consideración el hecho de que cuando el sujeto enfrenta un ítem de opción múltiple, la elección de una entre varias alternativas de respuesta se ve afectada por la presencia de las otras alternativas, lo cual permite caracterizar matemáticamente dicha elección.

En el desarrollo del modelo que se presenta en este trabajo se parte de la posibilidad de interpretar los modelos logísticos desde una perspectiva dinámica; es decir, tomando en consideración la variación en la probabilidad de responder correctamente al ítem con la variación del rasgo a medir. Se hace, en primer término, una generalización basada en la Teoría de Respuesta a los ítems; posteriormente, mediante la introducción de las probabilidades de elección de las diferentes alternativas de respuesta y mediante la consideración de parámetros que caracterizan a cada una de dichas alternativas y de parámetros que introducen en el modelo la posibilidad de que el sujeto que responde dude al presentársele estas alternativas se elabora un mocelo matemático de medida.

Por ello se mostrará, en primer lugar, cómo pueden derivarse los modelos logísticos tradicionales a partir de una perspectiva de naturaleza dinámica. Posteriormente se hará una propuesta de generalización mediante la introducción de nuevos parámetros que pueden ser relacionados con la teoría cognitiva que permitirá el diagnóstico.

El análisis de los resultados que se obtienen de la solución del modelo propuesto muestra que es posible obtener una gran variedad de formas para las curvas que describen la probabilidad de elección de cada una de las opciones de respuesta de cada ítem, sin necesidad de introducir modelos distintos para cada una de ellas. Este hecho hace al modelo que se presenta en este trabajo más versátil que otros que se encuentran en la literatura. Con una misma formulación matemática pueden hacerse aplicaciones a tests diseñados para medir conocimientos que a tests diseñados para medir estrategias o preferencias. Además, los parámetros del modelo se pueden interpretar con facilidad en términos de la información que proporcionan en el marco educativo. Otra característica, que constituye una aportación fundamental del modelo de medida que se presenta en este trabajo, consiste en que éste proporciona datos que permiten evaluar la efectividad de los distractores de cada ítem. No se encontró ningún en la literatura ninguna referencia que tuviera esta propiedad.

El modelo matemático diseñado en este trabajo y su posibilidad de interpretación en términos educativos se ponen a prueba a través de su aplicación al análisis de los resultados de un cuestionario diagnóstico que intenta medir la comprensión de los estudiantes de un concepto fundamental del álgebra elemental: el concepto de variable.

Una vez estimados los parámetros para cada ítem, es posible identificar intervalos de la variable que representa el rasgo a medir que podrían considerarse zonas de duda o confusión. Así, a diferencia de otros modelos que únicamente dividen a los sujetos en dos conjuntos: aquéllos que responden correctamente el ítem y aquéllos que lo responden incorrectamente, éste permite identificar un mayor número de intervalos que proporcionan información valiosa para el diagnóstico y para la toma de decisiones. Esto se muestra claramente en la aplicación del modelo al cuestionario diagnóstico.

En cuanto a su organización, el trabajo se divide en tres capítulos. En el primero se hace una revisión de algunos modelos de la teoría de los tests. En el segundo se presenta la aportación que se propone y en la tercera ésta se aplica a un cuestionario diagnóstico.

La primera parte de este trabajo introduce las ideas más importantes relativas al significado de la medida en educación. Se presentan en ella las hipótesis y los resultados de las dos teorías de los tests más frecuentemente utilizadas, que se mencionaron con anterioridad, la teoría clásica de los tests y la teoría de respuesta a los ítems, así como sus principales limitaciones. Esta parte concluye con una revisión de los resultados de la investigación en dos áreas relevantes en el contexto de este trabajo: el formato de los ítems de los tests y el diseño de algunos modelos de medida, que proveen un marco de referencia adecuado dentro del cual se puede ubicar el modelo que aquí se presenta.

En la segunda parte del trabajo, y con el fin de introducir el significado del cambio de paradigma desde el punto de vista matemático, se hace una deducción de los modelos logísticos a partir de una formulación dinámica del problema, en términos de ecuaciones diferenciales. Se muestra con esta deducción que ambos puntos de vista, el estadístico y el dinámico, conducen, en este caso particular a modelos equivalentes.

A continuación se introduce el modelo matemático que constituye la contribución central de este trabajo. Se desarrolla el modelo en forma general y se interpretan sus parámetros en términos matemáticos y desde el punto de vista de su significado en términos educativos. Se analizan, posteriormente, sus consecuencias en términos matemáticos para el caso de ítems que contienen únicamente dos o tres opciones de respuesta. Se discuten con detalle sus posibilidades en términos de las posibles curvas que representan a las probabilidades de elección de cada una de las opciones de respuesta que se pueden obtener y la forma en la que el análisis puede llevarse a cabo. Se discute asimismo, detalladamente, la forma en la que las curvas antes mencionadas se pueden analizar en el contexto educativo.

La segunda parte concluye con la presentación de una posible forma de estimar los parámetros del modelo. Dada su complejidad y su naturaleza no lineal, la estimación de los parámetros requiere de la introducción de métodos numéricos que requieren del uso de conceptos matemáticos avanzados. Esto hace necesario que, en esta parte, el lenguaje matemático empleado sea mucho más formal que en el resto del trabajo.

Con el fin de poner a prueba el modelo, de mostrar la posibilidad de utilizarlo conjuntamente con una teoría cognitiva y de verificar su utilidad en el marco del diagnóstico, en la tercera parte del presente trabajo se introduce una discusión acerca del papel del diagnóstico en la educación y se presenta un modelo, basado en las ideas de Piaget, que es adecuado para el diagnóstico del aprendizaje de conceptos matemáticos. Enseguida se discute la importancia del concepto de variable en la enseñanza del álgebra elemental y se revisan algunas de las investigaciones que se han llevado a cabo en el ámbito de la didáctica de las matemáticas sobre este concepto, poniendo énfasis en los resultados obtenidos en ellas.

El modelo de diagnóstico para la enseñanza de las matemáticas se aplica, posteriormente, al concepto de variable en el álgebra elemental. Esta aplicación permite el diseño de un cuestionario cuya finalidad es adentrarse en la forma en la que los estudiantes comprenden el concepto de variable. De entre los ítems de ese cuestionario se eligen treinta para diseñar un cuestionario de opción múltiple en el que las opciones de respuesta están sustentadas por el modelo educativo y por los resultados de la investigación en didáctica de las matemáticas. Se presenta, en esta parte del trabajo un análisis detallado de cada uno de los ítems desde la perspectiva del modelo educativo.

A partir de los datos obtenidos de la aplicación del cuestionario, se estiman los parámetros del modelo para cada uno de los ítems. Esta parte del trabajo termina con el análisis cuidadoso de los resultados obtenidos de la aplicación del modelo, tanto en términos matemáticos cuanto en términos del modelo de diagnóstico del concepto de variable. En esta parte se muestra claramente la viabilidad del modelo, se discuten algunas de sus ventajas y sus limitaciones.

El trabajo que aquí se presenta termina con una discusión de las posibilidades de aplicación del modelo de medida diseñado, de sus limitaciones y de las aportaciones en el ámbito de la teoría de los tests de un modelo de esta naturaleza. Se señalan, por último, algunas preguntas de investigación que surgen del presente trabajo. ◆

CAPÍTULO I

LA MEDIDA EN EDUCACIÓN

Medir, en cualquier ciencia, está relacionado con la asociación de números o cantidades a alguna propiedad o atributo de un objeto en estudio. En el caso de las ciencias naturales las cantidades que se desean medir son atributos físicos de la materia a los que generalmente se puede acceder directamente a través de instrumentos diseñados con ese fin. El error en la medida depende de la precisión con la que dichos instrumentos accedan a las propiedades que se desean medir y, mediante el avance en la tecnología y en el conocimiento de las propiedades, se logra mejorar paulatinamente la precisión.

En este capítulo se pretende presentar una breve introducción a los desarrollos en la teoría de la medida en educación.

SECCIÓN I.1 PROPÓSITO DE LA MEDIDA EN EDUCACIÓN

La definición más común de medida en psicología es la propuesta por Stevens en su obra clásica sobre matemáticas y teoría de la medida (1946), según la cual, medición en su sentido más amplio es "la asignación de números a objetos de acuerdo con ciertas reglas". Torgerson (1958) y Lord y Novick (1968), dieron una definición más precisa de la medición señalando que ésta se aplica a las propiedades o características de los objetos más que a éstos en sí mismos. Cuando las características que se desean medir son "intangibles" como es el caso de la inteligencia o el conocimiento y otras características asociadas con los objetos de estudio de las ciencias sociales, las medidas son descripciones de la conducta o comportamiento de los sujetos que, se supone reflejan las características a medir. Por tanto, medir en educación significa asignar un valor cuantitativo a una evidencia de conducta observable que, se supone refleja la característica a medir de un individuo, a través de cualquier tipo de instrumento, al que suele llamarse test.

En el caso de las ciencias sociales, en particular de la psicología y de las ciencias de la educación, las propiedades que se desean medir nunca pueden ser directamente confirmadas; sin embargo, y a pesar de que esas características cambian constantemente, existe algún grado de consistencia en las conductas que permite definir constructos teóricos que son el reflejo observable de lo que se desea medir y diseñar instrumentos para medir dichos constructos y hacer inferencias acerca del grado en que caracterizan a un individuo a partir de los datos que se obtienen.

Puesto que es imposible acceder directamente a los atributos psicológicos de las personas, es necesario definir los atributos a través de otras propiedades del comportamiento que se relacionan con los atributos que se desean medir, pero que son más directamente accesibles a la observación y a la medición. De esta manera es posible utilizar estos atributos como indicadores con los cuales es posible diseñar instrumentos destinados a dar cuenta de manera indirecta de los atributos de las personas.

Así, en el intento de explicar el comportamiento humano, en todas las ciencias sociales se han definido conceptos que intentan dar cuenta de él. La existencia de esos constructos no se puede confirmar directamente, por lo que el grado en el que caracterizan el comportamiento de un individuo, o de un conjunto de individuos, sólo se puede inferir de la observación de ciertos rasgos observables del comportamiento y del establecimiento de relaciones entre los observables y el constructo teórico.

La medida de un constructo requiere del establecimiento de una regla de correspondencia entre él y los comportamientos observables que se consideran como los indicadores que dan cuenta legítimamente del constructo en cuestión. Esta regla constituye la definición operacional del constructo. Una vez definido operacionalmente el constructo, se procede a generar la forma en la que los comportamientos seleccionados como indicadores pueden observase y registrarse. Esta fase incluye normalmente el desarrollo de un instrumento de observación, que, en el caso de la educación, queda constituido en muchas ocasiones por un examen o test. En el ámbito de la educación la palabra test se usa de una manera muy general en la que el término test no se asocia exclusivamente a un cuestionario o a un examen sino que se refiere a cualquier tipo de procedimiento estándar para obtener una muestra del comportamiento óptimo, o del comportamiento típico, según sea el caso, de un sujeto, o de un conjunto de sujetos, en un dominio específico. La palabra test puede ser empleada tanto para un cuestionario cuanto para un guión de observación o para una entrevista.

La medición del atributo en cuestión se lleva a cabo cuando se le asigna un valor cuantitativo a la muestra obtenida mediante el test. A partir de esa medida, es posible establecer, mediante inferencias, algunas propiedades del constructo definido, así como acerca de las relaciones entre constructos diferentes, de forma tal que sea posible construir una teoría educativa o psicológica.

El diseño de los instrumentos, en el caso de este tipo de ciencias, juega un papel primordial. Por una parte se debe justificar de una manera aceptable que el instrumento mide aquello para lo que se diseñó y por otra parte, debe aclararse también el significado de la precisión en este tipo de medida.

Es claro que el diseño de instrumentos que permitan el registro de los observables asociados al constructo presenta grandes dificultades. Entre ellas se pueden citar las siguientes:

- No existe ningún procedimiento de medición de algún constructo que sea aceptado universalmente, diferentes procedimientos de medida pueden conducir a conclusiones diferentes acerca del nivel del constructo en el individuo.

- Las medidas psicológicas se basan generalmente en muestras limitadas de comportamientos. La determinación del número de cuestiones que se van a observar y la variedad de contenido necesario para obtener una muestra adecuada del dominio de comportamiento es un problema fuerte en el desarrollo de un procedimiento sólido de medida.

- La medida obtenida está siempre sujeta a error. Un problema importante que se debe resolver es la forma de estimar el error en la medición de un conjunto dado de observaciones.

- La definición de las escalas de medida es sumamente compleja. La definición de las propiedades de la escala de medida, la elección de las unidades y la interpretación de los valores derivados son asuntos que deben considerarse en el diseño de cualquier instrumento de medida y del procedimiento de calificación. No hay un acuerdo universal respecto a esta cuestión.

- Los constructos psicológicos no se pueden definir únicamente mediante definiciones operacionales. Deben tener además relación con otros constructos y fenómenos observables de manera que el proceso de medida sea consistente.

Una medida psicológica, aun basada en observaciones, puede tener poco sentido y utilidad, a menos que se pueda interpretar a la luz de los constructos teóricos subyacentes. Lord y Novick (1968) proponen, por esta razón, que además de la definición operatoria del constructo, se cuente con una definición de su relación lógica o matemática con otros constructos dentro del sistema teórico. La definición de las relaciones lógico matemáticas entre los constructos de un sistema permite sentar las bases para la interpretación de las medidas obtenidas. Sin embargo, la obtención de evidencia de la forma en la que un conjunto de medidas psicológicas se relaciona con las medidas de otros constructos o con otros eventos en la realidad, constituye el mayor reto del desarrollo de los tests.

El estudio de estos problemas y, sobre todo de los métodos para resolverlos, se conoce como teoría de los tests. Esta teoría trata principalmente con los métodos para establecer relaciones lógico matemáticas entre los constructos teóricos y los observables relacionados con ellos, con los métodos para estimar hasta qué grado los problemas antes mencionados influyen sobre las mediciones tomadas en una situación particular y con el diseño de métodos para resolver o minimizar estos problemas.

La teoría de los tests constituye un marco general para el proceso de desarrollo de instrumentos. Los métodos matemáticos que se han empleado tradicionalmente no dependen de una teoría psicológica o pedagógica particular. Pueden ser igualmente útiles para la medición de muy diversos atributos y se han derivado teniendo en cuenta las necesidades específicas de la investigación en educación y en las ciencias sociales.

Dentro de la teoría de los tests caben diversas formas de resolver los problemas que se han mencionado anteriormente. Cada acercamiento parte de hipótesis diferentes, aporta elementos nuevos de discusión y tiene asimismo limitaciones diferentes. La teoría de los tests es un área joven de conocimiento que se encuentra en pleno desarrollo. Además de los problemas teóricos antes mencionados, los problemas operativos para poner en práctica los métodos construidos en el marco de esta disciplina, son también enormes. Los métodos correlacionales y el análisis factorial son las técnicas estadísticas más comunes para estudiar y validar los tests; sin embargo, gracias a los avances tecnológicos en el ámbito de la computación y a los avances en las diversas ramas de las matemáticas, hoy se cuenta con una vasto instrumental que permite la exploración para construir nuevos métodos de estudio y de validación de los tests. Los problemas prácticos que se presentan todavía son grandes, pero las posibilidades teóricas son enormes. La teoría de los tests puede encontrar en a ellas un campo fructífero dentro del cual se abren nuevas posibilidades de desarrollo.

Dado que el propósito del presente trabajo se inscribe dentro del marco de la teoría de los tests es conveniente revisar los métodos tradicionales de medida, sus posibilidades y sus limitaciones, así como los métodos que han surgido recientemente haciendo énfasis en sus aportaciones y las dificultades de su aplicación.

La teoría clásica de los tests ha sido la teoría dominante desde su creación hasta la actualidad. Dada su importancia y su vigencia es importante revisarla someramente. A esto se dedicará la primera parte de este capítulo.

La búsqueda de métodos alternativos que permitan subsanar las limitaciones de la teoría clásica es la que ha dado origen a los nuevos desarrollos. Dentro de ellos destacan los modelos de respuesta a los ítems que en la actualidad han adquirido un estatus paralelo al de la teoría clásica. Por su importancia y puesto que este trabajo pretende desarrollar un modelo de medida que se sustenta en los modelos logísticos de respuesta a los ítems se presenta también una revisión de estos modelos, de sus posibilidades de aplicación y de sus limitaciones.

Los últimos años se han caracterizado por el nacimiento de nuevos modelos de medida. Algunos de ellos basados en los modelos logísticos y otros que parten de diversas ramas de las matemáticas y aprovechan los métodos del análisis factorial para explorar territorios nuevos. La última parte de este capítulo se dedicará a la revisión de los métodos que están impactando más fuertemente el desarrollo de la teoría de los tests en la actualidad. Dicha revisión permitirá además contar con un marco de referencia dentro del cual puede situarse el modelo que se presenta en este trabajo.

Sección I.2 Teoría Clásica de los tests

En esta parte del trabajo se desarrollarán brevemente los postulados y los principales resultados de la teoría clásica de los tests así como sus limitaciones.

La teoría clásica de los tests nace de los trabajos de Spearman (1904,1907,1913) y se caracteriza por la sencillez de los modelos matemáticos empleados y por la versatilidad que presenta su aplicación, lo que le ha garantizado una larga vida. El modelo lineal clásico sigue siendo la base de la mayoría de los instrumentos de medición utilizados actualmente en las ciencias sociales; es por ello que la comprensión de los conceptos y técnicas de la teoría clásica resultan básicos para poder entender los avances en la teoría de la medida y los modelos que recientemente se han desarrollado con el fin de mejorar las técnicas de medida que se utilizan en la evaluación psicológica y educativa.

El modelo lineal clásico propuesto por Spearman supone que la evaluación, o puntuación empírica que un sujeto *j* obtiene mediante la aplicación de un test *i* consta de dos componentes aditivos: uno que es la verdadera puntuación del sujeto en el test *V* constante para cada sujeto, pero desconocida, y otro, el error *e* que inevitablemente va asociado a todo proceso de medición y que afecta la puntuación del sujeto *j* en un test *y*; este error, en este caso en particular, es debido al instrumento de medición. Es decir, de acuerdo a este modelo:

$$X = V + e$$

A partir de este modelo y algunos supuestos básicos, la teoría desarrolla todo un conjunto de deducciones encaminadas a estimar la cantidad de error que afecta a las puntuaciones en los tests. La lógica general para la estimación de los errores es común a todas las ciencias, pero la naturaleza no observable directamente de lo psicológico añade algunas complicaciones adicionales.

La puntuación empírica que obtiene un sujeto cuando se le aplica un test, en un momento dado no coincide exactamente con su verdadera puntuación, pues en ese momento el sujeto está afectado por factores de difícil control que inciden en su conducta. Si estos factores perjudican al sujeto, obtendrá una calificación empírica más baja de la que verdaderamente le correspondería; si le benefician, obtendrá una superior. Cuando se aplica un test a un sujeto, lo único que se puede obtener es su puntuación empírica, es decir, los puntos que obtiene en el test; su puntuación verdadera se estima basándose en los supuestos del modelo.

Siempre que a un sujeto se le aplica un test, su puntuación en él puede considerarse como una variable aleatoria. Por ejemplo, si el test tiene veinte ítems, la puntuación de un sujeto puede tener un valor entre 0 y 20, según una densidad de probabilidades. Antes de aplicar un test a un sujeto, no se puede saber qué factores le beneficiarán o le perjudicarán al momento de ser examinado; por ello, antes de aplicar un test a un sujeto puede verse la puntuación del sujeto en el test como una posibilidad de asumir uno de varios valores de acuerdo con la densidad de probabilidades desconocida. Esta distribución de las puntuaciones que puede obtener un sujeto se puede considerar como una variable aleatoria y la puntuación que se observa una vez que se le aplica el test es una realización de esa variable aleatoria.

Es importante señalar que la puntuación de cada uno de los sujetos a los que se les aplica un test representa una variable aleatoria diferente; es decir, la probabilidad de obtener una cierta puntuación se determina en forma independiente con una distribución diferente para cada sujeto.

Antes de exponer los supuestos básicos del modelo clásico es importante definir la notación para las variables de interés:

Sea

V la puntuación verdadera del sujeto, es decir, de cualquier sujeto de una muestra de sujetos, por tanto, una variable aleatoria,

V_j la puntuación verdadera de un sujeto, que es un parámetro no observable,

X la puntuación empírica de un sujeto en un test, que es una variable aleatoria,

X_j la puntuación empírica de un sujeto j en un test y que es también una variable aleatoria,

j_e el error que afecta la puntuación de un sujeto en un test, que es una variable aleatoria,

e_j el error que afecta la puntuación de un sujeto j en un test y que es también una variable aleatoria.

Supuesto 1.-

- La puntuación verdadera de un sujeto *j* es la esperanza matemática de la empírica, es decir, $V = E(x)$. Esto significa que si un test se aplica un número infinito de veces al mismo sujeto, suponiendo que cada aplicación no afecta a las otras y que el sujeto no cambia en el curso de las aplicaciones, la puntuación verdadera del sujeto en el test sería la media de las puntuaciones empíricas obtenidas en las aplicaciones. Este supuesto es en realidad una definición de la puntuación verdadera. A partir de los valores de las puntuaciones empíricas y bajo otros supuestos que se verán más adelante, la teoría clásica de los test permite hacer estimaciones probabilísticas razonables acerca del valor de la puntuación verdadera.

A partir de este primer supuesto pueden deducirse los siguientes resultados:

- La esperanza matemática de los errores de medida que afectan la puntuación de un sujeto *j* es cero, de otra forma: $E(X_j) = 0$.
- Suponiendo que se examina a un grupo de sujetos, a cada uno de los examinados se les asocia una seria de posibles puntuaciones empíricas en este instrumento. La media de las puntuaciones verdaderas es igual a la media de las puntuaciones empíricas.
- La media de los errores que afectan a las puntuaciones de un grupo de examinados es cero. Es importante señalar que este resultado no es garantía de que, en cada ocasión que se aplica un test, el promedio de los errores que afectan las puntuaciones de los examinados en ese test sea cero. Es el valor colectivo de la media de la población de los errores para todos los examinados la que tiene el valor esperado de cero. Aplicar un

test i a un grupo de n sujetos es equivalente a seleccionar una muestra de valores de entre los errores, al obtener aleatoriamente un valor de los errores para cada examinado. La media de esta muestra puede ser o no ser cero.

Supuesto 2.-

- No existe correlación o asociación lineal entre las puntuaciones verdaderas de los sujetos en un test y sus respectivos errores de medida. Esto puede expresarse como: $\rho(V, e) = 0$, en la que $\rho(V, e)$ denota la correlación lineal entre las puntuaciones verdaderas de una población de sujetos

A partir de este supuesto pueden deducirse los siguientes resultados:

- Las puntuaciones verdaderas no covarían con los errores: $Cov(V, e) = 0$.
- La covarianza entre las puntuaciones empíricas y las verdaderas es igual a la varianza de las verdaderas: $Cov(V, e) = Var(V)$.
- La varianza de las puntuaciones empíricas es igual a la varianza de las puntuaciones verdaderas más la de los errores: $Var(X) = Var(V) + Var(e)$.
- La correlación entre las puntuaciones empíricas y los errores es igual al cociente de la desviación típica de los errores entre la de las puntuaciones empíricas: $\rho(X, e) = \sigma_e / \sigma_X$

Supuesto 3.-

- Los errores de medida de los sujetos en un test no están correlacionados con sus errores de medida en otro test distinto: $\rho(e_i, e_j) = 0$

A partir de este primer supuesto pueden deducirse los siguientes resultados:

- La covarianza entre las puntuaciones empíricas de dos tests *j* e *i*, (X_i, X_j), es igual a la covarianza entre las verdaderas (V_i, V_j).

Es importante aclarar que ninguno de los supuestos del modelo es comprobable empíricamente de un modo directo. El modelo y las deducciones que se han presentado están formulados para los valores paramétricos de la población.

Si bien las hipótesis en las que la teoría clásica de los tests no se pueden comprobar experimentalmente, a partir de ellas se pueden hacer inferencias que si se pueden contrastar experimentalmente.

Estadísticos de la teoría clásica de los test

En términos generales, la teoría clásica supone que los estadísticos obtenidos en muestras suficientemente amplias constituyen estimadores apropiados de los valores de la población. Cuanto más amplias sean las muestras más pertinente será esta lógica. La teoría clásica de los tests supone que para muestras muy grandes, las puntuaciones obtenidas en un test tienen una distribución normal y las posiciones de los sujetos, respecto de la característica que se está midiendo, se pueden expresar en una escala de intervalo.

El supuesto de la distribución normal de las puntuaciones obtenidas en un test cuando la muestra es grande se basa en el Teorema Central del Límite que dice que si una población tiene una varianza σ y una media μ, entonces, a medida que el tamaño de la muestra, *n* aumenta, la distribución de la media de la muestra tiende a la distribución normal con varianza σ y media μ.

Dado que una colección de datos procede de una muestra de una población, se debe hacer una distinción entre los resultados derivados de la muestra y los que se derivan de la población. La razón más importante para ello es que el análisis que se realiza para una muestra, en ocasiones se hace con el propósito de generalizarlo a la población.

A continuación se presentan algunas operaciones formuladas para los valores paramétricos de la población que se aplican a los instrumentos de medida utilizando estadísticos como estimadores de la población. Los estimadores que se utilizan para la media y la varianza de las puntuaciones de la población son los siguientes:

El estimador máximo verosímil de la media de las puntuaciones de la población está dado por

$\mu = \sum_{i}^{n} X_i / n = \overline{X}$, donde $\overline{X}$ es la media muestral, X_i es la i-ésima puntuación observada y n es el tamaño de la muestra.

El estimador máximo verosímil de la varianza de las puntuaciones está dado por $\sigma^2 = s^2 = (X_i - X)^2 / (n-1) = \overline{X}$, donde s^2 es la varianza muestral, X_i la i-ésima puntuación observada, $\overline{X}$X la media de la muestra y n el tamaño de la muestra. El estimador de la varianza de la población aquí expuesto es un estimador insesgado y es el de mayor aplicación en la práctica para estimar σ^2.

Siempre que se utiliza un instrumento de medida para evaluar las características de algún objeto o individuo, se busca la seguridad de que se obtendrán los mismos resultados cada vez que se mida la misma característica en condiciones similares. A esta consistencia o reproductibilidad de los resultados se le denomina fiabilidad. Por esta razón, uno de los aspectos más importantes que requiere definir una teoría de la naturaleza de la teoría de los tests es la noción de fiabilidad.

Cualquier instrumento de medida, en este caso un test, se considera fiable si las medidas en que se utiliza carecen de errores de medida; es decir, si son consistentes. En términos prácticos, la fiabilidad representa el grado con el que las desviaciones de los individuos permanecen consistentes sobre la aplicación repetida del mismo test o de formas alternativas del test. Un test es fiable si cada vez que se aplica a los mismos sujetos se obtiene el mismo resultado. Esta medida debe describir el grado en el que los errores aleatorios influyen en los resultados de la aplicación del instrumento.

Los errores de medida de los que se ocupa la fiabilidad son aquellos que no se pueden controlar, que son inevitables en todo proceso de medición ya sea, por ejemplo, físico, químico o psicológico. En muchas ocasiones la diferencia entre una medición y otra no depende sólo de estos errores, sino que puede además explicarse por errores tales como los cambios operados en los sujetos, la imprecisión del instrumento o los errores al calificar un test. Es importante señalar que no es necesario, ni recomendable, tal vez, hacer mediciones repetitivas utilizando el mismo test para estimar su fiabilidad. Es más conveniente hacerlo a través de instrumentos de medida equivalentes llamados test paralelos. Dada la dificultad de construir formas paralelas de un test y de aplicarlas a los sujetos, suele emplearse otros métodos para obtener el coeficiente de fiabilidad, entre ellos destacan el método test-retest, el método de construir dos formas distintas del test y el método de las dos mitades.

Dos formas o test i y j se denominan paralelos si, para ambos, la media y la varianza de las puntuaciones empíricas, la varianza de los errores y las puntuaciones verdaderas de los sujetos son las mismas. Esta es una definición teórica que supone implícitamente que se pueden construir de hecho estos tests. Pero, en la práctica, es imposible que estas condiciones se cumplan. De un modo menos formal que el señalado en la definición de paralelismo de tests, se dice que dos tests se consideran paralelos si miden lo mismo pero con diferentes ítems.

Con el propósito de llegar a formas prácticas para medir la fiabilidad de un instrumento, se toma el resultado de un individuo en un test como una variable aleatoria, considerando que el resultado del individuo puede tomar un valor entre un conjunto de posibilidades de acuerdo a un conjunto desconocido de probabilidades. La distribución de los resultados potenciales para un examinado en particular se puede considerar como una variable aleatoria y el resultado obtenido se puede pensar como una realización de esa variable; por ello se define la calificación verdadera como la esperanza matemática de los resultados obtenidos mediante la aplicación del test.

En la teoría clásica de los tests se definen maneras diferentes de medir la fiabilidad de un test, cada una de ellas implica una problemática diferente. El índice de fiabilidad $\rho_{XX'}$ de un test se define como la correlación entre las puntuaciones obtenidas por los mismos sujetos en dos formas paralelas de un test. Entre los diversos métodos para estimar la fiabilidad de un test a través de formas paralelas el más empleado es el método de las "dos mitades". Este método es muy funcional ya que sólo exige una aplicación del test. El método consiste en dividir el test en dos mitades. El coeficiente de fiabilidad está dado por la correlación lineal entre las puntuaciones correspondientes a cada una de esas dos mitades e indica la correlación o consistencia interna entre ellas, más una correlación para obtener la fiabilidad del test total. La estimación así obtenida constituye un indicador de la consistencia interna del test.

El empleo del método de las dos mitades plantea diversos problemas experimentales relacionados con la validez interna. El modelo clásico no da especificaciones concretas respecto de este problema, quedando al criterio del investigador lo que debe hacerse en cada situación nueva. Dichos problemas se refieren a que no se obtiene un único estimador del coeficiente de fiabilidad del test, ya que existen varias maneras posibles de dividir un test en dos mitades. Un test con *k* ítems tiene como posibles mitades el número de combinaciones de estos *k* ítems tomados de *k/2* en *k/2*, pero hay que garantizar que las mitades de test sean paralelas. No es recomendable, por ejemplo, considerar como mitades la primera parte del test por un lado y la segunda por otro pues los sujetos llegarán más cansados a la segunda, por lo que esta parte resultará más difícil que la primera. Para evitar ésto es frecuente tomar como una mitad los ítems pares y otro los impares, o usar algún otro tipo de apareamiento de los ítems. Este es, en definitiva, un problema de control experimental.

Al dividir un test en dos formas paralelas, a y b, la puntuación total, *X*, para cada examinado está basada en una puntuación compuesta: $X = Xa + Xb$, donde X_a es la puntuación obtenida en la forma *a* y X_b la puntuación obtenida en la forma *b*.

El índice de fiabilidad es un indicador de la estabilidad de las medidas hechas, ya que, al aplicar las formas paralelas del test a una misma muestra de sujetos en dos tiempos distintos, debe esperarse una correlación perfecta, es decir, $\rho_{XX'} = 1$, puesto que miden lo mismo. El grado en que el coeficiente de fiabilidad se aleje de 1 proporciona una medida de qué tanto las medidas obtenidas mediante el instrumento se ven afectadas por los errores aleatorios de medición.

¿Cómo se puede calcular este coeficiente empíricamente para una muestra determinada? En la práctica el coeficiente de fiabilidad se puede definir como la correlación entre las calificaciones obtenidas en dos formas paralelas del test y resulta ser igual al cuadrado del índice de fiabilidad. Para obtener estos coeficientes se deben, entonces, construir dos formas paralelas del test, aplicarlas a una muestra amplia de sujetos representativos de la población en la que se quiere hacer la medida y calcular la correlación entre las puntuaciones obtenidas en las dos formas del test.

En el modelo clásico, se utilizan con frecuencia dos maneras de expresar la fiabilidad de un test en términos de la propiedades estadísticas de sus componentes:

El primer método, que utiliza la fórmula de Spearman – Brown, permite estimar la fiabilidad de un test compuesto por formas paralelas cuando la fiabilidad de una de las formas es conocida. El segundo método utiliza un estadístico llamado alfa de Cronbach y permite estimar el coeficiente de fiabilidad de un test compuesto cuando se conoce la varianza de su puntuación y las covarianzas de sus componentes.

Si se tienen *k* tests o formas paralelas de un test, la puntuación total está dada por la suma de las puntuaciones en cada una de las partes que componen el test global. La puntuación verdadera de un sujeto está compuesta por su puntuación verdadera en cada una de las formas paralelas del test. Se define el coeficiente de fiabilidad para cualquier pareja de formas paralelas de un test como $\rho_{ii'} = \sigma^2{}_{\mathrm{Vi}} / \sigma_i^2$ en la que $\sigma^2{}_{\mathrm{Vi}}$ es la varianza de las puntuaciones verdaderas en el test *i*, $\sigma^2{}_i$ es la varianza de las puntuaciones obtenidas en el test i y $\rho_{ii'}$ es la correlación lineal entre las puntuaciones observadas en dos formas paralelas i e i' del test. A partir de esta definición de los resultados derivados de los supuestos del modelo, puede encontrarse una manera de calcular empíricamente el coeficiente de fiabilidad, ya que de la definición es imposible hacerlo, dado que el valor de $\sigma^2{}_V$ no se puede obtener de las respuestas de los sujetos a los ítems; no obstante, es útil para dar una idea conceptual de lo que representa el coeficiente de fiabilidad. Puesto que $\rho_{ii'}$ indica la proporción de la varianza verdadera respecto a la empírica, si no hubiese errores de medida, el coeficiente de fiabilidad, como ya se dijo anteriormente debiera ser igual a 1. A este coeficiente se le denomina también coeficiente de equivalencia puesto que indica el grado en el que ambas formas son equivalentes. En la práctica este coeficiente se calcula a partir de la fórmula de correlación lineal de Pearson, que es la que se utiliza para derivar la fórmula para el coeficiente de fiabilidad. La forma general de Spearman – Brown está dada por la siguiente ecuación:

$$\rho_{XX'} = k\rho_{ii'} / (1 + (k-1)\rho_{ii'})$$

Esta ecuación se denomina frecuentemente como la profecía de Spearman – Brown porque a partir de ella se puede observar que dado que k es el número de formas, tests o ítems, paralelos si se aumenta la longitud del tests, entendida como el número total de formas del test, la fiabilidad del nuevo test alargado viene dada por esta fórmula, donde $\rho_{ii'}$ sería la fiabilidad del test original. Esta fórmula funciona siempre y cuando las formas añadidas sean paralelas a las previamente existentes. Es interesante notar que al aumentar k veces la longitud de un test, su varianza verdadera aumenta proporcionalmente más que su varianza empírica pues, mientras la verdadera original resulta multiplicada por k la empírica se multiplica por $k(1+(k-1)\rho_{ii'})$, expresión cuyo valor sólo se igualaría k en el caso en que $\rho_{XX'} = 1$, hecho poco probable en la práctica. Es decir, la fiabilidad de un test aumenta a medida que se incrementa su longitud, ya que el numerador crece más rápidamente que el denominador, alcanzando teóricamente el valor 1 para un test compuesto por un número infinito de formas. Mejorar la fiabilidad de un test a base de aumentar su longitud puede ser útil y recomendable, por ejemplo, cuando el test original ya tiene una fiabilidad razonable con no muchas formas, pero si el test ya tiene un número considerable de formas y, sin embargo, es poco fiable, más que en su longitud hay que pensar en construir otro.

Se suele denominar índice de fiabilidad ρ_{XV} a la correlación lineal entre las puntuaciones empíricas de un test, *X*, y las verdaderas, *V*, y puede encontrarse como la raíz cuadrada del coeficiente de fiabilidad. Puede demostrarse que este coeficiente está dado por:

$$\rho_{XX'} = k^2 \ Cov(i,l) / \sigma^2{}_X$$

cuando todos los tests son paralelos

Como en la práctica no se puede estar completamente seguro de que todos los tests o las formas del test que intervienen en la puntuación total o compuesta, sean estrictamente paralelas, es posible utilizar la suma de las covarianzas del test y la varianza compuesta para estimar el límite inferior de la fiabilidad del test compuesto. Este coeficiente debe ser menor al coeficiente de fiabilidad.

Cuando los *k* subtest de un test compuesto no son estrictamente paralelos, por lo menos para uno de ellos, la varianza de su puntuación verdadera es mayor o igual a su covarianza con cualquier otro subtest. Además, es posible demostrar que para cualquier par de test que no sean estrictamente paralelos, la suma de las varianzas de sus puntuaciones verdaderas es mayor o igual que el doble de su covarianza y, también que, la suma de *k* varianzas de la puntuación verdadera de tests que no sean paralelos será mayor o igual a la suma de sus *k(k-1)* covarianzas divididas entre *k-1*. Utilizando este hecho y el coeficiente de fiabilidad se puede deducir la fórmula para el coeficiente alfa de Cronbach (1951):

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_i \sigma_i^2}{\sigma_X^2}\right)$$

Si $\rho_{XX'} = \alpha$ entonces los subtest o formas de un test son paralelas.

Cuando un test está compuesto por formas no paralelas se puede estimar el límite inferior de su coeficiente de fiabilidad utilizando el coeficiente alfa. Como puede observese, su cálculo requiere del conocimiento del número de subtest o formas, la varianza de las puntuaciones compuestas y la suma de las varianzas entre los subtests. Por otra parte, un test puede considerarse como una variable compuesta por cada ítem como subtest; con esta consideración, el coeficiente alfa de Cronbach refleja, más que la estabilidad de las medidas, el grado en el que covarían los ítems que constituyen el test.

Existen otras maneras de estimar la fiabilidad del test en términos de la consistencia interna. Dado que α constituye una solución más general al problema, basta aquí citar cuatro de las formas más empleadas dentro del modelo clásico, como casos particulares de α. Estas son el coeficiente de Rulon (1939), la fórmula de Gutman (1945)– Flanagan(1937) y las fórmulas de Kuder-Richardson (1937), KR20 y KR21, denominadas de esta forma por ser precisamente el número 20 y el 21 de las presentadas por estos autores.

Otra manera de calcular el coeficiente alfa es mediante el uso del modelo de análisis de varianza de medidas repetitivas, puesto que los ítems pueden ser considerados como medidas repetidas de la misma característica: aquélla que se trata de evaluar con el test.

El análisis de varianza trata de analizar la variación de una respuesta, en este caso la puntuación, y de asignar porciones o componentes de esta variación a cada una de las variables de un conjunto de variables independientes. El razonamiento se basa en que las variables de respuesta se modifican por la variación de algún conjunto de variables independientes desconocidas. Como el investigador rara vez incluirá todas las variables que afectan a la respuesta en un experimento, es posible observar una variación aleatoria en la respuesta, aún cuando se mantengan constantes todas la variables independientes consideradas. El objetivo del análisis de varianza es identificar variables independientes importantes en un estudio y determinar cómo interactúan y afectan la respuesta. El análisis de varianza divide la suma de los cuadrados de las desviaciones en partes; cada una de estas partes se atribuye a una de las variables independientes en el estudio, más un residuo que se asocia con el error aleatorio.

Para los casos que se consideran y de acuerdo con la hipótesis de que la respuesta, en este caso, la puntuación no está relacionada con las variables independientes, se puede demostrar que cada una de las componentes de la suma total de los cuadrados de las desviaciones dividida entre una constante apropiada, proporciona un estimador insesgado para la varianza del error de medida. Cuando una variable está muy relacionada con la respuesta, su porción, llamada la suma de los cuadrados para esta variable, estará exagerada. Una manera de considerar esta fuente de variación está determinada por las diferencias entre todos los posibles pares de medias de la puntuación de un sujeto *j*, mientras más grandes sean las diferencias, más grande será la fuente de variación. La fiabilidad de la media de las *k* observaciones es la variabilidad que se debe a las puntuaciones verdaderas dividida entre la suma de la variabilidad que se debe a las puntuaciones verdaderas y la que se debe al error de medida.

Una vez que se ha encontrado el coeficiente alfa para una muestra, resulta interesante plantearse algunas preguntas acerca del valor del coeficiente para la población, utilizando la información de la muestra. Para ello se puede plantear una prueba de hipótesis en la que se confronte la significación estadística de este parámetro. Las hipótesis nula y alternativa son las siguientes α igual a cero y α diferente de cero. La teoría del error muestral para el coeficiente alfa ha sido desarrollada por diferentes autores, entre ellos destaca el trabajo de Kristof

(1963) y Feldt (1965) quienes propusieron la estadística de prueba $F=(1-\alpha)/(1-\hat{\alpha})$ que sigue, bajo la hipótesis nula, una distribución *F* con *(n-1)* y *(k-1)(n-1)* grados de libertad y donde *n* es el número de sujetos en la muestra, *k* es el número de ítems en el test, α el valor de alfa para la población y α el estimador de alfa. La decisión de rechazar la hipótesis nula o no rechazarla se realiza comparando el valor calculado *F* con el valor crítico de *F*, para las dos colas de la distribución, ya que se tiene una prueba de hipótesis bilateral. Si el valor calculado de *F* no cae en la región crítica entonces se rechaza la hipótesis nula.

La fiabilidad de un test no depende únicamente de las características propias del test sino del tipo de muestra de sujetos utilizados para calcularla, lo cual constituye una seria limitación para el modelo clásico, ya que se describe un instrumento de medida, como lo es el test, en función de los objetos medidos: los sujetos. La variabilidad de la muestra influye en la fiabilidad. La fiabilidad aumenta al aumentar la variabilidad. Por tanto, un dato imprescindible para la interpretación del coeficiente de fiabilidad es la variabilidad de la muestra en la que se calculó; en otras palabras, un test no tiene un coeficiente de fiabilidad fijo, éste depende de la muestra en la que se calcula.

Una vez que se conoce la fiabilidad de un instrumento, se pueden hacer ciertas estimaciones acerca de las puntuaciones verdaderas de los sujetos y del error que afecta a las puntuaciones empíricas. La estimación de las puntuaciones verdaderas se puede realizar a través de una tabla de frecuencias de las puntuaciones observadas, que es otra manera de presentar las puntuaciones de los sujetos en un test. Una vez estimada la puntuación media se desea saber qué tan variables fueron las puntuaciones de los examinados. La desviación de una puntuación es la distancia que existe entre la puntuación de un examinado y la media del grupo. La varianza muestral de un conjunto de puntuaciones es el cociente de la suma de los cuadrados de las desviaciones y *(n-1)* y resulta ser un estimador insesgado de la varianza de las puntuaciones de la población. Al obtener dichos estimadores se pueden hacer estimaciones acerca de la cantidad de error que afecta a las puntuaciones empíricas.

Las estrategias más comunes en la aplicación de la teoría clásica de los tests para la descripción de grupos, en lugar de sujetos particulares, es a través del cálculo del error típico de medida como la desviación típica de los errores de medida, que nos indica la discrepancia entre la puntuación verdadera de un examinado y la empírica. Para interpretar la puntuación verdadera a partir del error típico de medida que afecta a la puntuación empírica es necesario establecer un intervalo de confianza, éste puede basarse en la distribución normal de los errores o en la estimación de las puntuaciones verdaderas, de acuerdo a un modelo de regresión lineal.

Extensiones a la teoría clásica de los tests

A través de los años y dada la utilidad de la teoría clásica, se han ido añadiendo algunos elementos que la hacen más aplicable a las necesidades de quienes diseñan o evalúan tests. Entre estos desarrollos destacan la teoría de la generalizabilidad y el análisis de los ítems.

La teoría de la generalizabilidad analiza las fuentes del error de medida de una forma sistemática y desglosada. Ya se ha mencionado que en las práctica, es imposible encontrar condiciones de medida estrictamente paralelas. Las llamadas formas paralelas de un mismo instrumento no siempre presentan una equivalencia respecto de contenido, media y varianza. Por otra parte, resulta del todo imposible uniformar las situaciones de aplicación (mismo examinador, mismas formas paralelas).

En caso de que existieran las condiciones que permitieran lograr condiciones de medida idénticas, se estarían sesgando las puntuaciones de los sujetos, es decir, dichas condiciones podrían beneficiar sistemáticamente a algunos sujetos y perjudicar a otros. Esto no sería un obstáculo si la finalidad del estudio fuera analizar la estabilidad de un instrumento de medida, pero sí los es cuando se requiere estimar la puntuación verdadera del sujeto.

La teoría clásica tiende a la uniformidad de las condiciones de medida, la de la generalizabilidad, en cambio, toma en cuenta los múltiples factores que pueden producir variaciones en las puntuaciones de los sujetos y, mediante la aplicación de técnicas del análisis de varianza, permite estimar la varianza atribuible a cada uno de esos factores, así como a sus interacciones.

La teoría clásica de los tests define la puntuación verdadera de un sujeto como la media de sus puntuaciones en un amplio (infinito) número de tests estrictamente paralelos. En el marco de la teoría de la generalizabilidad, la puntuación universo (verdadera de un sujeto) se define como la media a través de todas las condiciones incluidas en el universo de generalización. A diferencia de la teoría clásica que sólo considera una puntuación verdadera para cada sujeto, la teoría de la generalizabilidad considera tantas puntuaciones universo como universos de generalización haya en el estudio.

En el proceso real de construcción de un test, se empieza por elaborar un número elevado de ítems, dos o tres veces más de los que el test tendrá finalmente; posteriormente se aplican esos ítems a una muestra de sujetos semejantes a los que el test se destina y se descartan aquellos ítems que no sean pertinentes, por ello es necesario contar con herramientas de análisis de la pertinencia de los ítems.

Otra forma de ampliar la información que se obtiene de la aplicación de instrumentos de medida en el ámbito de la teoría clásica de los tests consiste en realizar un análisis detallado de cada uno de los ítems que componen el instrumento. Este tipo de análisis proporciona información acerca de la pertinencia de cada uno de los ítems en el instrumento, a través de un estudio de las propiedades de los ítems que están directamente relacionadas con las propiedades del test y, en consecuencia, influyen en ellas.

Para llevar a cabo el análisis de los ítems se definen el índice de dificultad y el índice de discriminación. El índice de dificultad se define en términos de la proporción de sujetos que lo aciertan contra la proporción de sujetos que intentan responderlo y está directamente relacionado con la puntuación media del test. Este índice depende fuertemente de la muestra de sujetos en la que se calcula, es decir, no constituye una propiedad intrínseca del ítem. A nivel práctico, la teoría clásica reduce este inconveniente al calcular el índice de dificultad en muestras similares en competencia con aquéllas con las que se utilizará el instrumento. Dado que sería de esperar que las propiedades de los instrumentos de medida no dependiesen de los objetos medidos, este recurso resulta, a nivel teórico, poco convincente.

El índice de discriminación se define como la correlación lineal entre las puntuaciones de los sujetos en el ítem y sus puntuaciones en el test.

Existen varios métodos de correlación aplicables para estimar el índice de discriminación, dependiendo de las características de las variables que se van a correlacionar, en este caso el ítem y el test. Dado que habitualmente los ítems son dicotómicos (se aciertan o se fallan) y que el test constituye una medida cuantitativa discreta, con frecuencia se utiliza el coeficiente de correlación biserial puntual para calcular el índice de discriminación. Este coeficiente es una aplicación de la correlación lineal de Pearson cuando una de las variables es dicotómica y la otra cuantitativa continua o discreta. Esta correlación siempre obtiene un número entre -1 y 1. Un valor igual a 1 indica que la información que se obtiene del ítem es consistente con la que se obtiene del test como un todo; es decir, si el examinado obtuvo una buena puntuación en el ítem casi se podría pronosticar que obtendrá una buena calificación en el test, y si el índice es igual a -1, la información que proporciona el ítem no es consistente con el test, es decir, una buena puntuación en el ítem no implica nada acerca de la puntuación global del test.

Limitaciones de la teoría clásica de los tests

La teoría clásica es, como se dijo con anterioridad, un modelo simple que proporciona información útil, que puede ser empleado en una gran diversidad de situaciones y que no impone restricciones fuertes sobre el tamaño de la muestra que se requiere emplear, sin embargo tiene algunas limitaciones importantes.

La teoría clásica de los tests no proporciona información acerca de cómo se comportan los examinados de distinto nivel de habilidad en el rasgo medido en el ítem. Además, los parámetros que caracterizan a los ítems en la teoría clásica, como el índice de dificultad y el índice de discriminación, no son invariantes sobre diferentes grupos de examinados, es decir, dependen de la muestra de examinados a partir de la cual fueron obtenidos. Esta es una limitación fuerte para una teoría de medida, ya que en este caso, las estadísticas asociadas a los ítems son útiles únicamente para poblaciones de examinados semejantes en habilidad a la muestra con la que se calcularon las estadísticas, y la fiabilidad del test está relacionada con la variabilidad de las calificaciones del mismo.

Otras limitaciones de la teoría clásica de los tests consisten en que la comparación de los examinados, en la habilidad medida por el test, se limita a situaciones en las que se les administra el mismo test o uno paralelo; que no provee una base para determinar el rendimiento de un examinado frente a un ítem en particular, y que los procedimientos no son adecuados para detectar sesgos en las preguntas o para proporcionar una base adecuada sobre la cual construir pruebas "a la medida".

Para superar estas limitaciones es necesario construir modelos de medida para tests que puedan proporcionar información acerca de la precisión de una calificación, una estimación de la habilidad del sujeto y en los que los parámetros de los ítems no dependan de la muestra de examinados con la que se prueba un test. Los modelos logísticos de respuesta a los ítems, que se presentan a continuación, representan una opción frente a la teoría clásica de los tests.

Sección I.3 Teoría de respuesta a los ítems (IRT)

Cuando se elabora un test, la selección de las preguntas se hace, generalmente, basándose en los índices de dificultad y de discriminación de las mismas. Esta práctica funciona bien en muchas situaciones de diseño y de construcción de tests, pero podría mejorar si se usara más información acerca de las respuestas que los individuos dan a cada uno de los ítems.

Como se discutió en la sección anterior, los índices de dificultad y discriminación que se obtienen a partir del uso de la teoría clásica de los test dependen fuertemente de la muestra de sujetos elegidos para hacer las pruebas del test. Las limitaciones del modelo clásico pueden superarse si se desarrolla un modelo de medida que centre la atención en el análisis de los ítems propiamente dichos. Existen varios modelos que intentan superar por esta vía las limitaciones de la teoría clásica de los test; entre ellos destacan los llamados modelos logísticos que son la base de la teoría de respuesta a los ítems (IRT).

La teoría de respuesta a los ítems proporciona una herramienta alternativa a la teoría clásica de los tests para llevar a cabo la validación del instrumento de medición diseñado.

Esta teoría es un acercamiento al desarrollo de los tests que pretende producir un panorama más completo de cómo funciona un ítem. Los primeros trabajos que se realizaron en la dirección del estudio de los ítems binarios fueron los de Richardson (1936), Lawley (1943), Tuker (1946) y Lord (1952). En todos estos trabajos se utilizaba un modelo de ogiva normal para representar la curva de probabilidades de respuesta a los ítems. Rash (1960) y Birnbaum (1968) introdujeron los llamados modelos logísticos que son los que se emplean más frecuentemente en la actualidad y son con los que el presente trabajo se relaciona.

Análogamente al caso del modelo clásico, la teoría de respuesta a los ítems parte de algunos supuestos que sientan las bases sobre las que descansan todos los resultados.

Supuestos de la teoría de respuesta a los ítems

La hipótesis fundamental de esta teoría es que las respuestas a los ítems de un test responden a un cierto número de características que se quieren medir y que el número de estas propiedades es menor que el número de reactivos utilizados en el test.

La teoría de respuesta a los ítems se desarrolla sobre un modelo matemático que intenta describir la forma en que los examinados de diferentes niveles de habilidad responderían a un ítem. Este modelo es importante puesto que si el conocimiento de la forma en la que sujetos de diferentes niveles de habilidad en un cierto rasgo responden a un ítem en particular, permite establecer comparaciones entre los resultados de examinados que toman diferentes tests que contienen ítems equivalentes, y hacer posible, además, la aplicación de los resultados del análisis de un ítem particular a grupos con diferentes niveles de habilidad a los grupos utilizados como muestra para hacer el análisis de prueba de dicho ítem o incluso de un test.

La teoría de respuesta a los ítems toma como base al ítem, a diferencia del modelo clásico que se centra en el estudio del test como un todo.

En la teoría de repuesta a los ítems se supone que, si existen un conjunto *k* de habilidades que subyacen al desempeño de un examinado en un conjunto de ítems de un test, estas *k* habilidades definen un espacio en el que la posición de cada examinado se determina mediante su posición en cada uno de los ejes definidos por las habilidades. Si se supone, además, que un único rasgo determina la habilidad del examinado, el modelo se considera unidimensional.

A pesar de que la hipótesis de unidimensionalidad es la más común cuando se desarrollan tests utilizando la teoría de respuesta a los ítems, es claro que es difícil que un test cumpla esta condición. En la práctica lo que se requiere es que exista un factor dominante que influya en el desempeño en el test, factor al que se suele referir como la habilidad medida por el test.

Además de los modelos unidimensionales existen algunos modelos multidimensionales en la teoría de respuesta a los ítems. Sin embargo, el desarrollo técnico de éstos es todavía muy restringido, por lo que, en esta sección, nos limitaremos al estudio de los modelos unidimensionales.

Otro supuesto, relacionado con el de unidimensionalidad en la teoría de respuesta a los ítems, es el supuesto de independencia local. Éste se refiere a la hipótesis de que las respuestas de un examinado a los distintos ítems que conforman un test son estadísticamente independientes, es decir, que el comportamiento de un examinado al responder a una pregunta no afecta sus respuestas a los otros ítems del test.

En la teoría de respuesta a los ítems se cambia el uso de la noción de correlación entre reactivos para hablar de la relación entre las distintas variables, que es central a la teoría clásica de los tests, por los conceptos más generales de dependencia o independencia estadística. El concepto de independencia estadística puede expresarse de la siguiente manera: dos puntuaciones en dos ítems dicotómicos, *i* y *j*, son estadísticamente independientes si $p(+,+) = p_i(+)p_j(+)$; es decir, si la probabilidad de obtener un patrón de respuesta a ambos ítems, se puede calcular conociendo únicamente las probabilidades de respuesta de cada de ellos.

Un test es unidimensional si la dependencia estadística entre los ítems puede ser explicada mediante un rasgo único. Es decir, un test es unidimensional si los ítems que lo componen son estadísticamente dependientes en la población completa y si existe un único rasgo tal que los ítems son estadísticamente independientes en cada subpoblación de examinados cuyos miembros son homogéneos con respecto al rasgo. Puesto que la independencia está definida para una subpoblación de examinados localizados en un punto específico de la escala del rasgo, generalmente se identifica a este concepto como independencia local.

Es importante recalcar que los conceptos de independencia local y unidimensionalidad no son conceptos equivalentes. La dimensionalidad del test está relacionada con el número de rasgos necesarios para lograr independencia local.

Si dos ítems son estadísticamente dependientes, comparten una misma fuente de varianza. Si existe un rasgo tal que para los examinados que son homogéneos con respecto al rasgo, los ítems son localmente independientes, entonces la varianza común se puede explicar totalmente por ese único rasgo y el test cumple el requerimiento de unidimensionalidad.

En la práctica, nunca se puede decir con certeza que existe un solo rasgo (o cualquier otro número) tal que los ítems de un test son localmente independientes, por lo que la independencia local y el número de rasgos deben suponerse. Lo que sí es posible es verificar la validez de estas suposiciones.

Decir que un rasgo subyace la actuación en los ítems de un test es lo mismo que afirmar que el rasgo da cuenta de la dependencia estadística entre los ítems. Es importante destacar, sin embargo, que como tal, el rasgo no es necesariamente una medida válida del constructo que se intenta medir con el test; es posible que ese rasgo subyazca a los resultados que se obtienen de la aplicación de un conjunto de ítems, y que éste sea una medida pobre del constructo que los ítems intentan medir. Por esta razón es necesario mantener la diferencia entre la noción de rasgo y la de constructo.

Conceptos fundamentales de la teoría de respuesta a los ítems

La teoría de respuesta a los ítems utiliza como unidad de análisis el ítem, en lugar del test completo. Intenta poner en relación la capacidad del sujeto en un cierto rasgo que se desea medir con su rendimiento en el ítem, de tal forma que se logren superar las dos limitaciones fundamentales de la teoría clásica, es decir intenta desarrollar un modelo de medida en el que el resultado de la medición no dependa del test utilizado y en el que las propiedades del instrumento de medida no estén en función de los sujetos a quienes se aplica el instrumento.

Los modelos que se utilizan en la teoría de respuesta a los ítems asumen una relación funcional entre los valores de la variable que miden, por ejemplo la capacidad de comprensión de un concepto, la habilidad en un área o las preferencias, que se representa por la variable independiente θ, y la probabilidad de responder acertadamente a cada uno de los ítems. Esta relación funcional se representa mediante una curva de forma sigmoide que se denomina curva característica del ítem. La forma de la curva se adopta con base en la idea de que la probabilidad de acertar un ítem dicotómico es 1 si se posee la capacidad que se intenta medir y 0 en caso contrario. Si bien la forma de las curvas características de todos los ítems es la misma, la curva que representa a cada uno de ellos es específica y diferente; está definida por parámetros específicos que la distinguen de las demás.

El concepto de curva característica del ítem es central a todos los modelos de respuesta. En todos ellos se parte de esta relación funcional entre la probabilidad de responder correctamente a un ítem y la habilidad que se intenta medir mediante el reactivo. Una curva característica del ítem es así una función matemática que relaciona la probabilidad de responder correctamente un ítem con la habilidad medida por el conjunto de ítems de la prueba en la que ese ítem está contenido. Usualmente se considera a la curva característica como la función de regresión no lineal de la calificación del ítem sobre el rasgo que se quiere medir con el test.

En los modelos de la teoría de respuesta a los ítems la probabilidad de acertar un ítem depende exclusivamente de la variable θ. Este hecho nos indica que la teoría de respuesta a los ítems supone implícitamente en su formulación que las preguntas destinadas a medir la variable θ son unidimensionales. Si se cumple el supuesto de unidimensionalidad puede demostrarse que existe independencia local entre los ítems, es decir, que la respuesta de un sujeto con un valor determinado en la variable θ a un ítem particular no se ve influida por sus respuestas en los otros ítems.

La diferencia fundamental entre los distintos modelos de respuesta a los ítems es la forma que se le asigna a la curva característica. Si fuera posible conocer tanto las calificaciones en el test como las de la habilidad medida, se podría descubrir la forma de la curva tomando como punto de partida alguna consideración acerca de la distribución de las calificaciones de los ítems para un nivel fijo de la habilidad. Es decir, la curva característica podría encontrarse calculando la media de cada distribución y posteriormente uniendo las medias de cada una de ellas.

Si el espacio completo de la habilidad está definido para toda una población de interés, las distribuciones de las calificaciones, a un nivel dado de habilidad, deben ser las mismas a través de las poblaciones. Por otra parte, si las distribuciones de probabilidad son idénticas, las curvas que conectan las medias de las distribuciones deben ser también idénticas; por ello puede decirse que la curva característica permanece invariante ante las poblaciones de examinados para los que se definió el espacio de habilidades.

Cuando se mide sólo una habilidad, la curva se conoce como una curva característica del ítem. Cuando el modelo es multidimensional, la función obtenida se conoce como función característica del ítem.

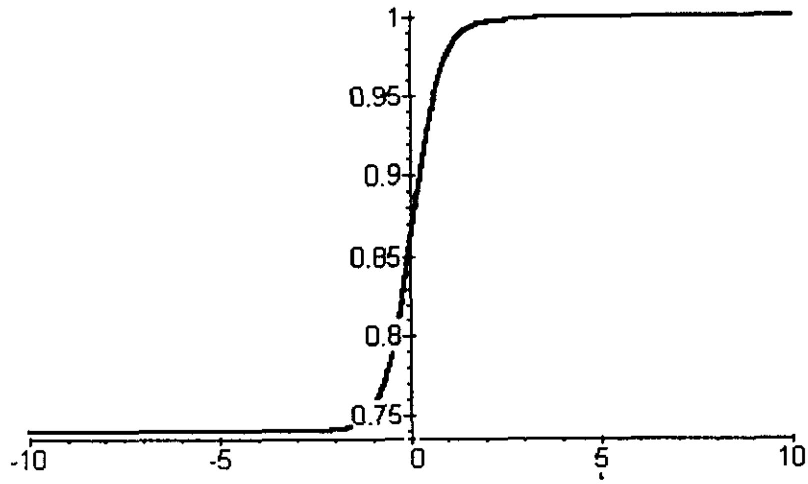
Puesto que la probabilidad de que un examinado responda correctamente a un ítem depende únicamente de la forma de la curva característica, esa probabilidad es independiente de la distribución de la habilidad del examinado en la población de examinados de interés; entonces, la probabilidad de que un examinado responda correctamente al ítem no depende de cuántos examinados se encuentren en ese mismo nivel de habilidad.

La invarianza de la curva característica frente a la población de examinados que se utiliza para calibrar la prueba es una de las propiedades más atractivas de este tipo de modelos. Las implicaciones de este resultado son sumamente importantes en distintos tipos de aplicación de la teoría de medida, como por ejemplo la construcción de pruebas a la medida, la elaboración de bancos de preguntas y el estudio del sesgo en los ítems.

La elección del tipo de función característica que se usará como modelo depende del tipo de test al que se quiera aplicar y del tipo de herramienta de análisis de que se disponga. En todos estos modelos es frecuente suponer que la respuesta al ítem es dicotómica, es decir, que se acierta o se falla el ítem, independientemente del número de alternativas que se tengan o de que las preguntas sean abiertas.

En la mayoría de las aplicaciones de esta teoría se supone que la curva característica del item tiene forma de S, como la que se muestra en la figura I.3.1, de manera que, conforme la habilidad en el rasgo aumenta, la probabilidad de responder correctamente también aumenta. La importancia de la curva característica del ítem, en contraste con los índices de dificultad y de discriminación de la teoría clásica, reside en que hace posible ver la dependencia la probabilidad de contestar correctamente al reactivo con el rasgo a medir.

Figura I.3.1



Si bien la forma sigmoide, de S, o de función escalón, para la curva característica de los ítems es, como se acaba de mencionar, la más empleada en la construcción de tests, no es la única posible y, de hecho, existen modelos que utilizan otras formas de curva para la curva característica. La curva en forma de escalón implica que existe un valor mínimo θ' para la puntuación del rasgo, por abajo de la cual, los examinados no pueden responder correctamente al ítem; sin embargo, un examinado con un nivel de habilidad mayor o igual que θ' lo responderá correctamente. Las curvas escalón no se usan tan frecuentemente en la construcción de tests como las de forma de S porque los datos reales son generalmente más consistentes con la curva en forma de S y porque esta última es más adecuada para el manejo matemático.

Los modelos de respuesta a los ítems que se emplean con mayor frecuencia en la práctica son aquéllos que suponen para la curva característica de los items una distribución acumulativa logística y por ello se les conoce también como modelos logísticos.

La definición de cada curva característica depende, generalmente, de tres parámetros: el índice de discriminación que permite decidir si la pregunta distingue entre los sujetos que tienen la capacidad para responder acertadamente el ítem y aquellos que no la tienen, el índice de dificultad, que, a diferencia de la teoría clásica no se define en términos del número de sujetos que aciertan el ítem, sino directamente de la misma función como el valor de θ en el que la pendiente a la curva característica alcanza su máximo valor, y por último la probabilidad de acertar el ítem al azar.

La función logística utilizada en este tipo de modelos es de la forma:

$$p_g(\theta) = \frac{e^x}{1 + e^x}, \text{ con } x = x(\theta)$$

Entre los modelos logísticos que se emplean con mayor frecuencia se encuentran los de uno, dos y tres parámetros.

El modelo de dos parámetros se puede expresar con la función logística antes mencionada con $x(\theta) = Da_g(\theta - b_g)$, en esta expresión D es una constante arbitraria, a_g es el índice de discriminación del ítem g, b_g es el índice de dificultad del ítem g y $p_g(\theta)$ se interpreta como la proporción de examinados que tienen un nivel θ de habilidad y que responden correctamente al ítem. Es común elegir la constante D como D = 1.7 ya que puede demostrarse que, de esa manera, la probabilidad obtenida mediante la función logística y mediante la ojiva normal (distribución normal acumulada) no difieren por más de 0.01 para cualquier valor de θ (Lord y Novick, 1968).

El modelo logístico de un parámetro, que es el más empleado, se obtiene si se toma el índice de discriminación como constante; es decir, si se supone que todos los ítems tienen el mismo índice de discriminación. A este modelo también se le conoce como modelo de Rasch.

Cuando se aplica el modelo logístico a datos obtenidos de pruebas de opción múltiple, se debe tener en cuenta que, en esas condiciones, se presenta la posibildad de que haya aciertos al ítem obtenidos de respuestas al azar. Si se admite esta posibilidad, la ecuación para la curva característica se convierte en:

$$p_{g\,3}(\theta) = c_g + (1 - c_g)\, p_{g\,2}(\theta),$$

en la que $p_{g\,3}(\theta)$ es la probabilidad de acertar el ítem g en el modelo de tres parámetros, $p_{g\,2}(\theta)$ es la probabilidad de acertar al ítem g en el modelo de dos parámetros y c_g se conoce como el parámetro de pseudo-aciertos al azar.

En la práctica, es imposible encontrar directamente la variable que se va a medir. No es posible siquiera verificar qué tan apropiada es la elección de modelo, ya que las regresiones se hacen sobre una variable que no se puede medir directamente. La única manera de verificar qué tan apropiado resulta el modelo elegido consiste en analizar la validez de las predicciones obtenidas mediante su aplicación; si son aceptables, el modelo se considera adecuado. Por ello es necesario elegir el modelo que se desea emplear de manera arbitraria.

La ventaja más importante que representa el uso de estos modelos consiste en que, dado un conjunto de ítems de un test que se ha ajustado a un modelo de respuesta, es posible estimar la habilidad de un examinado en la misma escala de habilidad de cualquier subtest de ítems en el dominio de los ítems que fueron ajustados al modelo. Los modelos de respuesta a los ítems proveen así una forma de comparar a los examinados aunque hayan respondido a diferentes subconjuntos de ítems del test.

La estimación de las variables θ para los sujetos es una estimación condicional de máxima verosimilitud que se hace conjuntamente con la de los otros parámetros. La forma en la que esta estimación se hace se describirá con mayor amplitud en el siguiente capítulo. A partir de los valores obtenidos para los parámetros: *a* = índice de discriminación, *b* = índice de dificultad y *c* = factor de aciertos al azar, se pueden graficar las curvas características correspondientes a cada uno de los ítems.

Una manera de identificar aquéllos ítems para los cuales la estimación de los parámetros es pobre es mediante el estadístico inferencial x^2. Estos ítems se caracterizan por tener una x^2 muy grande en relación con las demás.

Los resultados que se obtienen del análisis de los ítems se pueden utilizar en la teoría de respuesta a los ítems para obtener una caracterización del cuestionario completo. Esta caracterización permite asimismo establecer un puente entre los resultados obtenidos mediante esta teoría y los obtenidos mediante la aplicación del modelo clásico. Para ello se define la curva característica del test. Esta curva se obtiene sumando las curvas características de los ítems.

Una forma de comparar los resultados obtenidos al analizar un test utilizando la teoría de respuesta a los ítems con aquellos obtenidos a partir de la teoría clásica de los tests es mediante la obtención de la puntuación verdadera en el test de un sujeto o sujetos a los que se ha estimado una cierta puntuación θ_j en la variable a medir con la teoría de respuesta a los ítems. Esta puntuación verdadera se estima calculando la suma de las probabilidades dadas por las curvas características de los ítems que componen el test para el valor θ_j. Cuando las puntuaciones se obtienen de esta manera, los resultados quedan expresados en la misma escala que las puntuaciones empíricas, en lugar de quedar expresadas en la escala de las θ, que es diferente. La transformación de escala se logra mediante la curva característica del test, con la ventaja adicional de que estas puntuaciones verdaderas, de acuerdo a los supuestos de la teoría de respuesta a los ítems, no dependen del test utilizado.

Con el fin de comparar los resultados obtenidos mediante la teoría de respuesta a los ítems se puede calcular también la curva característica esperada para algunos sujetos y las curvas características correspondientes a esos mismos sujetos y comparar las diferencias. La curva característica empírica del sujeto se obtiene graficando el índice de dificultad *b* contra la proporción de ítems acertada por el sujeto. La curva teórica se obtiene, en cambio, graficando el índice de dificultad *b* contra la proporción de aciertos que el sujeto debiera tener según la curva característica de los ítems, es decir, la media de los valores $p(\theta)$ de los ítems de cada categoría para un valor dado de θ para el sujeto en cuestión.

Para lograr un análisis más completo del test como un todo, es importante contar con un indicador de su precisión. Con este propósito se define la función de información del test. Esta función se define de manera tal que cuando la función de información sea grande, el error típico de medida sea pequeño y la información aportada por las estimaciones sobre el parámetro θ sea grande.

La función de información de cada ítem está dada por:

$$I(\theta)=\frac{[Pi(\theta)]^2}{Pi(\theta)Qi(\theta)}$$

y la función de información del test se obtiene sumando las funciones de información de cada uno de los ítems. La función de información permite además, elegir la zona de θ para la cual el test es más adecuado y puede emplearse para establecer intervalos de confianza para distintos valores de θ. En el modelo logístico de tres parámetros, la información aportada por los ítems es máxima para un valor de θ dado por

$$\theta = b + 1/\ Da\,[\ \ln(1/2 + 1/2\ \sqrt{1+8c}\ )]$$

El cálculo del error de medida es fundamental en la definición de cualquier teoría de medida. Este nos proporciona información acerca de la diferencia entre las puntuaciones verdaderas y las empíricas. De hecho, el error típico de medida es la desviación típica de esas diferencias. En la teoría de respuesta a los ítems, el valor de este error no es el mismo para todos los sujetos; está en función de θ, lo cual indica que la precisión con la que miden los tests no es uniforme a lo largo de la escala de la variable θ.

La forma más conveniente de calcular el error típico de medida es mediante la función de información del test:

$$Se = \frac{1}{\sqrt{I(\theta)}},$$

en la que Se es el error de medida e $I(\theta)$ es la función de información del test.

Ventajas de la teoría de respuesta a los ítems

Los modelos de respuesta a los ítems presentan ventajas grandes cuando se comparan con la teoría clásica de los tests. Entre ellas pueden mencionarse las siguientes:

- Suponiendo la existencia de una población grande de examinados, los descriptores de un ítem de la prueba son independientes de la muestra de examinados que se utilizó para calibrar el test.

- El modelo proporciona una estadística que indica la precisión con la que se estimó la habilidad de cada examinado.

- El concepto de confiabilidad de formas paralelas se reemplaza por el concepto de estimación estadística y los errores estándares asociados a ella.

Estas ventajas han hecho que la Teoría de Respuesta a los Ítems sea la que se usa más frecuentemente en la medida a través de las pruebas estandarizadas que se realizan en todo el mundo.

Limitaciones de los modelos IRT en cuanto a la posibilidad de diagnóstico

Los modelos de respuesta a los ítems presentan algunas desventajas. En términos teóricos las desventajas de las teorías de respuesta a los ítems radican, por una parte, en que se basan en hipótesis fuertes que limitan su aplicabilidad a muchos conjuntos de datos y, por otra parte, en que la estimación de los parámetros requiere de la aplicación de los ítems a muestras muy grandes de sujetos.

Los modelos estándar logísticos IRT caracterizan a los sujetos en términos de su propensión a responder acertadamente a un ítem, sin hacer ninguna consideración relativa a la forma en la que los sujetos aprenden una disciplina particular, ni sobre la estructura misma del rasgo que se desea medir.

Los parámetros que caracterizan a los sujetos se estiman, en estos modelos, mediante métodos de máxima verosimilitud y criterios de índole puramente estadística. Por esta razón, dichos parámetros están fuertemente relacionados con el porcentaje de respuestas correctas.

Si bien estos modelos son apropiados para pruebas en las que todos los sujetos emplean la misma estrategia para resolver los ítems, en cuyo caso los parámetros del los ítems reflejan el número o la complejidad de las operaciones necesarias para resolver un ítem dado (Fisher, 1973 y 1995), resultan menos satisfactorios cuando los sujetos emplean estrategias diferentes en la solución de los problemas. En estos casos, es posible cuestionarse acerca de la validez del uso de calificaciones que proporcionan poco más que el porcentaje de aciertos para comparar a los estudiantes y tomar decisiones que pueden, en ciertos casos, afectar su futuro.

Por otra parte, la información que se obtiene de los modelos logísticos es insuficiente para determinar las razones por las cuales un ítem particular puede ser fácil para sujetos que siguen una estrategia, pero difícil para sujetos que siguen otra, e incluso para determinar si el diseño mismo del ítem favorece cierto tipo de estrategias en la respuesta.

Cuando se usan los modelos IRT tradicionales, la única caracterización que se puede hacer de un examinado es en términos de su propensión a responder correctamente en un cierto dominio de ítems. Consecuentemente, los procesos y los resultados del aprendizaje sólo se pueden plantear como cambios en el desempeño global. Esta caracterización resulta limitada cuando se tiene en consideración el avance de las ciencias cognitivas y de la investigación en los procesos de enseñanza y de aprendizaje de disciplinas concretas, tales como la física, la química, las matemáticas o la lecto-escritura. Los resultados de la investigación en estas áreas brindan información útil respecto a la forma en la que las personas resuelven problemas, adquieren conocimientos e incrementan sus habilidades y sus conocimientos (Snow y Lohman, 1989).

Ahora se sabe que los aprendices se vuelven más competentes, no a través de la acumulación de más información, sino a través de la reestructuración de su conocimiento previo. Cuando se planea o se evalúa la instrucción, no resulta de tanto interés preguntarse por ejemplo ¿cuántos ítems puede resolver un estudiante? o ¿qué proporción de la población puede obtener calificaciones menores que la de ese mismo estudiante? Es, por el contrario, más interesante intentar responder ¿qué es lo que esta persona puede estar pensando para que sus acciones tengan sentido desde esta perspectiva? (Thompson, 1982) o ¿qué organización tiene en mente el estudiante de forma que sus acciones se puedan representar mediante un patrón coherente de respuestas?, ¿cómo se puede caracterizar a cada sujeto en términos del tipo de estructura mental y de la disciplina particular?, ¿cuáles son sus necesidades a futuro?, entre otras. Para responder estas preguntas, se requiere de modelos de medida que puedan dar cuenta del estado de desarrollo de los sujetos de acuerdo a indicadores del nivel en el que se encuentran en cuanto a sus estructuras de conocimiento, en cuanto a su habilidad y al tipo de procesos cognitivos que suelen emplear en la solución de diversos tipos de problemas.

La extensión de los modelos IRT hacia modelos que incluyan estrategias múltiples de respuesta o que intenten caracterizar el estado del conocimiento de los sujetos a la luz de los resultados de las ciencias cognitivas en la actualidad tiene una gran diversidad de aplicaciones potenciales: Puede, por ejemplo, proveer de un marco riguroso para probar teorías alternativas sobre procesamiento cognoscitivo; puede ayudar a la obtención de estimaciones de la forma en la que los estudiantes resuelven los problemas específicos; puede proporcionar información diagnóstica más completa y confiable. Estas y otras aplicaciones puede ser mucho más valiosas para propósitos de diagnóstico, remediales y de revisión de currícula que simplemente conocer cuántos problemas puede resolver correctamente un sujeto (Messik, 1984).

En contraste con los modelos clásicos y con los modelos IRT, los modelos que permiten la introducción de estrategias alternativas en la solución de un ítem o que permiten determinar las zonas en las que el sujeto entra en conflicto al resolver cierto tipo de problemas, o que permiten acercarse a las estructuras cognoscitivas, formadas o en formación, en los sujetos, debe comenzar por proponer explícitamente los procesos mediante los cuales los sujetos alcanzan cierto tipo de respuesta. La formulación de estos modelos requiere entonces de la existencia de una teoría psicológica o educativa sólida que dé cuenta de los procesos involucrados en la solución del ítem, además de un modelo matemático que permita caracterizar al ítem y obtener la información deseada.

Tanto la teoría clásica de los tests como la de respuesta a los ítems son modelos matemáticos en los que no se incorpora ninguna consideración de tipo cognitivo que pueda afectar de una u otra manera la forma en la que los sujetos responden los ítems. A medida que las teorías cognitivas se han ido consolidando, se ha hecho un esfuerzo por construir modelos de medida que incluyan un modelo matemático al mismo tiempo que un modelo cognitivo. El resultado de este esfuerzo ha sido la aparición de una gran diversidad de nuevos modelos de medida que se considerarán más adelante.

El modelo que se plantea en el capítulo II de este trabajo intenta subsanar la incapacidad de las teorías clásica de los tests y de respuesta a los ítems de ser utilizadas como herramienta de diagnóstico.

Sección I.4 Otros modelos en la teoría de los tests

En los últimos años se ha visto surgir un movimiento de reforma que busca nuevos paradigmas para la construcción e interpretación de los tests. Muchas de las prácticas y de las ideas tradicionales concernientes a la teoría de los tests se han cuestionado. Las discusiones entre especialistas han puesto de manifiesto la inminente necesidad de introducir nuevos modelos que incluyan ítems de mayor calidad, mejores distractores y que proporcionen más información acerca de las estrategias y de los procesos mentales de los sujetos involucrados en la solución de las preguntas planteadas. Este movimiento ha ido creciendo en la actualidad a pesar de que las limitaciones para la aplicación de los nuevos modelos siguen vigentes: la tecnología requerida para ajustar las funciones que se proponen en estos modelos para las curvas características de los ítems o para la conformación de mapas de trayectorias de conocimiento está todavía lejos de llegar a una forma en la que estos modelos tengan una aplicación práctica.

En la literatura reciente tanto las formas de construcción de los tests cuanto los modelos tradicionales de medida se han puesto en tela de juicio. Sin descartar su utilidad y los beneficios que se obtienen de su aplicación, se ha considerado necesario tomar en cuenta que actualmente las posibilidades de medida se han visto beneficiadas, por una parte, por los desarrollos técnicos en los ámbitos de las matemáticas y de la computación y, por otra parte, por la integración, dentro de la teoría de los tests, de las teorías cognitivas. Ambos desarrollos permiten contar con interpretaciones de los datos más adecuadas y de mayor utilidad.

Los tests tienen hoy en día una gran diversidad de aplicaciones: se utilizan con fines de diagnóstico, para la certificación, con propósitos de clasificación y como una herramienta de investigación. A partir de los resultados que de ellos se obtienen se toman decisiones de diversa índole que afectan a una gran cantidad de personas. Esta multiplicación de usos hace que la necesidad de profundizar en el estudio de los tests, en su formato, en su validación y en la interpretación de los resultados que se obtienen de su aplicación se haga cada día más patente.

Dos debates que tienen importancia desde la perspectiva de este trabajo se han llevado a cabo durante los últimos años. El primero centrado en el diseño de los ítems y, en particular, en su formato y el segundo alrededor de las técnicas o modelos de medida que son más pertinentes para el análisis de los ítems. Analizaremos brevemente los resultados de cada uno de ellos.

Debate en torno al formato de los ítems de los tests

En cuanto al diseño de los tests, una de las cuestiones que ha recibido mayor atención es la referente al formato de los ítems. Durante las décadas de 1970 y 1980 el tipo de examen más comúnmente empleado fue referido al criterio, es decir aquél en el que se utilizaban ítems referidos a un objetivo específico planteado con la mayor claridad posible, que se constituían en los criterios de diseño de cada uno de los ítems.

Este tipo de tests se diseña a partir de objetivos conductuales y provee, de manera efectiva, una definición operativa del aprendizaje que se desea lograr en los estudiantes. Sin embargo, más recientemente ha surgido un movimiento que apoya con mayor fuerza los tests en los que se usan los ensayos y la investigación en las pruebas de desempeño y de aptitudes. Entre los nuevos modelos de test que han surgido destacan las llamadas pruebas de evaluación auténtica.

El formato de opción múltiple ha sido criticado tanto por los diseñadores de instrumentos de medida como por los educadores, sobre todo porque, de acuerdo a los críticos, este formato limita a los sujetos a realizar actividad mental de bajo nivel que no es lo deseable cuando se intenta medir su desempeño o su habilidad (Frederiksen, 1984; Snow, 1993). Algunos críticos han incluso hecho acusaciones en el sentido de que este tipo de formato ha creado un estilo de enseñanza que es dañino debido a que no estimula el pensamiento crítico (Shepard, 1993).

El formato de opción múltiple, sin embargo, tiene cualidades que le han permitido salir victorioso de los ataques recientes. Se ha demostrado, por ejemplo, que los formatos de opción múltiple y los llamados formatos abiertos o de respuesta construida están fuertemente correlacionados; que con los ítems de opción múltiple se pueden evaluar muchos de los objetivos de mayor profundidad y complejidad del aprendizaje de una manera tan rigurosa como la que se logra con los de formato abierto; y que los ítems de opción múltiple permiten la detección y eliminación de sesgos de una manera objetiva (Snow, 1993). Se ha encontrado, además, que es posible transformar pasajes escritos a ítems de opción múltiple (Bormouth, 1970) y las revisiones integradoras realizadas por Albanese (1992), Downing (1992), Frisbie (1992) y Soloway (1990), entre otros, ofrecen una guía para entender mejor la variedad de formatos de opción múltiple que pueden emplearse hoy en día en la construcción de ítems.

Dentro de esta discusión se ha llegado a sugerir que en la medida de procesos cognoscitivos complejos y dinámicos son tal vez más útiles los formatos de respuesta construida o abierta (Bennett, et al. 1990), mientras que los ítems de opción múltiple son útiles en la medida de conocimiento estático. Estas aseveraciones han estado basadas muy fuertemente en concepciones de sentido común y se ha hecho poca investigación que realmente favorezca un tipo de diseño sobre el otro. Entre las investigaciones que se han llevado a cabo de forma sistemática, se encuentra la realizada por Birnbaum y Tatsuoka (1987) quienes examinaron el efecto del formato de respuesta en el diagnóstico y concluyeron que los ítems de opción múltiple pueden no proporcionar la información apropiada para la identificación de las ideas previas de los alumnos. Por otra parte, las investigaciones de Snow (1993) revaloran la importancia del formato de opción múltiple en cuanto a sus posibilidades de relación con los resultados de las ciencias cognitivas.

Frente a la incertidumbre generada por este tipo de discusiones y a la dificultad de asegurar la efectividad de cualquier formato de diseño de ítems es conveniente tomar una posición más balanceada. Varios investigadores (Snow,1993) han propuesto una agenda de investigación cuyo fin sea precisamente el de examinar a profundidad qué tipo de constructos importantes no se han tomado en cuenta en ninguno de los formatos existentes, así como las ventajas y limitaciones de los distintos tipos de formato para los ítems ante diversas situaciones, problemas y en diferentes áreas del conocimiento. Los resultados obtenidos no inclinan la balanza con claridad hacia ninguna de las dos posiciones, por lo cual, en el diseño de instrumentos, se ha seguido una política que responde más directamente a las necesidades de la aplicación del instrumento que a las posturas teóricas subyacentes.

Los test que utilizan ítems de opción múltiple se han revalorado a su vez en los últimos años a través de la relación que puede establecerse entre el tipo de distractores empleados y los resultados de la investigación en las ciencias cognitivas y en la didáctica de los contenidos específicos. La tendencia actual se aboca al desarrollo de instrumentos sustentados fuertemente en las teorías cognitivas, tanto en la parte relativa al diseño de los ítems, cualquiera que sea el formato empleado, cuanto en la parte correspondiente a los modelos de medida.

El diseño de los ítems debe responder a aquello para lo que se intentan utilizar los resultados del proceso de medida. De acuerdo a Haladyna (1994), el diseño de los ítems debe modelar el rango de comportamiento que distingue a los aprendices con respecto a lo que se está aprendiendo y midiendo. Otra característica deseable en el diseño de ítems consiste en tomar en consideración el valor diagnóstico de las elecciones de los sujetos al responder, esto implica que los distractores diseñados deben basarse con mayor firmeza en los errores comunes de los estudiantes (Haladyna y Downing, 1989) y en las tendencias y estrategias de respuesta que aparecen reportados cada vez con mayor frecuencia en la literatura. La idea central del diseño actual consiste en generar ítems con un conjunto predecible de parámetros a partir de los cuales sea posible hacer interpretaciones claras en cuanto a las posibilidades cognitivas o psicológicas del sujeto.

A la evaluación de los distractores de los ítems de opción múltiple, se le ha prestado poca atención. El índice de discriminación de los modelos logísticos más empleados en la actualidad no proporciona suficiente información acerca de cuáles distractores deberán ser retirados o revisados para mejorar el desempeño del ítem y, por ende, el del test. En la calificación de los ítems polítomos, la discriminación se conceptualiza de manera diferente. Este tipo de calificación ve la información diferencial contenida en los distractores de manera más sensible que en el caso de la corrección dicotómica. Los distractores discriminantes, sin embargo, son poco frecuentes; por ello, de acuerdo con Haladyna y Downing (1997) es importante prestar atención al número adecuado de distractores en cada ítem, para obtener información suficiente. Por otra parte Huynh (1996) demostró que, en el caso específico del modelo de Rasch, no es necesario escribir ítems con más de tres categorías de respuesta, porque cada uno de esos ítems es teóricamente equivalente a un conjunto de ítems independientes con sólo dos o tres categorías de respuesta. Esta reconceptualización de la discriminación implica la evaluación de los distractores así como la consideración de los patrones de respuesta de cada distractor relativo a los otros distractores. Los ítems que tienen patrones similares de respuesta, a menos que reflejen diferentes preconceptos o estrategias, pueden no ser muy útiles en el diseño de tests.

Si los distractores se diseñaran teniendo en cuenta, como ya se ha mencionado, las ideas previas de los estudiantes, sus estrategias más comunes y el tipo de razonamiento que suelen emplear frente a conceptos diversos, las técnicas de evaluación permitirían la validación empírica de las respuestas a los mismos. Esto redundaría necesariamente en una mayor y mejor posibilidad de diagnóstico. La evaluación de los distractores, la conceptualización de la discriminación del ítem y el análisis de los distintos factores que entran en juego en la respuesta a un ítem se presentan entonces, como un imperativo a los futuros modelos en la teoría de los tests.

Se requiere de un programa amplio de investigación dirigido hacia la comprensión del impacto del diseño de los ítems en las respuestas y en el tipo de estrategias seguidas por el sujeto que responde, así como a la profundización en el significado de la fiabilidad y de la validez de test basados en diferentes tipos de diseño.

Los desarrollos de modelos de medida no pueden permanecer al margen de esta discusión. Frente a los retos que el diseño de los tests plantea, se ha respondido mediante distintos tipos de acercamiento a los problemas teóricos.

Debate en torno a otros modelos de medida

Si bien el formato de los ítems es un factor decisivo en la estructuración de tests de todo tipo, la posibilidad de estudiar las consecuencias de los usos de los distintos formatos y de profundizar acerca de la forma en la que los instrumentos dan cuenta del constructo que están destinados a medir, radica fuertemente en la teoría de medida. Medir es conocer. Es a través de la interpretación de los datos obtenidos de las respuestas a los ítems como se llega a la toma de decisiones. La investigación en esta área de la teoría de los tests es igualmente importante si se desea llegar a un proceso de toma de decisiones a partir de los resultados de las pruebas que sea más profundo y confiable. Por ello, en los últimos años la medida ha sido un área fecunda de discusión y de estudio.

La revisión de los avances en esta disciplina muestra tendencias diferentes de investigación, que se discutirán a continuación.

Una de las líneas de investigación en medida ha permanecido en la tradición que se centra únicamente en el desarrollo de teorías matemáticas generales para la medida en educación. Así, han aparecido los modelos logísticos multivariados. Estos modelos toman en cuenta la elección de las diferentes opciones de respuesta que se presentan en un ítem, tomando como punto de partida los modelos logísticos tradicionales y generalizan los modelos dicotómicos integrándolos a modelos categóricos multivariados. (Abrhamowicz y Ramsay, 1992; Gloneck y McCullagh, 1995).

Dentro de estos esquemas se pueden incluir algunos de modelos de crédito parcial (Huynh y Ferrara, 1994; Yen,1993; Andersen, 1966 y 1985; Andrich, 1978; Masters, 1982; Muraki, 1992) que permiten manejar los datos recogidos de ítems polítomos en los cuales las respuestas se califican en escalas que contienen más de dos categorías. También es posible incluir en esta categoría los modelos de respuesta graduada (Bock, 1972; Samejima, 1972; Thissen y Steinberg, 1984; Muraki, 1992; Andrich, 1995) en los cuales la escala se considera un continuo. Para este tipo de modelos Thissen y Steinberg presentan una taxonomía (1986), a la que nos referiremos más adelante, en la que muestran que los únicos acercamientos diferentes son el de Samejima (1979) de respuesta graduada y el modelo nominal de Bock (1972); los otros modelos pueden reducirse a uno de estos dos acercamientos.

El modelo de Samejima es un modelo de diferencia que requiere que la variable independiente se mida en una escala ordinal. Las alternativas en las opciones se ordenan con respecto a su grado de acercamiento a la solución del problema. Los resultados de la estimación de estas funciones, sin embargo, dependen fuertemente del ordenamiento de las respuestas y por ello no se prestan a la evaluación de los distractores.

Los modelos del tipo del de Bock son modelos divididos por el total. En ellos se requiere de la estimación de *k* funciones diferentes de probabilidad, una para cada una de las opciones de respuesta al ítem. Las funciones de probabilidad incluyen la restricción de exhaustividad, es decir que la suma de las *k* funciones de probabilidad para un valor de la habilidad dado sea uno. Este tipo de modelos permite el análisis de datos nominales y ordinales de opción múltiple. En ellos la interpretación de los parámetros es, sin embargo, difícil y éstos deben verse sólo como un medio para obtener las formas de las distintas funciones de probabilidad; no pueden, de manera alguna, ligarse a las características de origen cognitivo de los distractores. Los resultados del análisis multicategórico de este tipo de modelos se representan gráficamente, son estas gráficas las que proporcionan la mayor parte de la información que puede obtenerse de ellos (Thissen, Steinberg y Fizpatrik, 1989).

Por otra parte, y como ya se ha mencionado anteriormente, hay muchos investigadores que consideran que el campo de la medida psicológica y educativa requiere hoy en día de la inclusión de los resultados obtenidos de las teorías cognitivas dentro de los modelos de medida. Esta tendencia considera que los resultados de los modelos generales no son fácilmente aplicables si no se cuenta con información específica acerca de la disciplina de la que el constructo proviene.

Con el advenimiento de la investigación en procesamiento de información, en las habilidades de solución de problemas y en el aprendizaje de las disciplinas específicas (Carroll, 1976; Embreston, 1985a y d), se ha ido acumulando gradualmente evidencia respecto a los tipos de habilidades de procesamiento cognitivo y de las estrategias involucradas en el desempeño de distinto tipo de tareas. Más recientemente, los resultados de la investigación cognitiva, que ha dirigido sus esfuerzos a la consideración de los dominios específicos, ricos en conocimiento, de las diferentes asignaturas escolares y al estudio de las pruebas de desempeño en estos dominios (Snow y Lohman, 1989), han enriquecido el acervo de información de la que el desarrollo de modelos de medida puede disponer.

Muchos investigadores consideran que la perspectiva psicométrica tradicional y las teorías cognitivas, tienden a basarse fuertemente en el análisis de constructo de los tests existentes, ya sea mediante análisis factorial o mediante análisis de tareas, y que debieran tomarse en cuenta las teorías del dominio del constructo como una guía en el diseño de tests relevantes para el dominio en cuestión (Messick, 1993). Los modelos de construcción de tests y de medida no pueden ignorar estos resultados. Todos estos trabajos deben tener impacto en el desarrollo de las nuevas teorías de medida.

Podría decirse que la psicología cognitiva, que pretende servir de base a este nuevo tipo de modelos para el diseño y evaluación de los tests no es aún una ciencia consolidada, ni siquiera unificada en su base teórica; sin embargo, la mayor parte de los psicólogos cognitivos parten de una base común, las hipótesis de que los aprendices desarrollan modelos mentales internos para resolver problemas, que estos modelos se desarrollan con base en la experiencia personal y que se usan para resolver las situaciones similares que se encuentran en la vida cotidiana. Estas hipótesis son las que se han tomado como punto de partida en la búsqueda de nuevos modelos de tests.

En la psicología cognitiva la atención se centra en cómo aprenden las personas más que en qué aprenden. En lugar de ver al aprendizaje como la acumulación de conocimientos, se piensa en él como un conjunto de esquemas o estructuras mentales en las que el conocimiento está organizado. El movimiento de reforma al diseño y medida en la educación ha cuestionado, a la luz de las hipótesis antes mencionadas, la utilidad del objetivo instruccional. Actualmente, en lugar de este término, se usa el de resultado de aprendizaje o el de evaluación auténtica buscando con ello una relación más cercana entre la teoría de los tests y la psicología cognitiva.

La psicología cognitiva ha tenido mucho impacto en la comprensión de habilidades cognitivas complejas tales como la solución de problemas, el pensamiento y la inferencia. Esta teoría está siendo empleada cada vez con mayor frecuencia como sustento, tanto de teorías de instrucción, cuanto de teorías de aptitud (Glaser, 1985 y 1993). Es importante que los nuevos modelos de medida contengan teorías matemáticas acerca de los patrones de elección de opciones y que éstas sean compatibles con las características del razonamiento complejo. Partiendo de la hipótesis de que la tareas cognitivas se pueden caracterizar por los procesos que las componen, por las estrategias utilizadas o por las estructuras de conocimiento involucradas en la solución de una tarea, se pueden elaborar nuevos modelos de tests que profundicen más en los aspectos diagnósticos.

Las calificaciones de los tests, por otra parte, deben reflejar las habilidades que se enfatizan en esas teorías cognitivas e incluir maneras de diagnosticar dificultades de aprendizaje haciendo uso de las respuestas incorrectas de los estudiantes.

Es importante destacar que la investigación ha mostrado que las estrategias de solución de problemas cambian en forma dinámica, es decir, dichas estrategias tienden a variar frente a los retos del ambiente (Ginssgburg, 1977) y se ha demostrado que este cambio de estrategias de solución de problemas es más frecuente en estudiantes que no dominan el tema que en alumnos más avanzados (Tatsuoka, Birnbaum y Arnold, 1989). Este tipo de estudios sugiere que debe procederse cautelosamente al elegir la unidad de diagnóstico y que más que los errores frecuentes que cometen los estudiantes en la solución de problemas, son las componentes más generales, tales como las componentes cognitivas relevantes para la instrucción o las que conforman las estructuras cognitivas, las que deben considerarse como unidades de diagnóstico.

Otra dificultad que se enfrenta en el diseño de modelos de medida consiste en que los resultados de las teorías cognitivas establecen que el crecimiento del conocimiento es lento e irregular. Por ello los nuevos modelos de medida no pueden tener pretensiones de que las calificaciones de los tests muestren claramente el crecimiento del aprendizaje.

Un problema crítico en el desarrollo de teorías de aprendizaje y de modelos de medida es el problema de la transferencia de las tareas aprendidas a otro tipo de situaciones. Los nuevos modelos de tests deben intentar abordar este problema sustentándose sobre bases muy sólidas de investigación y diseñando cuidadosamente el tipo de modelos y de parámetros que pueden dar cuenta de este fenómeno.

Por otra parte la enorme diversidad de formas en las que la psicología cognitiva representa el pensamiento de alto nivel se presenta como otro de los retos que se deben vencer en el desarrollo de los modelos de medida. Los términos que se emplean en la psicología rara vez están definidos de manera adecuada como para poder identificar o construir ítems que midan rasgos específicos. No existe una taxonomía validada del comportamiento cognitivo complejo; quizás la única existente es la de Royer, Cisero y Carlo (1993) basada en la teoría cognitiva propuesta por J.R. Anderson (1990) y aún ésta, está lejos de la etapa de implementación. Los desarrollos en la teoría de medida requieren de un esfuerzo mayor y de apoyarse en teorías parciales que sean adecuadas a las características especiales de lo que se desea medir, en los resultados de las investigaciones recientes en las ciencias cognitivas y en las disciplinas específicas relacionadas con los constructos a medir, y dejar a un lado las pretensiones de encontrar un modelo general, que dominaron los desarrollos anteriores.

Las situaciones para las cuales la aplicación de los modelos estándar IRT es menos satisfactoria involucran, ya sea un mejor conocimiento de los procesos psicológicos subyacentes a la respuesta del ítem, o teorías acerca de los procesos que hacen que la medición sea más precisa o más válida. Las características especiales de estas dos situaciones han conducido a nuevos modelos de medida que pretenden generalizar los modelos logísticos y pueden ser fuente de nuevos desarrollos.

Cuando lo que se pretende es desarrollar tests tal vez sea indispensable considerar que la psicología cognitiva y la didáctica de los conocimientos específicos son dos medios que permiten conceptualizar más adecuadamente los rasgos que se quieren medir. Hasta hace poco, la teoría cognitiva había influido poco en el contenido real de las pruebas o en la forma de calificarlas. Incluso hoy sigue imperando la influencia de las teorías IRT o de rasgo latente en el diseño y en la evaluación de los tests. Estas teorías, como ya se ha visto, enfatizan los métodos óptimos de calificación de las respuestas en relación a una habilidad central y dejan de lado la atención en los procesos, estrategias o estructuras de conocimiento que subyacen a la solución de un ítem.

El mayor impacto que la psicología cognitiva puede tener en el desarrollo de tests reside, tal vez, en la conceptualización de la respuesta del examinado al ítem. Esto, a su vez tiene implicaciones directas sobre la validez del test pues permite distinguir por una parte la representación de los constructos, es decir las componentes, estrategias y estructura del conocimiento involucradas al responder un ítem específico y, por otra parte, la utilidad del test como una medida de las diferencias individuales a través de la relación entre las calificaciones del test y otras medidas de dichas diferencias. El cambio en los objetivos de medida implica la búsqueda de una nueva teoría psicométrica en la que puedan introducirse formas más complejas y más generales de calificación.

Los desarrollos en la teoría cognitiva sugieren que las nuevas pruebas de desempeño deben reflejar los aspectos importantes de la capacidad de los sujetos, incluyendo los procesos cognitivos que subyacen a la solución de problemas, el cambio dinámico en las estrategias de los estudiantes y la estructura o representación del conocimiento y de las habilidades (Glaser, 1985). Estos objetivos de medida requieren de una nueva teoría de los tests que es, a la vez, de naturaleza cuantitativa y cualitativa. Las medidas de desempeño deben ser descriptivas e interpretables en términos de los procesos que determinen el desempeño. En este contexto, los objetivos de la información requerida de un test dependerán del uso que pretenda dársele. De ellos podrá obtenerse información útil para propósitos de selección, pero también para el diseño de estrategias remediales de enseñanza.

En contraste con los modelos clásicos y con los modelos IRT, un modelo que permite la introducción de estrategias alternativas en la solución de un ítem, debe proponer explícitamente los procesos mediante los cuales los sujetos alcanzan cierta respuesta. Este tipo de modelos requiere entonces de la existencia de una teoría sólida psicológica o educativa que respalde los procesos involucrados en la solución del ítem.

La incorporación de variables cognitivas al modelo de medida puede hacerse a partir de la separación de los parámetros de los ítems de los de los sujetos en un modelo de tipo logístico. Podría, por ejemplo, reemplazarse el parámetro de dificultad del ítem por un modelo explicativo que represente la complejidad del mismo, o utilizar parámetros de discriminación definidos en términos de las estrategias empleadas o de las estructuras de conocimiento inferidas de la respuesta a los ítems. Es precisamente de esta manera como se ha empezado a abordar en los últimos años el problema de medida en la teoría de los tests.

Los modelos de medida que se reportan actualmente en la literatura toman frecuentemente como punto de partida la teoría IRT, enriqueciéndola con la inclusión de consideraciones diversas en las que entra en juego algún aspecto de las teorías cognitivas.

En los modelos logísticos, la calificación global no proporciona información acerca de la forma en la que el sujeto llega a la respuesta final de un conjunto complejo de actividades. En los modelos más recientes, a través del análisis de los patrones de respuesta a los diferentes reactivos, se puede obtener información relativa a la efectividad de un procedimiento empleado en la solución de un problema, o hacer inferencias acerca del aprendizaje logrado.

Dado que los modelos que han surgido en los últimos treinta años utilizan una amplia gama de consideraciones para diseñar distractores mejores, el formato de opción múltiple ha cobrado nueva vigencia. La posibilidad de dar diferentes pesos a las opciones de ítems politómos proporciona evidencias de una mayor confiabilidad (Haladyna y Sympson, 1988) y las teorías de respuesta a los ítems formuladas por Bock (1972), Samejima (1979), Sympson (1983a, b y c) y Thiessen y Steimberg (1986) que están siendo apenas experimentadas, parecen ser herramientas efectivas de diagnóstico.

Una posibilidad de desarrollar modelos basados en la teoría actual de los tests que permitan incorporar en ellos variables de procesamiento cognitivo y cuantificarlo, se muestra también en el variado trabajo de Embreston (1983a y b;1985a y b; 1993). Sus modelos contienen parámetros que toman en consideración el impacto de las variables de procesamiento cognitivo en las respuestas a los ítems.

En las que podríamos llamar nuevas teorías de los tests, aunque no sean ya tan recientes, se ha intentado modelar el comportamiento cognitivo mediante teorías estadísticas que examinan los patrones de respuesta a los ítems en lugar de centrar la atención en el patrón de respuesta a los ítems en referencia a la calificación global (Frederiksen, Mislevy y Bejar, 1973; Fisher, 1973; Scheiblechner, 1985, Spada y McGaw, 1985; Jannarone, 1986; Embreston y Waxman, 1989; Mislevy y Verhelst, 1990; Mislevy, 1993b). Se está trabajando también con conjuntos de ítems cuya respuesta depende del contexto (Wainer y Kiely, 1987); estos conjuntos de ítems suelen denominarse testlets.

En la literatura reciente referente a la medida en educación, se encuentran así una gran diversidad de modelos para analizar datos en términos de estructuras sugeridas por la teoría cognitiva y la teoría del desarrollo. Entre ellos se pueden mencionar el de Tatsuoka, (1990) quien hace una extensión del modelo IRT usando análisis cognitivo de tareas; aquellos que ya han sido citados con anterioridad como el de Embreston (1985a) y Samejina (1983), quienes hacen la extensión del modelo IRT a partir del análisis de estrategias de respuestas alternativas cuando se pueden observar resultados de subtareas; los de Haertel (1984) y Paulson (1986), quienes construyen sus modelos a partir de las combinaciones de habilidades que cada tarea requiere usando modelos de clase latente, y los de Kelderman (1989), Kelderman y Macready (1990), Mislevy y Verherst (1987 y 1990), Rost (1990) y Yamamoto (1987), que combinan IRT con modelos de clase latente. Se describirán a continuación algunos de estos modelos con un poco más de detalle.

En 1983 Tatsuoka estudió el desempeño en matemáticas en términos de la aplicación de reglas correctas o incorrectas, localizando vectores en un espacio bidimensional en el que la primera dimensión es una estimación del parámetro de habilidad del modelo IRT estándar y, el segundo, es un índice de falta de ajuste a ese modelo. Paulson (1985) analizó datos similares con menos reglas usando modelos de clase latente para relacionar la probabilidad de respuesta correcta a un ítem con las características del mismo y las reglas que los sujetos pueden estar empleando para resolverlo. Yamamoto (1987) combina aspectos de estos dos modelos, tomando subpoblaciones de sujetos que responden de acuerdo a un modelo IRT y de sujetos que considera por sus respuestas no escalables asociados con patrones esperados de respuesta particulares.

Samejima (1983) y Emberston (1985b y c) ofrecen modelos para estrategias alternativas en situaciones en las que los resultados de las subtareas pueden ser observados, además del hecho de que la respuesta sea correcta o incorrecta en un ítem. Mislevy y Verhelst (1987, 1990) describen una familia de modelos IRT de estrategia múltiple que aplican cuando cada sujeto pertenece a una de varias clases mutuamente excluyentes que corresponden a estrategias de solución del ítem, y las respuestas de todos los sujetos en una clase dada se clasifican de acuerdo a un modelo IRT estándar; suponen además que para cada ítem, los parámetros calculados mediante un modelo IRT para cada clase de estrategias se pueden relacionar con características conocidas del ítem a través de alguna teoría psicológica o de una teoría sustantiva. Otro ejemplo de este tipo de modelos es el de Scheiblechner (1985).

Mislevy y Verhelst (1990) tratan situaciones en las que diferentes personas pueden elegir diferentes estrategias a partir de alternativas específicas, pero parten de las competencias globales para encontrar posibilidades de establecer comparaciones significativas entre personas que usan la misma estrategia. Suponen que la estrategia no es directamente observable pero que se puede inferir, aunque sin certidumbre, de los patrones de respuesta o a partir de una base teórica. Suponiendo que la teoría sustantiva permite diferenciar las expectativas acerca de los patrones de respuesta bajo estrategias diferentes y que el sujeto aplica la misma estrategia en todos los ítems, es posible estimar los parámetros del modelo IRT para cada estrategia. Además es posible calcular las probabilidades de que cada sujeto haya empleado cada una de las estrategias alternativas y de estimar la habilidad del sujeto bajo cada estrategia posible. La diferenciación de la verosimilitud de patrones de respuestas bajo estrategias diferentes es la clave de la aplicación de este modelo. El test se debe construir para maximizar diferencias en estrategia, por ejemplo incluyendo ítems que son difíciles bajo una estrategia, pero fáciles bajo otra.

Los modelos mentales, por otra parte, son una herramienta pedagógica y una guía en la solución de los problemas. La posibilidad de inferir el modelo mental que una persona está usando, basándose en el conocimiento de los datos observables, puede servir como guía para el entrenamiento posterior. Si un modelo de Rasch se aplica a los ítems para cada estrategia de solución, se podría inferir qué modelo mental está usando un sujeto. Analizaremos a continuación las características de algunos modelos que intentan este acercamiento con mayor detalle.

El modelo Saltus de Wilson (1984, 1989) es un ejemplo de este tipo de modelos. Éste es un modelo de desarrollo jerárquico que extiende el modelo de Rasch a patrones de desarrollo en los que las capacidades aumentan en etapas discretas, por niveles. El modelo incluye parámetros para cada etapa, así como la consideración de las habilidades de las personas y parámteros de las etapas junto con dificultades para los ítems. Este modelo exhibe, incluso, patrones en los que la probabilidad de éxito en ciertas tareas decrece.

El modelo Saltus de Wilson para el desarrollo jerárquico generaliza el modelo de Rash para ítems dicotómicos introduciendo *H* "niveles de desarrollo". Se supone que un examinado está exactamente en un nivel de desarrollo en el momento en el que hace la prueba, pero que la pertenencia a ese nivel no puede ser observada. Los ítems se clasifican igualmente en H clases. Se supone que para cada nivel de desarrollo vale un modelo de Rasch y que la distancia relativa entre los ítems de la misma clase es la misma, independientemente del nivel de desarrollo. Las dificultades relativas entre los ítems pueden variar de un nivel a otro por cantidades especificadas por los parámetros del modelo. Los parámetros del modelo proporcionan información acerca de la forma en que ciertos ítems se vuelven más fáciles relativos a otros, conforme los estudiantes reconceptualizan el dominio o añaden nuevas reglas a su repertorio, o acerca de la manera en la que ciertos ítems se pueden volver más difíciles cuando los estudiantes pasan de un nivel inferior a otro superior, si antes los respondían correctamente por razones incorrectas.

Este modelo está basado en un modelo psicológico que se apoya en la adquisición de reglas de sustracción en una jerarquía de aprendizaje de Gagné (1968) y en el desarrollo de las habilidades de pensamiento proporcional ilustradas en los datos recogidos por Siegler en el estudio relativo a los problemas de balanzas (Siegler, 1981). Las clases en las que los items se agrupan están basadas en la teoría de desarrollo dentro de la cual se generan los items. Los requisitos que se ponen entonces a la teoría de desarrollo consistente es que describa, por una parte, las clases por las cuales los estudiantes se desarrollan y, por otra, que sea suficientemente específica como para generar ítems que se puedan asociar claramente con predicciones de éxito relativo entre esas clases.

El modelo Saltus puede ser útil en los siguientes casos: Cuando hay individuos en la muestra que tienden a tener distrubuciones multimodales en términos de sus calificaciones globales y cuando los ítems tienden a formar "clusters", es decir, tienden a aparecen agrupados cuando se analizan estadísticamente en términos de porcentajes de ítems acertados.

En este modelo, cada sujeto se caracteriza por dos variables, una cualitativa y otra cuantitativa. El parámetro cualitativo denota la pertenencia a un nivel de desarrollo e indica la naturaleza de la capacidad; el parámetro cuantitativo indica el grado de capacidad, cuyas implicaciones están condicionadas por la pertenencia al nivel. Cada examinado se caracteriza mediante un parámetro de capacidad θ, y también mediante un parámetro de pertenencia a un nivel φ. Si hay H niveles, se tiene que $\varphi_i = (\varphi_{i1}, \varphi_{i2} \ldots \varphi_{iH})$ en el que φ_{ih} vale 1 si el examinado i pertenece al nivel h y 0 si no pertenece a ese nivel.

Aunque ninguno de estos parámetros es observable, las soluciones aproximadas en las primeras demostraciones de Saltus trataban a las estimaciones de la pertenencia a un nivel (basadas en calificaciones globales) como si fueran valores verdaderos de parámetros conocidos, seguidos por simulaciones, "a la medida", para corregir algunos de los efectos de esta sobresimplificación.

A cada ítem se le asocia un parámetro de dificultad β_j y también un indicador de la clase a la que pertenece el ítem b_j, $b_j = (b_{j1}, b_{j2}, \ldots, b_{jH})$ en el que b_{jk} toma el valor de 1 si el ítem j pertenece a la clase H y 0 si no pertenece a esa clase. Es importante aclarar que el parámetro b_j se establece a priori para todos los ítems.

Con toda esta información se construye una matriz de los parámetros del modelo $T = (t_{hk})$. La matriz T es de tamaño $H \times H$. En ella, el elemento t_{hk} indica un efecto en la dificultad de los ítems en la clase k que afecta a los examinados del nivel h. La probabilidad de que un examinado con parámetro de pertenencia a un nivel φ y capacidad θ responda correctamente al ítem j esta dada por

$$P(x_j = 1/\ \theta, \varphi, \beta_j, T) = \prod_h \prod_k \Psi(\theta - \beta j + t_{hk})^{\phi_h b_{jk}}.$$

El doble producto sobre h y sobre k sirve únicamente para elegir el parámetro Saltus para el ítem j que corresponde al nivel de desarrollo del examinado, puesto que el exponente es 1 en ese caso y 0 en los demás. El modelo supone que dados θ y φ las respuestas a los diferentes ítems son independientes. P_h es la probabilidad de obtener un patrón de respuestas x dados θ y la pertenencia al estado h.

El modelo Saltus es una mezcla de modelos IRT, una clase de modelos estudiados por Kelderman (1989), Kelderman y Macready (1990), Mislevy y Verhelst (1990) y Rost (1990). Este tipo de modelo contiene una mezcla de modelos de Rasch, por una parte, que gobiernan las respuestas a los ítems distinguidos por diferentes patrones de dificultad y, por otra parte, modelos de clase latente que se aplican a los examinados.

El modelo Saltus puede considerarse un caso especial del acercamiento de Mislevy y Verhelst. En este último, se toma una mezcla de modelos logísticos pero las diferencias entre los parámetros están condicionadas mediante una teoría psicológica.

La estimación del modelo se hace en dos fases: En la primera, el modelo Saltus IRT proporciona las probabilidades condicionales de respuestas correctas sobre membresía a un nivel y capacidad. En la segunda, se aplica un modelo de poblaciones a las proporciones de la población de examinados en cada nivel y a la distribución de capacidades entre los estados.

Tatsuoka (1983, 1990) considera que las teorías estadísticas de uso frecuente manejan principalmente la calificación global del test y las de los ítems, y que esta información puede considerarse de nivel macro, mientras que los procesos cognitivos pueden considerarse como información de nivel micro y requieren de un análisis más fino del desempeño observable que el que puede obtenerse de la *información de nivel macro.* *Es por ello que en su modelo,* intenta explorar las propiedades de cada uno de estos niveles y las relaciones entre el nivel micro y las tareas invisibles para predecir los resultados de esa relación, que son observables como respuesta a los ítems.

Tatsuoka propone un método que pretende explicar fenómenos de nivel macro derivados de tareas del nivel micro a través del uso de la teoría elemental de gráficas. En este modelo, se lleva a cabo un análisis detallado de los objetivos *y subobjetivos de las tareas y de las diversas trayectorias de solución.* A partir de este análisis se genera un diagrama de flujo con el que se pretende simular el desempeño de los estudiantes en un test, con la ventaja de que el mismo tipo de análisis puede hacerse para pruebas de habilidades o para pruebas de contenido.

Una vez definidos los atributos cognitivos necesarios para llevar a cabo una tarea, o resolver un ítem y las relaciones entre ellos, se define una matriz de incidencia con ellos y se establece una función mediante la cual se comparan las *respuestas de los sujetos con los patrones de respuesta ideales derivados de* la matriz de incidencia. Esta comparación hace posible una inferencia acerca del desempeño de los sujetos en los atributos no observables.

Para evaluar los ítems se pueden usar diferentes fórmulas: Si un sujeto resuelve correctamente todos los atributos involucrados en un ítem, se puede evaluar el ítem como correcto y si comete errores como incorrecto. Se puede, por otra parte, utilizar evaluaciones graduadas aplicando a los datos el modelo de crédito parcial de Masters (1982) o modelos generales de respuesta graduada como el de Samejima (1988). La estimación de la habilidad de los sujetos puede hacerse con el modelo general continuo IRT de Samejima (1995) o con el modelo politómico de Bock (1972).

El modelo lineal logístico de rasgo latente LLTM de Fisher (1973, 1983) pertenece a la familia de modelos de Rasch y relaciona el impacto de las características del estímulo con la dificultad de procesamiento de las componentes. Su aplicación requiere que las características del estímulo se califiquen y, posteriormente, se usen para modelar la dificultad del ítem. Este modelo fue el primero que intentó relacionar la psicología cognitiva con la teoría de los tests.

Antes de usar el modelo, se califican los ítems en M factores que representan su complejidad en los estímulos que influyen en el procesamiento de la información y, puesto que únicamente se pueden observar los resultados finales de la aplicación del ítem como un todo, se supone que las dificultades componentes se combinan aditivamente.

En el modelo LLTM se restringen los parámetros de tal manera que la dificultad del ítem esté dada por una combinación lineal del producto de los pesos de dificultad de los factores por el parámetro de complejidad del factor más una constante de normalización. Este modelo es análogo a una regresión de las dificultades de los ítems sobre los factores de complejidad. En realidad, el modelo LLTM es una restricción del modelo de Rasch en la que las dificultades de loa ítems se modelan como funciones lineales de los pesos de las componentes.

Las aplicaciones de este modelo permiten enlazar las aptitudes con las variables asociadas a los procesos cognitivos. Las teorías cognitivas se operacionalizan mediante los valores asociados a la complejidad cognitiva. A partir de una prueba de bondad de ajuste, basada en la verosimilitud de los datos obtenidos del modelo, se puede poner en evidencia si los examinados utilizan los procesos y estrategias que han sido tomados en consideración en el modelo. Por otra parte, los valores de los factores de complejidad son análogos a los pesos asociados a cada variable cognitiva para explicar la dificultad del ítem. Por esta razón, el modelo puede explicar la representación del constructo del test a través del impacto de las variables cognitivas en las respuestas a los ítems.

Una aplicación de este modelo es la comparación de teorías cognitivas alternativas en cuanto a su poder explicativo. Dicha comparación se define y se lleva a cabo mediante el análisis de qué tan adecuadamente los parámetros describen las componentes cognitivas del modelo. Este modelo puede emplearse, además, en el desarrollo o selección de ítems por su complejidad cognitiva, lo que es útil para controlar la representación del constructo en el test. Por último, las habilidades se pueden interpretar en términos del potencial de las personas para procesar adecuadamente los estímulos que aparecen en los ítems. Todas estas aplicaciones requieren un apoyo fuerte de una teoría, externa al modelo de medida, acerca de la forma en la que la complejidad de los estímulos influye en el procesamiento de la información.

El modelo multicomponente de rasgo latente MLTM de Whitely (1980) relaciona los resultados de las componentes cognitivas de las tareas con las calificaciones de los tests. La idea central en este modelo es que la probabilidad de éxito al responder un ítem es una función multiplicativa de las probabilidades de éxito en las subtareas que dicho ítem involucra. Su aplicación requiere tanto de un resultado para cada una de las subtareas, cuanto de los datos del ítem. Todas las instancias de una componente siguen supuestamente un modelo de Rasch, lo que implica que cada subtarea tiene su propio parámetro de dificultad, dependiente de la instancia en la que aparece; esto significa que la dificultad de construir una regla puede diferir dependiendo del ítem.

El modelo MLTM estima tanto las habilidades de los sujetos cuanto las dificultades del ítem en cada componente constitutiva del modelo. La dificultad de procesamiento se estima directamente de los resultados obtenidos para las tareas y, aunque se requiere de una teoría que de cuenta del procesamiento de las componentes, esta teoría no necesita especificar la relación de las características de los estímulos con la complejidad de procesamiento de la información.

Para estimar los parámetros en este modelo, se requieren dos tipos de datos a) las respuestas al ítem estándar y b) las respuestas a las subtareas necesarias para la solución del ítem, para representar las componentes. Para esto último, se construyen otros ítems que sólo requieren de una respuesta para cada componente. Las probabilidades de responder correctamente a los componentes están dadas por un modelo logístico de un parámetro en términos de la habilidad de la persona en esa componente y de la dificultad del ítem en esa misma componente. No es necesario calcular los parámetros utilizando máxima verosimilitud, ya que los parámetros se pueden estimar a partir de las subtareas únicamente (Maris, 1995).

Este modelo se puede utilizar para elegir las componentes de procesamiento cognitivo que se desean reflejar en un test o para seleccionar ítems que sean tales que ciertas componentes específicas tengan una contribución mayor en las diferencias individuales.

Los modelos LLTM y MLTM tienen algunas características en común. Ambos son modelos de componentes que permiten el análisis de los ítems en términos de subtareas; ninguno de estos dos modelos supone parámetros de peso para las componentes. Estos modelos difieren, sin embargo en que el primero introduce un solo parámetro por componente, mientras que el segundo tiene tantos como lo requieran los ítems compuestos. El modelo LLTM descompone las dificultades de los ítems en la suma de sus contribuciones por componente, mientras que el modelo MLTM descompone las probabilidades de respuesta en el producto de las probabilidades de respuesta a las componentes.

El modelo de componente general de rasgo latente GLTM de Embreston (1985a, 1989) incorpora los modelos LLTM y el MLTM en un nuevo modelo, por lo que su aplicación requiere de los datos necesarios en ambos. En este modelo, como en el MLTM, se incluye una relación multiplicativa entre los resultados de las subtareas, pero se postula además una relación lineal entre los factores de complejidad y los componentes de los parámetros de dificultad, como en el LLTM.

Este modelo evalúa dos tipos de variables cognitivas: la relación entre las tareas componentes de un ítem y el ítem total y la relación entre las características de los estímulos de los ítems con los resultados obtenidos para las componentes. El modelo GLTM tiene las mismas ventajas que el MTLM pero además es capaz de diagnosticar cómo las características de los estímulos controlan la dificultad de las componentes y por ello ofrece una herramienta más completa para el diseño de pruebas.

El modelo de Butter y sus colaboradores (1998) toma del modelo LLTM la descomposición aditiva de las dificultades introduciendo, además, parámetros específicos de dificultad para las componentes como en el modelo MTLM. Pero, a diferencia de ellos, introduce parámetros de peso, un peso diferente para cada componente. En él, los elementos de la matriz de las dificultades de las componentes de los ítems deben estimarse. Este modelo se denomina MIRID, que significa modelo con restricciones internas sobre la dificultad de los ítems.

Otro tipo de modelos que se han desarrollado paralelamente a los anteriores puede ejemplificarse con el trabajo de Mislevy (1993b). Este autor utiliza redes de inferencia en las que se representa gráficamente el cúmulo y las conexiones entre las diferentes tareas que conforman un comportamiento cognitivo complejo. Estas redes pueden ponerse en correspondencia con una base estadística que refleja el razonamiento acerca de la causalidad de los factores que se pueden observar. Dentro de las redes pueden incluirse ítems de opción múltiple o de respuesta construida, pues cada uno de ellos se ajusta a un cierto tipo de inferencia. En el modelo de Mislevy se describe así un modelo causal de razonamiento acerca de las observaciones y una teoría estadística apropiada que se puede usar para modelar el comportamiento del estudiante que está aprendiendo.

Un aspecto esencial de la investigación cognitiva es la hipótesis de que los niveles de habilidad cambian diferencialmente y que las diferencias individuales tienen relaciones significativas con otras medidas. Un modelo reciente en esta dirección es el modelo multidimensional de Rasch de aprendizaje y cambio MRMCC de Embreston (1985b) que requiere de la medición de dos o más habilidades. Conceptualmente se distingue dentro del modelo la habilidad inicial de la modificabilidad. La medida tradicional de habilidad se toma como la inicial en la primera acción, mientras que la modificabilidad se considera como el cambio entre dos medidas sucesivas.

La investigación en medida ha incorporado también una nueva dirección que consiste en el análisis de los efectos del cambio en el contexto cognitivo en una sesión de examen. Por ejemplo, se evalúan los efectos de proporcionar ayudas o del nivel de construcción sobre los niveles de habilidad (Campione, et al., 1985, Lohman, 1989) en un intento de separar la habilidad actual con el nivel potencial o zona de desarrollo próximo (Vigosky, 1978).

Existen, por último, modelos que parten directamente de una perspectiva basada en alguna noción de la psicología cognitiva. Un ejemplo de ellos es el modelo de Marshal y sus colaboradores (1989) que parte de una perspectiva basada en la noción de esquema. En él se parte de una definición precisa de lo que los autores entienden por esquema de conocimiento en una cierta área de conocimiento específico. Su definición requiere del establecimiento claro del contenido del esquema y de su estructura, es decir las relaciones entre los contenidos. Lo que se pretende encontrar es cuál es el contenido del esquema que maneja cada sujeto en particular, la forma en la que dichos contenidos están relacionados y qué tan sólidas son las relaciones.

Para determinar los esquemas primarios de un dominio se parte de aquello que los investigadores esperarían que el estudiante fuera capaz de hacer si dominara la materia. De este modo se identifican los nodos conceptuales del esquema ideal junto con sus relaciones. El conocimiento de un individuo se representa como una gráfica o red grande que contiene redes más pequeñas que corresponden a los esquemas individuales. Es posible que en esos esquemas exista más de una relación o ligadura entre dos nodos y que las trayectorias entre los diferentes nodos no sean necesariamente directas. Los esquemas primarios se diseñan con base en una teoría cognitiva específica y con la ayuda de expertos en el dominio de conocimiento dentro del que se encuentran los conceptos a medir.

Para establecer la calidad del esquema de un individuo se examina su conectividad, es decir, se intenta encontrar hasta qué grado sus nodos constituyentes están conectados unos con otros. Se considera que cada elemento de información tiene el potencial de activar otra información con la que está relacionada. Si un individuo puede recordar una pieza particular de información es importante reconocer a cuántas otras puede acceder y procesar automáticamente. A mayor conectividad del esquema, mayor cantidad de información activada y de esta manera la conectividad del esquema puede asociarse con el fenómeno conocido en psicología como "spreading activation". El grado de activación puede medirse usando algún modelo como el ACT de Anderson (1990) o la teoría armónica (Smolensky, 1986); puede asimismo estimarse el número de ligaduras de la gráfica que representa la red (Marshal, 1990 y 1993). Una vez que se ha hecho el análisis de la conectividad de la gráfica, se hace otro análisis que se considera igualmente importante, de los tipos de conocimiento que están relacionados.

Los ítems, en estos modelos, proveen evaluaciones a varios niveles de la red. Pueden probar un nodo individual si se dirigen hacia la información del mismo o pueden probar distintos tipos de relaciones. La respuesta del estudiante es la base para determinar la presencia o ausencia del nodo correspondiente en la red o para inferir la conectividad. Cuando se agregan los resultados particulares de los ítems con la calificación global, los ítems deben tratarse diferencialmente para que reflejen el tipo y cantidad de información que poseen.

Una limitación de este tipo de modelos consiste en que cuando el nivel de complejidad aumenta, la dificultad del modelo crece mucho. Se requiere incluir un gran número de nodos y de relaciones para dar cuanta del estado de conocimiento de los sujetos, lo que hace sumamente difícil el manejo de la información. Por otra parte, la dificultad en el diseño de los ítems aumenta sustancialmente, puesto que éstos deben representar un gran número de dimensiones del esquema de conocimiento.

Es cierto que con las posibilidades que ofrece actualmente la computación, las dificultades en el manejo de la información pueden ser superadas. El diseño de la red cognitiva, sin embargo, requiere de investigación profunda en el dominio de conocimiento en el que se inscribe el test. Esta investigación es ardua, requiere de una inversión fuerte de tiempo y de un esfuerzo considerable. Los resultados de los que hoy en día se dispone en este sentido son insuficientes para lograr los exigencias que estos modelos plantean.

Una posible taxonomía de modelos

Dada la gran cantidad de modelos de medida que se han desarrollado en los últimos veinte años, se hace necesaria una clasificación de acuerdo a criterios específicos. Thiessen y Steimberg (1986), por ejemplo, proponen una clasificación o taxonomía que puede considerarse válida hasta hoy en día.

Los modelos considerados parecen, a primera vista, muy diferentes unos de otros; sin embargo, pueden organizarse como miembros de tres clases diferentes en las que los modelos pueden distinguirse ya sea por sus supuestos o por las restricciones en sus parámetros: Modelos de diferencia, en los que las respuestas se consideran ordenadas, modelos divididos por el total que se pueden emplear para respuestas ordenadas o nominales y modelos aditivos por la izquierda que se utilizan para ítems de opción múltiple en los que es posible introducir la posibilidad de adivinar la respuesta. El propósito de una taxonomía de esta naturaleza es hacer notar la relación existente entre los modelos existentes. Esta taxonomía muestra que modelos distintos se pueden considerar como generalizaciones o casos específicos uno de otro.

La mayor parte de los modelos de medida se desprenden de los modelos binarios, entre los cuales se encuentran el de ojiva normal, los logísticos de uno y dos parámetros y los de splines.

Los llamados modelos de diferencia toman los modelos de ojiva normal o los logísticos como punto de partida para encontrar la probabilidad de que los sujetos respondan a un ítem en un nivel o categoría específica que se considera con relación a otras categorías superiores. De esta manera, el problema original se divide en un conjunto de subproblemas más manejables. Estos modelos son adecuados generalmente cuando las respuestas a los ítems se pueden organizar o clasificar por niveles claramente definidos. El modelo más típico de diferencia es el de respuesta graduada de Samejima, aunque se puede incluir también en esta clase el de Wilson, los de Huynh y Ferrara (1994), el de Yen (1993) y el de Tatsuoka (1990).

Los modelos divididos por el total consideran también categorías ordenadas de respuesta, pero, en ellos, se considera que la respuesta a un ítem está compuesta por m categorías y que puede, por ello, dividirse en $m - 1$ ítems hipotéticos binarios. Las curvas correspondientes a estos modelos se encuentran calculando la probabilidad de alcanzar la categoría k dado que la respuesta está en la categoría $k - 1$ o k. El total está dado por la suma de las categorías en las que el sujeto contesta y el resultado de la función de probabilidad es una exponencial dividida entre la suma de todas las exponenciales que pueden aparecer en el numerador. Entre los modelos divididos por el total se podrían incluir los de crédito parcial de Masters, el de Andersen, el de escala graduada de Andrich, el de Masters y Wright, el de Muraki, el nominal de Bock y el de Butter y De Boeck que se han mencionado ya anteriormente.

En los modelos de suma por la izquierda se pueden identificar dos funciones, una creciente y generalmente logística que representa la probabilidad de obtener una respuesta correcta al ítem y otra que representa la probabilidad de adivinar la respuesta. Entre estos modelos se pueden incluir el logístico de tres parámetros y el de dos parámetros de Choppin (1983).

Existen además, modelos que pueden considerarse de suma por la izquierda y divididos por el total como el de opción múltiple de Samejima, el de Sympson y el de Thiessen y Steinberg.

Además de este tipo de clasificación podría incluirse una adicional que considerara la posibilidad de relacionar los parámetros del modelo con alguna teoría cognitiva. Este es el caso de los modelos de Wilson, Whitely, Flamange, Paulson, Haertel, Embreston, Campione y Marshal.

Se ha mostrado hasta aquí como en los últimos veinte años se han desarrollado una gran cantidad de modelos de medida. La teoría de los tests se encuentra en una etapa de desarrollo vigoroso. Dada la complejidad involucrada en la medida del conocimiento, ha sido necesario que los modelos sean cada vez más específicos en relación con las tareas que se desean medir y más complejos en cuanto a las técnicas matemáticas involucradas en el manejo y análisis de la información.

Los modelos matemáticos empleados en todos los modelos descritos descansan fuertemente en teorías estadísticas, aun cuando en ocasiones hacen uso de otras teorías matemáticas como la teoría de gráficas o la de redes. El modelo que se presenta en este trabajo intenta dar un paso en otra dirección al tomar como punto de partida un modelo determinista que se inscribe dentro de la rama del Análisis Matemático. En este modelo se utilizan ecuaciones diferenciales no lineales para describir las curvas de respuesta a los ítems y en el análisis de la información que se obtiene de ellas. El modelo que se desarrolla en el siguiente capítulo de este trabajo puede considerarse una extensión del modelo logístico de dos parámetros, pero, a diferencia de éste, introduce la posibilidad de dar sentido tanto a las respuestas correctas a los ítems cuanto a las incorrectas y de relacionarse con alguna teoría específica cognitiva que haga posible la interpretación de los resultados que se obtienen. ◆

Referencias

ABRAHAMOWITZ, M. y RAMSAY, J.O. (1992). Multicategorical spline model for item response theory. Psychometrika, 57, 5-27.

ALBANESE, M.A. (1992). Type K items. Educational Measurement: Issues and Practices, 12, 117-128.

ANDERSEN, E.B. (1966). Den diskrete malingsmodel of endelig orden med anvendelse pa et socialpsykologisk materiale. Kobenhavn, Denmark: Statens Tryningskontor.

ANDERSEN, E.B. (1985). Estimating latent correlations between repeated testings. Psychometrika, 50, 3-16.

ANDERSON, J.R. (1990). The adaptive character of thought. Hillsdale, N.J., Lawrence Erlbaum.

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

ANDRICH, D. (1995). Distinctive and incompatible properties of two common classes of IRT models for graded responses. Applied Psychological Measurement, 19, 101-119.

BENNETT, R.E., ROCK, D.A., BRAUN, H.I., FRYE, D., SPOHRER, J.C., y SOLOWAY, E. (1990). The relationship of constrained free-response to multiple-choice and open-ended items. Applied Psychological Measurement, 14, 151-162.

BIRNBAUM, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M.R. Novick, Statistical theories of mental test scores (pp. 17-20). Reading, MA: Addison Wesley.

BIRNBAUM, A. y TATSUOKA, K.K. (1987). Open ended versus multiple choice response formats- It does make a difference for diagnostic purposes. Applied Psychological Measurement, 11, 329-341.

BOCK, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. Psychometrika, 37, 29-51.

BORMOUTH, J.R. (1970) On a theory of achievement test items. Chicago, Universiy of Chicago Press.

BUTTER, R., DE BROECK, P. y VERHELST N, (1998). An Item Response Model with Internal Restrictions on Item Difficulty. Psychometrika, 63, 47-63.

CAMPIONE, J.C., BROWN, A.L., FERRARA, R.A., JONES, R. S., Y STEINBERG, E. (1985). Breakdown in flexible use of information: Inteligence related differences in transfer following equivalent learning performance. Intelligence, 9, 323-344.

CARROLL, J. B. (1976). Psychometric tests as cognitve tasks: A new structure of intellect. In L. B. Resnick (Ed.), The nature of intelligence (pp. 27-56). Hillsdale, NJ: Lawrence Erlbaum Associates.

CHOPPIN, B. (1983). A two-parameter latent trait model. (CSE Report No. 197). Los Angeles, CA: University of California, Center for the Study of Evaluation, Graduate School of Education.

CRONBACH, L.J. (1951): Coefficient alpha and the internal structure of tests. Psycometrika, 16, 297-334.

DOWNING, S.M. (1992). True-false and alaternate-choice item formats: A review of research. Educational Measurement: Issues and Practices, 11, 27-30.

EMBRESTON, S.E. (1983a). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93, 179-197.

EMBRESTON, S.E. (1983b) An incremental fit index for the linear logistic latent trait model. Paper presented at the annual meeting of the Psychometric Society, Los Angeles, CA.

EMBRESTON, S.E. (1985a). Psychometric models for learning and cognitive processes. En Test design: Developments in psychology and psychometrics (pp.195-218) Orlando, FL. Academic Press.

EMBRESTON, S.E. (1985b). Multicomponent latent trait models for test design. En S.E. Embreston (Ed.) Test design: Developments in psychology and psychometrics (pp.195-218) Orlando, FL. Academic Press.

EMBRETSON, S.E. (1985c). (Ed.), Test design Developments in psychology and psychometrics (pp. 195-218). Orlando, FL: Academic Press.

EMBRESTON, S. E. (1985d). Introduction to the problem of test design. In S. E. Embreston (Ed.) Test design: Developments in psychology and psychometrics. New york, Academic Press.

EMBRESTON, S.E. (1993). Psychometric models for learning and cognitive processes. En Frederiksen, N., Mislevy, R.J., & Bejar, Y. (Eds.). Test theory for a new generation of tests. Hillsdale, NJ: Lawrence Erlbaum Associates.

EMBRESTON, S.E. y WAXMAN, M. (1989). Models for processing and individual differences in spatial folding. Unpublished manuscript. Citado en Embreston, S.E. (1993).

FELDT, L. S. (1965): "The aproximate sampling distribution of Kuder-Richardson reliability coefficient twenty", Psycometrika, 30, 357-370.

FISHER, G. H. (1973). The linear logistic test model as an instrument in educational research. Acta Psychologica, 37, 359 –374.

FISHER, G. H. (1983). Logistic latent trait models with linear constraints. Psychometrika 48, 3-26.

FISHER, G.H. (1995). Some neglected problems in IRT. Psychometrika, 60,459-487.

FLANAGAN, J.C. , (1937). A note on calculating the standard error of measurement and reliability coefficients with the test scoring machine. Journal of Applied Psychology, 23, 529.

FREDERIKSEN, N. (1984) The real test bias: influences on testing on teaching and learning. American Psychologist 39, 193-202.

FREDERIKSEN, MILEVY, R.J. y BEJAR, (1993). Test theory for a new generation of tests. Hillsdale, N.J. Lawrence Erlbaum.

FRISBIE, D.A. (1992). The status of multiple true-false testing. Educational Measurement: Issues and Practices, 5, 21-26..

GAGNÉ, R.M. (1968). Learning hyerarchies. Educational Psychologist, 6, 1-9.

GINSSBURG, H. (1977). Children's arithmetic. The learning process. New York: Van Nostrand.

GLASER, R. (1985). The integration of instruction and testing. Paper presented at the ETS Invitational Conference on the Redesign of testing for the 21th century, New york.

GLASER, R. (1993, April). Criterion-referenced tests: Origins and unfinished business. In Criterion-referenced measurement: A 30-year retrospective. Symposium conducted at the annual meeting of the american Educational Research Association, Atlanta.

GLONEK, G. F. V. Y MCCULLAGH, P. (1995) Multivariate Logistic Models. J. R. Statist. Soc. B, 57, pp 533-546.

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

HAERTEL, E.H. (1984). An application of latent class models to assessment data. Applied Psychological Measurement 8, 336-346.

HALADYNA, T.M. (1994). Developing and validating multiple choice test items. Hillsdale, N.J. Lawrence Erlbaum.

HALADYNA T.M. y DOWNING, S.M. (1989). A taxonomy of multiple choice item writing rules. Applied Meassurement in Education 1, 51-78.

HALADYNA, T.M., y DOWNING, S.M. (1997). How many options is enough for a multiple-choice test item. Educational & Psychological Measurement.

HALADYNA, T.M., y SYMPSON, J.B. (1988, April). Empirically based polychotomous scoring of multiple-choice test items: A review. In New Development in Polychotomos Scoring. Symposium conducted at the annual meeting of the American Educational Research Association, New Orleans.

HUYNH , H. (1996). Decomposition of a Rash partial credit item into independent binary and indecomposable trinary items. Psychometrika 61, 31-39.

HUYNH , H. Y FERRARA, S. F. (1994). A comparison of equal percentile and partial credit equatings for performance-based assessments comprised of free response items. Journal of Educational Measurement, 31, 125-141.

JANNARONE, R. J. (1986). Conjunctive item response theory kernels. Psychometrika, 51, 357-373.

KELDERMAN, H. (1989). Item bias detection using loglinear IRT. Psychometrika 54, 681-697.

KELDERMAN, H. y MACREADY, G.B. (1990). The use of loglinear models for assessing differntial item functioning across manifest and latent examinee groups. Journal of Educational Measurement, 27, 307-327.

KRISTOF, W. (1963). The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. Psychometrika 28, 221-238.

KUDER, G.F., y RICHARDSON, M.W. (1937). The theory of the estimation of test reliability. Psychometrika 2, 151-160.

LAWLEY, D. N. (1943). On problems connected with item selection and test construction. Proceedings of the Royal Society of Edinburgh, 62-A (Pt. I), 74-82.

LOHMAN, D.F. (1989). Human intelligence: An introduction to advances in theory and research. Review of Educational Research, 59, 333-373.

LORD, F.M. (1952). A theory of test scores. Psychometric Monograph, No. 7.

LORD, F.M., y NOVICK, M.R. (1968). Statistical theories of mental test scores. Reading, MA. Addison Wesley.

MARIS, E. (1995). Psychometric latent response models. Psychometrika 60, 523-547.

MARSHAL, S.P. (1990, April). What students learn (and remember) from word problem instruction. In Penetrating to the Mathematical Structure of Word Problems. Symposium conducted at the annual meeting of the American Educational Research Association, Boston.

MARSHAL, S. P. (1993). Assesing schema knowledge. En N. Frederiksen (Ed.) Test theory for a new generation of tests. Hillsdale, Lawrence Erlbaum.

MARSHAL, S. P, BARTHULI, K. E., BREWER y M. A. ROSE, F. E. (1989). Story problem solver: A schema- based system of instruction (Tech. Rep. 89-01). Center for Research in Mathematics and Science Education, San Diego, CA.

MASTERS, G.N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

MESSIK, S. (1984). The psychology of educational measurement. Journal of Educational Measurement, 21, 215-237.

MESSIK, S. (1993). Trait equivalence as construct validity of score interpretation across multiple methods of measurement. In R.E. Bennett & W.C. Ward (Eds.), Constrction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment (pp. 60-73). Hillsdale, N.J. Lawrence Erllbaum Associates.

MISLEVY, R.J. (1993b). Test theory reconsidered. En Frederiksen, Mislevy, R.J. y Bejar, (1993). Test theory for a new generation of tests. Hillsdale, N.J. Lawrence Erlbaum.

MISLEVY , R. J.y VERHELST , N. (1987). Modeling item responses when different subjects employ different solution strategies (research Report RR-87-47-ONR). Princeton, NJ. Educational Testing Service.

MISLEVY , R. J. y VERHELST , N. (1990). Modeling item responses when different subjects employ different solution strategies. Psychometrika, 55, 195-215.

MURAKI, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement, 16, 159-176.

PAULSON, J. (1985). Latent class representation of systematic patterns in test responses (ONR Technical Report) Portland, OR: Portland State University.

PAULSON, J.A. (1986). Latent class representation of systematic patterns in test responses (Technical Report ONR-1). Portland, OR. Portland State University, Psychology Department.

RASH G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Denmarks Paedagogiske Institut.

RICHARDSON, M. W. (1936). The relationship between difficulty and the differential validity of a test. Psychometrika 1, 33-49.

ROST, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. Applied Psychological Measurement, 14, 271-282.

ROYER, J.M., CISERO, C.A., & CARLO, M.S. (1993). Techniques and procedures for assessing cognitive skills. Review of Educational Research, 63, 201-243.

RULON, P.J. (1939). A simpliefied procedure for determining the reliability of a test by split-halves. Harvard Educational Review, 9, 99-103.

SAMEJIMA, F. (1972). A general model for free- response data. Psychometric Monograph, No. 18.

SAMEJIMA, F. (1979) A new family of models for the multiple-choice item. (Office of Naval Research Report 79-4) Knoxville, TN, University of Tennessee.

SAMEJIMA, F. (1983). A latent trait model from differential strategies in cognitive processes(Tecnical Report ONR/RR-83-1). Knoxville, TN: University of Tennessee.

SAMEJIMA, F. (1988). Advancement of latent trait theory 9ONR Final Report). Knoxville, TN: University of Tennessee.

SAMEJIMA, F. (1995). A cognitive diagnosis method using latent trait models: competency space approach and its relationship with DiBello and Stout's unified cognitive/psychometric diagnosis model. In P.D. Nichols, S.E. Chipman & R.L. Brennan (Eds.), Cognitively diagnostic assessment. Hillsdale, NJ: Lawrence Erlbaum.

SCHEIBELCHNER,H. (1985) Psychometric models for speed-test construction: The linear exponential model. En S. Emberston (Ed.) Test Design: Developements in Psychology and Psychometrics (pp. 219-244). New York. Academic Press.

SHEPARD, L.A. (1993). The place of testing reform in educational reform-A reply to Cizek. Educational Researcher, 22, 10-13.

SIEGLER, R.S. (1981). Developmental sequences within and between concepts. Monograph of the Society for Research in Child Development, 46.

SMOLENSKY, P. (1986) Information processing in dynamical systems: Foundations of harmony theory. En D.E. Rummelhart, I.C. L. McClelland y The P.D.P. Research Group (Eds.) Parallel distributed processing. Vol. I: Foundations. Pp 194-281. Cambridge MA, MIT Press.

SNOW, R.E. (1993). Construct validity and constructed-response tests. In R.E. Bennett & W.C. Ward (Eds.), Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment (pp. 45-60). Hillsdale, ÑJ: Lawrence Erlbaum Associates.

SNOW, R.E. Y LOHMAN, D.F. (1989). Implications of cognitive psychology for educational measurement. In R.L. Linn (Ed.) Educational measurement (3rd ed. pp 262-331). New York. MacMillan.

SPADA H. y McGAW B. (1985) The assessment of learning effects with linear logistic tests models. En S. Inherston (Ed.) Test design: New directions in psychology and psychometrics, pp. 169-193.

SPEARMAN, C. (1904). The proof and measurement of association between two things, American Journal of Psychology, 15, 72-101.

SPEARMAN, C. (1907). *Demostration of formulae for true measurement of* correlation, American Journal of Psychology, 18, 161-169.

SPEARMAN, C. (1913). Correlations of sums and differences, British Journal of Psychology, 5, 417-426.

STEVENS, S.S. (1946). On the theory of scles in measurement. Science, 103,677-680.

SYMPSON, J.B. (1983a). A new IRT model for calibrating multiple choice items. Paper presented at the annual meeting of the Psychometric Society, Los Angeles, CA

SYMPSON, J.B. (1983b, April). Extracting information from wrong answers in computerized adaptive testing. In New Developments in Computerized Adaptive Testing. Symposium conducted at the annual meeting of the American Psychological Association, Washington, D.C.

SYMPSON, J.B. (1983, August). A new item response theory model for calibrating multiple-choice items. Paper presented at the annual meeting of the Psychometric Society, Los Angeles.

TATSUOKA, K. K. (1983). Rule-Space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20, 345-354.

TATSUOKA, K.K. (1990). Toward an integraation of item response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M.G. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition (pp. (453-488). Hillsdale, NJ: Lawrence Erlbaum Associates

TATSUOKA, K.K., BIRENBAUM, M., y ARNOLD, J. (1989). On the stability of students' rules of operation for solving arithmetic problems. Journal of Educational Measurement, 26, 351-361.

THISSEN, D. y STEINBERG, L. (1984). A response model for multiple- choice items. Psychometrika, 49, 501-519.

THISSEN, D. y STEINBERG, L. (1986). A taxomomy of item response models. Pshycometrika, 51, 566-577.

THISSEN, D., STEIMBERG, L. y FIZPATRIK, A.R. (1989). Multiple- choice models: The distractors are also part of the item. Journal of Educational Measurement, 26, 161-176.

THOMPSON, P.W. (1982). Were lions to spead, we wouldn't understand. Journal of Mathematical Behavior, 3, 147-165.

TORGERSON, W.S. (1958). Theory and methods of scaling. New York. John Wiley.

TUCKER, L.R. (1946). Maximum validity of a test with equivalent items. Psychometrika, 11, 1-13.

VIGOSTKY, L.S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

WAINER, H., & KIELY, G. (1987). Item clusters and computerized adaptive testing: A case for testlets. Journal of Educational Measurement, 24, 185-202.

WHITELY, S.E. (1980). Multicomponent latent trait models for ability tests. Psychometrika.45, 479-494.

WILSON, M. (1984). A Psychometric model of hyerarchical development. Unpublished doctoral dissertation, University of Chicago.

WILSON, M. (1989). Saltus: A psychometric model of discontinuity in cognitive development. Psychological Bulletin, 105 (2), 276-289.

YAMAMOTO, K. (1987). A hybrid model for item responses. Unpublished dissertation. University of Illinois.

YEN, W. M. (1993). Scaling performance assessments: Strategies for managing local item independence. Journal of Educational Measurement, 30, 187- 213.

CAPÍTULO II

LA TEORÍA DE RESPUESTA A LOS ÍTEMS DESDE UNA PERSPECTIVA DINÁMICA

Los modelos logísticos de los que se ha hablado en los capítulos anteriores de este trabajo parten de la consideración de la probabilidad de que un sujeto acierte en su respuesta a un ítem y modelan, como hemos visto, esa probabilidad en términos de una función logística. Estos modelos se han empleado abundantemente durante los últimos diez años y los resultados de su uso han sido satisfactorios, particularmente en aquellas aplicaciones destinadas a la evaluación a gran escala, a la selección y a la clasificación de estudiantes. Sin embargo, su utilidad es menor cuando cuando la meta de la aplicación es el diagnóstico de competencias específicas de los sujetos o cuando se intenta generar conocimiento para mejorar dichas competencias. Los modelos logísticos pueden usarse satisfactoriamente para predecir el éxito o el fracaso de los individuos ante una situación educativa determinada (Glaser, et al.; 1987), aunque es imposible, a partir de los resultados de la aplicación del modelo, comprender sus causas.

En la mayor parte de las aplicaciones, los modelos logísticos de uno, dos y tres parámetros se han utilizado para caracterizar a los sujetos en términos de sus respuestas correctas. Por ello, aún cuando los ítems se diseñen con varias alternativas de respuesta, el ajuste de los parámetros se hace tomando al ítem como si fuera dicotómico, es decir, codificando si la respuesta es correcta o incorrecta.

Los modelos logísticos son particularmente útiles para el diseño de pruebas en aquellos casos en los que se desea obtener información global del ítem, lo cual es muy conveniente en el caso de creación de tests o de bases de reactivos para estructuración de exámenes; cuando se desea obtener una perspectiva general del desempeño de los sujetos en un test determinado; y en situaciones en las que todos los sujetos emplean una misma estrategia en la solución de los ítems. En este último caso puede considerarse que los parámetros de los modelos reflejan el número o la complejidad de las operaciones necesarias para resolver el ítem (Fisher, 1973 y 1995).

Cuando se desea obtener un conocimiento más detallado, ya sea de la estrategia que utilizan los sujetos en la solución del ítem, del tipo de dificultades a las que suelen enfrentarse o del tipo de errores que es común detectar, los modelos que se han manejado hasta ahora son menos satisfactorios. Por ello, se hace necesario desarrollar modelos alternativos en los que los parámetros puedan proporcionar mayor información acerca de la competencia de los sujetos o de las estrategias que siguen al resolver los problemas planteados, a partir del reconocimiento de la alternativa de respuesta elegida por cada uno de ellos.

En la construcción de este nuevo tipo de modelos alternativos se pueden considerar las diferentes estrategias de solución al ítem, los errores comunes que cometen los sujetos en la solución de un ítem dado o alguna teoría acerca de la estructuración del conocimiento para diseñar los ítems. La inclusión de este tipo de consideraciones en el diseño de los tests y en los modelos de medida puede proporcionar información útil para el diagnóstico, para la clasificación y para el diseño de distractores adecuados en la construcción de los ítems.

Las posibilidades de diseño de tests con base en esos modelos son de enorme interés en el campo de la educación. A partir de las respuestas a los ítems así diseñados se podría obtener información acerca de la forma en la que diferentes sujetos resuelven un ítem particular, información que resulta más valiosa que el simple conocimiento del número de ítems que los sujetos pueden resolver correctamente. Por otra parte, el diseño de distractores adecuados, proporciona herramientas que permiten que los tests resulten más adecuados para lo que se pretende medir.

La formulación de un modelo de esta naturaleza requiere de la explicitación de los procesos por los cuales los sujetos llegan a distintas respuestas. Por ello, además del modelo matemático, hace falta contar con un modelo de naturaleza psicológica o cognitiva que sustente las posibilidades de elección entre las alternativas al ítem.

En la literatura más o menos reciente se encuentran, como se ha mostrado anteriormente, ejemplos de modelos que intentan implementar algunas de estas ideas. Por ejemplo Tatsuoka (1983) estudió el éxito de los sujetos al resolver tests de matemáticas en términos de la aplicación de reglas correctas e incorrectas; Paulson en 1985 empleó modelos cognitivos para relacionar la probabilidad de responder correctamente a un ítem con las reglas que los sujetos pueden emplear al resolverlo y Wilson (Mislevy y Wilson; 1996) desarrolló un modelo que da cuenta de los cambios de un nivel a otro de conocimiento.

Todos estos modelos presentan alternativas diferentes para el diseño de tests más acordes a las necesidades actuales. Cada uno de ellos aborda una problemática específica y representa un avance en la teoría de medida. Las limitaciones de todos ellos radican principalmente en la interpretación de los parámetros y en su estimación. La complejidad matemática de los modelos aumenta notablemente por lo que se limita su posibilidad de aplicación.

En el presente trabajo se desarrolla un modelo de medida, con la intención de que, como los anteriores, sea útil para el diagnóstico. En él, la selección de las distintas opciones de respuesta al ítem está ligada al tipo de razonamiento que siguen los sujetos al resolverlo y a un modelo de desarrollo cognitivo sustentado en las ideas de Piaget. El modelo, por otra parte, toma en consideración el hecho de que cuando el sujeto enfrenta un ítem de opción múltiple, la elección de una entre varias alternativas de respuesta se ve afectada por la presencia de las otras alternativas.

El modelo parte del supuesto de que en caso de no haber posibilidad de elección de opciones, es decir, cuando el ítem es de respuesta abierta, la curva que describe la probabilidad de responder correctamente a un ítem con una sola posibilidad de respuesta correcta, se comporta de la misma manera que la curva característica de un modelo logístico de la teoría de respuesta a los ítems. Asimismo supone que las curvas que modelan las diferentes alternativas de respuesta se comportan de acuerdo a modelos de tipo logístico, cuando se consideran cada una independiente de las otras, considerando que estas opciones se han elegido sobre la base de las respuestas más frecunetemente empleadas por los sujetos ante una situación dada. Por último se pretende que cada uno de los parámetros que el modelo asocia al ítem se pueden relacionar con características específicas del ítem a través de la teoría cognitiva antes mencionada.

Una particularidad del modelo que aquí se presenta es que en lugar de desarrollarse a partir de consideraciones basadas en la estadística, se toma como punto de partida la teoría de los sistemas dinámicos. Esto representa un cambio en el paradigma que ha sustentado tradicionalmente el desarrollo de modelos dentro de la teoría de los tests.

Para desarrollar el modelo que se presenta en este trabajo se parte de la posibilidad de interpretar los modelos logísticos desde una perspectiva dinámica; es decir, tomando en consideración la variación en la probabilidad de responder correctamente al ítem. Por ello se mostrará en primer lugar cómo pueden derivarse los modelos logísticos tradicionales a partir de una perspectiva de esta naturaleza. Posteriormente se presentará el nuevo modelo en el cual se plantea una generalización del modelo dinámico simple. En ella se introducen nuevos parámetros que podrán ser relacionados con la teoría cognitiva que permitirá el diagnóstico.

Sección II.1 Derivación dinámica del modelo logístico

Algunos de los modelos de la teoría de respuesta a los ítems son modelos logísticos que ajustan la probabilidad de que un sujeto acierte a un ítem en términos de una función logística. Este tipo de función tiene una gran variedad de aplicaciones en el seno de las matemáticas y, en particular, puede interpretarse desde una perspectiva dinámica, en la que la función logística se representa mediante una ecuación diferencial de la forma:

$$dp/d\theta = kp - \lambda p^2 = kp\left(1 - \frac{\lambda}{k}p\right) \qquad (1)$$

en la que λ y k son constantes positivas y θ, la variable independiente, representa el rasgo que se desea medir.

El término kp de esta ecuación nos indica el cambio en la probabilidad de respuesta de cierta forma al ítem cuando aumenta θ. Por ejemplo:

Si p es pequeña, $\frac{dp}{d\theta} \approx kp$, es decir, la probabilidad aumenta proporcionalmente al valor de p. Este crecimiento es exponencial o Malthusiano. Por otro lado, cuando p está cerca de k/λ, se tiene que $\frac{dp}{d\theta} \approx 0$, esto es, p casi no crece, o crece muy lentamente acercándose al valor fijo k/λ donde se estabiliza.

La solución a esta ecuación diferencial, para una condición inicial $p(\theta_o) = p_o$, es precisamente la función logística que se plantea en los modelos IRT como se muestra a continuación:

La ecuación diferencial (1) es una ecuación diferencial de primer orden separable, es decir que se puede integrar directamente. Para ello la podemos reescribir como

$$\frac{dp}{kp - \lambda p^2} = d\theta$$

Integrando ambos lados de la ecuación se tiene que la primera integral $\int \frac{dp}{p(k - \lambda p)}$ se puede descomponer en fracciones parciales de la siguiente manera

$$\frac{1}{p(k - \lambda p)} = \frac{1}{kp} + \frac{\lambda}{k(k - \lambda p)} = \frac{1}{k}\left(\frac{1}{p} + \frac{\lambda}{k - \lambda p}\right)$$

Integrando tenemos que

$$\int \frac{dp}{p(k - \lambda p)} = \frac{1}{k}\int\left(\frac{1}{p} + \frac{\lambda}{k - \lambda p}\right) dp = \frac{1}{k}\left[\int \frac{dp}{p} + \int \frac{\lambda dp}{k - \lambda p}\right] = \frac{1}{k}\left[ln\ p - ln(k - \lambda p)\right]$$

y la segunda integral es

$$\int d\theta = \theta + c$$

combinando estas dos ecuaciones se obtiene la solución de la ecuación diferencial:

$$\frac{1}{k}\left[ln\ p - ln(k - \lambda p)\right] = \theta + c$$

en la que c es una constante de integración arbitraria.

La solución a la ecuación diferencial no es única; consiste en una familia de curvas. Si se tiene información adicional, como por ejemplo, el valor de la función

p para un valor dado de θ, a lo que se le conoce como condición inicial, se puede encontrar la solución única de la ecuación diferencial que es compatible con esa condición.

Utilizando la condición inicial $p(\theta_o) = p_o$ en la solución, y tomando, sin perder la generalidad, la condición inicial cuando $\theta_0 = 0$ se encuentra el valor de la constante,

$$c = \frac{1}{k}\left[ln\ p_o - ln(k - \lambda p_o)\right] - \theta_o \ = \frac{1}{k}\, ln\left[\frac{p_o}{(k - \lambda p)}\right]$$

por lo que la solución de la ecuación (1) toma la forma

$$\frac{1}{k}\left[ln\ p - ln(\ k - \lambda p)\right] = \theta + \frac{1}{k}\left[ln\ p_o - ln(\ k - \lambda p_o)\right], \quad o$$

$$\frac{1}{k}\ln(p/(k - \lambda p) = \theta + \frac{1}{k}\ln(p_o/(k{-}\lambda p_o).$$

Rearreglando los términos de esta última ecuación se obtiene que

$$\frac{1}{k}\, ln\left(\frac{p}{k - \lambda p}\right) - \frac{1}{k}\, ln\left(\frac{p_o}{k - \lambda p_o}\right) = \theta$$

$$\frac{1}{k}\ln\left[\frac{p(k - \lambda p_o)}{p_o(k - \lambda p)}\right] = \theta$$

$$\ln\left[\frac{p(k - \lambda p_o)}{p_o(k - \lambda p)}\right] = k\theta$$

Exponenciando ambos lados de esta ecuación se obtiene

$$\frac{p(k - \lambda p_o)}{p_o(k - \lambda p)} = e^{k\theta}$$

simplificando, podemos reescribir esta ecuación como

$$p_o(k - \lambda p)e^{k\theta} = p(k - \lambda p_o)$$

lo cual, resolviendo para p, queda como

$$p_o k e^{k\theta} - p_o \lambda p e^{k\theta} = p(k - \lambda p_o)$$
$$p_o k e^{k\theta} = p\left[p_o \lambda e^{k\theta} + (k - \lambda p_o)\right]$$

o sea,

$$p = \frac{p_o k e^{k\theta}}{p_o \lambda e^{k\theta} + (k - \lambda p_o)}$$

multiplicando por $\left(\frac{1}{\lambda p_o}\right)e^{-k\theta}$ en el denominador y el numerador del término de la derecha se obtiene

$$p = \frac{(k / \lambda)}{1 + \left(\frac{k}{\lambda p_o} - 1\right)e^{-k\theta}} \qquad o,$$

$$p(\theta) = \frac{k / \lambda}{1 + \left(\frac{k / \lambda}{p_o} - 1\right)e^{-k\theta}} \qquad (2)$$

La ecuación (2) es la solución única de la ecuación diferencial (1). Hasta este momento no se ha hecho ninguna referencia a que la ecuación con la que estamos trabajando se refiere a la probabilidad de respuesta a un ítem.

Tomando en consideración que $p(\theta)$ es una función de probabilidad, ésta puede tomar únicamente valores entre cero y uno; además, por tratarse de la probabilidad de acertar en la respuesta a un ítem, la función debe cumplir la condición de que cuando $\theta \to \infty$, $p \to 1$. Tomando el límite cuando θ tiende a infinito de la ecuación (2) se encuentra

$$\lim_{\theta \to \infty} p(\theta) = \lim_{\theta \to \infty} \frac{k / \lambda}{1 + \left(\dfrac{k / \lambda}{p_o} - 1 \right) e^{-k\theta}}$$

pero como $\lim_{\theta \to \infty} p(\theta) = 1$,

entonces $\dfrac{k}{\lambda} = 1$, o lo que es lo mismo $k=\lambda$.

Incorporando este resultado, la ecuación (2) se reduce a

$$p(\theta) = \frac{1}{1 + \left(\dfrac{1}{p_o} - 1 \right) e^{-k\theta}}. \qquad (3)$$

Esta ecuación puede reescribirse en la forma

$$\begin{aligned} p(\theta) &= \frac{1}{1 + \left(\dfrac{1 - p_o}{p_o} \right) e^{-k\theta}} \\ &= \frac{1}{1 + e^{\gamma - k\theta}} \qquad \text{con } e^{\gamma} = \left(\frac{1 - p_o}{p_o} \right) \\ &= \frac{1}{1 + e^{-k\theta + \gamma}} \end{aligned}$$

y finalmente como

$$p(\theta) = \frac{1}{1 + e^{-k(\theta - \gamma / k)}} \qquad (4)$$

Esta última ecuación es equivalente a la ecuación logística de un modelo logístico de dos parámetros como el que se revisó en la sección I.3 del capítulo anterior.

¿Cuál es el significado de esos parámetros? Para interpretar los parámetros se puede regresar a analizar la ecuación diferencial (1) de la cual se partió.

Derivando la ecuación (1) implícitamente con respecto a θ se obtiene la segunda derivada:

$$\frac{d^2p}{d\theta^2} = k\frac{dp}{d\theta} - 2\lambda p\frac{dp}{d\theta} = \frac{dp}{d\theta}(k - 2\lambda p)$$

Puesto que para que la función represente la curva del modelo logístico que interesa $k = \lambda$, la segunda derivada puede expresarse como

$$\frac{d^2p}{d\theta^2} = k(1 - 2p)\frac{dp}{d\theta}$$

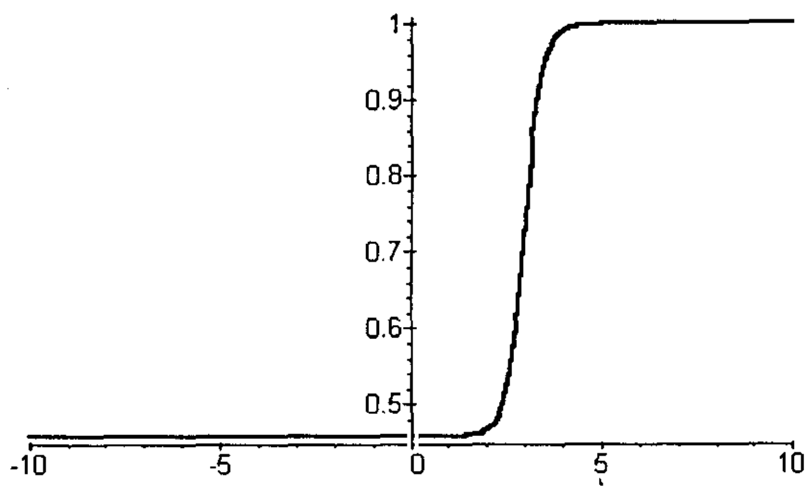
En esta última ecuación observamos que $\frac{d^2p}{d\theta^2} = 0$ cuando $p = \frac{1}{2}$, lo que nos indica que la derivada $\frac{dp}{d\theta}$ tiene un punto crítico en ese nivel de probabilidad.

Como además $\frac{d^2p}{d\theta^2} > 0$ si $p < 1/2$ y $\frac{d^2p}{d\theta^2} < 0$ si $p > 1/2$, la función de probabilidad tiene un máximo en el nivel de probabilidad 1/2.

Interpretando estos resultados podemos decir que cuando $p = ½$, la probabilidad está creciendo lo más rápidamente posible y la curva de probabilidad tiene un punto de inflexión; si $p<½$ la segunda derivada es positiva de manera que la curva de probabilidad es cóncava hacia arriba en esas condiciones, en cambio cuando $p>½$ la segunda derivada es negativa, lo que nos indica que la curva es cóncava hacia abajo. Por otra parte, puesto que $\frac{dp}{d\theta}$ es positiva siempre, la probabilidad aumenta en él.

Todos estos elementos nos permiten inferir que la curva tiene la forma de una sigmoide, como se muestra en la figura II.1.1.

Figura II.1.1



El valor de θ en el que cambia la concavidad de la gráfica en los modelos de respuesta a los ítems se asocia al índice de dificultad del ítem. A este índice suele denotársele por b. Como se ha visto anteriormente, en ese punto $p = 1/2$; sustituyendo se encuentra que en el modelo que aquí se presenta

$$b = \frac{\gamma}{k}$$

Puesto que el parámetro γ está relacionado con la probabilidad p_o de acertar el ítem cuando θ es cero, si p_o es grande, el índice de dificultad será pequeño. Por otra parte, el parámetro k es la razón a lo que cambia la probabilidad de acertar al ítem para valores de p cercanos a p_o. Si k es grande, la probabilidad de acertar el ítem crece muy rápidamente, en cuyo caso el índice de dificultad será pequeño.

La razón a la que crece la probabilidad en el punto de inflexión puede asociarse al índice de discriminación, ya que una curva muy suave, que crece lentamente no permite distinguir fácilmente entre la probabilidad de acertar o no acertar al ítem para valores distintos de θ; en cambio, una curva que se levanta bruscamente, permite una mejor distinción. Así, el índice de discriminación está asociado a la razón de crecimiento de la probabilidad en el punto θ correspondiente al índice de dificultad. Para encontrarlo, se calcula la derivada de la función $p(\theta)$ en el punto $\theta = b$:

$$\frac{dp}{d\theta} = \frac{ke^{-k(\theta-\gamma/k)}}{\left[1+e^{-k(\theta-\gamma/k)}\right]^2}$$

$$\frac{dp}{d\theta}(b) = \frac{dp}{d\theta}\left(\frac{\gamma}{k}\right) = k$$

El índice de discriminación suele denotarse por a y es proporcional a $p'(b)$; se puede decir entonces que

$$a \propto k.$$

lo que indica que el índice de discriminación del ítem es proporcional a la razón a la que cambia la probabilidad de acertar al ítem para valores cercanos a p_o.

Como puede notarse, tanto el índice de dificultad como el de discriminación están relacionados con la razón a la que cambia la probabilidad de acertar al ítem para valores de θ cercanos a cero. Si, además de estas consideraciones tomamos en cuenta que es conveniente que la curva logística se ajuste lo más posible a la curva normal acumulada, es necesario introducir la constante D=1.7 que se derivó en el capítulo anterior, de manera que el parámetro k puede escribirse en término del índice de discriminación y de la constante D como

$$k=aD \qquad (5),$$

Incorporando toda esta información, la solución a la ecuación diferencial de la que se partió, ésta se transforma en

$$p(\theta)=\frac{1}{1+e^{-Da(\theta-b)}}=\frac{e^{-Da(\theta-b)}}{1+e^{-Da(\theta-b)}},$$

que coincide con el modelo logístico de respuesta a los ítems de dos parámetros.

En el caso particular en el que todos los ítems tienen el mismo índice de discriminación, bastaría ajustar los valores que toma la ecuación

$$p(\theta)=\frac{1}{1+e^{-k(\theta-\gamma/k)}}$$

a los de la curva normal acumulada, haciendo k=1.7=D, de tal manera que la curva característica del ítem quedaría como

$$p(\theta)=\frac{1}{1+e^{-D(\theta-b)}}$$

que corresponde a la ecuación del modelo logístico de un parámetro.

Por último, si en el modelo anterior tomamos en consideración la posibilidad de que el ítem sea acertado al azar, será necesario añadir a la expresión de la probabilidad una constante que representa esa posibilidad.

La ecuación diferencial en sí no se ve afectada, pues la forma en que la probabilidad de responder acertadamente el ítem crece de la misma forma que anteriormente con la variación de la variable θ.

Si llamamos c a la probabilidad de acertar el ítem al azar, la curva característica del ítem vendría dada por

$$p(\theta) = c + (1 - c)p_2(\theta)$$

en la que $p_2(\theta)$ es la función de probabilidad de dos parámetros.

Esta ecuación indica que la función de probabilidad total está dada por la suma de la probabilidad de responder al azar, más la probabilidad de responder correctamente el ítem. La primera es simplemente c y la segunda es $(1-c)p_2(\theta)$ por lo que la curva característica del ítem puede escribirse como

$$p(\theta) = \frac{c + (1 - c)e^{Da(\theta - b)}}{1 + e^{Da(\theta - b)}}$$

De esta manera se ha demostrado que los modelos logísticos clásicos de la teoría de respuesta a los ítems se pueden representar en forma equivalente, sea a través de una función dinámica expresada en la forma de una ecuación diferencial o mediante la ecuación de probabilidad dada por la curva logística.

La expresión dinámica con la que dió inicio esta sección puede utilizarse a continuación como punto de partida en la construcción de un nuevo modelo que permita hacer una extensión de los modelos de la teoría de respuesta a los ítems.

Sección II.2 Posibilidad de extensión del modelo IRT usando una perspectiva dinámica

El modelo logístico que se ha considerado hasta aquí se aplica a ítems diseñados para respuestas dicotómicas, del tipo verdadero o falso, correcto o incorrecto o que pueden reducirse a este tipo de respuesta. En este tipo de modelos, aunque el sujeto elige entre diferentes opciones, para efecto de medida, se considera únicamente si la elección es la correcta o es incorrecta. El propósito de este trabajo consiste en extender el modelo logítico para incluir en él el efecto que ejerce el hecho de que haya alternativas de respuesta entre las que el sujeto puede elegir. La inclusión de ese factor, al que se denominará parámetro de interacción permite tomar en consideración en el modelo la elección de cada una de los diferentes distractores.

La respuesta a un ítem de opción múltiple es, como se ha visto en el capítulo anterior, de naturaleza diferente a la solución de un ítem de respuesta construida. En el caso de ítems de opción múltiple, la inclusión de las opciones alternativas de respuesta tiene necesariamente un efecto sobre el sujeto que responde. Este efecto depende de la seguridad con la que el sujeto se desempeña frente al ítem o de su estado de conocimiento. La posibilidad de medir el efecto que la presencia de las opciones alternativas ejerce sobre quien responde permite adecuar el modelo para que sea útil por una parte con fines de diagnóstico y por otra, con fines de prueba del propio diseño de los ítems.

Considerando las posibilidades de diagnóstico del modelo se puede decir que en la construcción del modelo se parte de la consideración de que al enfrentar una pregunta en un examen de opción múltiple, en el que las opciones de respuesta se han diseñado cuidadosamente y con base en una teoría cognitiva, el sujeto que responde debe tomar una decisión en la elección entre ellas. La toma de decisiones implica que el sujeto pondere las opciones antes de elegir entre ellas e incluye la posibilidad de que el sujeto dude entre las distintas opciones de respuesta que se le ofrecen. Este proceso puede concebirse, en términos del modelo de medida para el ítem, como una situación en la que en el sujeto se da una especie de competencia entre las posibles opciones por la atención del sujeto, o como una situación de interacción entre las opciones presentadas que tiene lugar en la mente del sujeto.

Dependiendo del tipo de ítem del que se trate, el resultado de esta competencia, que se objetivará en la respuesta del sujeto, será diferente. Por ejemplo, si el ítem que se presenta intenta medir el conocimiento del sujeto respecto a un cierto tema, la respuesta del sujeto dependerá del estado que guarda el conocimiento del individuo en esa materia en el momento de responder al ítem. Si las opciones se diseñan tomando en cuenta los resultados de las teorías cognitivas o de las didácticas específicas de las disciplinas, en base a los errores que se cometen más frecuentemente, se esperaría que conforme el conocimiento de un individuo avanza, la probabilidad de elegir opciones incorrectas al ítem debiera decrecer, mientras que la de elegir la opción correcta debería crecer; la probabilidad de elección de cualquiera de las opciones será un indicador del nivel de conocimiento del sujeto.

Existen, sin embargo, otros tipos de ítem que no se diseñan con el fin de detectar el nivel de conocimiento de una persona en un área o disciplina específica, sino otras cosas, por ejemplo sus preferencias o sus actitudes. En estos casos, la probabilidad de elección entre una u otra opción depende de los valores o de circunstancias contextuales del individuo en el momento de responder al ítem, por lo que el comportamiento de la curva característica correspondiente a cada opción puede tener formas diferentes: puede ser creciente, decreciente o exhibir ambos comportamientos; en cualquier caso, la información acerca de la elección particular entre las distintas opciones en un caso de esta naturaleza, es un indicador que permite medir con mayor certeza el rasgo de interés.

Se puede intentar modelar matemáticamente situaciones de esta naturaleza de una forma similar a la que se modelan mediante la representación de cada posible opción de respuesta a un ítem con un modelo dinámico en el que se incluya, además de los parámetros que caracterizan a cada opción, un factor que represente la competencia entre las posibles opciones de respuesta a un ítem. Este factor puede modelarse matemáticamente mediante la inclusión de un término que describa cómo el hecho de tener frente a sí diversas opciones de respuesta genera, en el sujeto que responde, la necesidad de valorar las distintas opciones y de tomar una decisión de elección. Este término puede considerarse como una interacción entre los distractores de un ítem.

El modelo matemático aquí descrito puede contener un número cualquiera de términos de interacción, dependiendo del número de distractores que contenga el ítem. Aunque se presentará el modelo general, con el fin de simplificar la situación y dado que en los ítems de opción múltiple lo más frecuente es que aparezcan tres o cuatro dictractores, se desarrollará un modelo para la interacción entre dos posibles opciones de respuesta a un mismo ítem, tomando como punto de partida el modelo logístico de la teoría de respuesta a los ítems de dos parámetros. Este modelo se puede generalizar para tomar en consideración cualquier número de opciones de respuesta y la interacción entre ellas.

Desde una perspectiva dinámica, es decir, en términos de ecuaciones diferenciales, la curva característica del ítem se puede representar, como ya se ha visto, mediante la función logística:

$$\frac{dp}{d\theta} = ap(1 - p)$$

Esta ecuación rige el comportamiento del crecimiento de la probabilidad $p(\theta)$ de la respuesta al ítem y esta probabilidad tiende a un valor límite de 1 cuando θ tiende a infinito. El parámetro *a*, como se vió en la sección anterior, está relacionado tanto con el índice de dificultad del ítem cuanto con el índice de discriminación.

Si en lugar de considerar el ítem completo se pone la atención en las opciones de respuesta, se puede modelar la probabilidad de elegir entre dos posibles opciones de respuesta a la misma pregunta. La probabilidad de elegir cada uno de los distractores puede representarse mediante curvas logísticas y se pueden asignar curvas diferentes a cada una de las opciones. Dichas curvas tienen diferentes características dependiendo del tipo de ítem de que se trate y del tipo de respuesta que se espere para él.

Por ejemplo, en una situación en la que se pretende medir el desempeño de estudiantes frente a problemas algebraicos, aquellos estudiantes que manejan el álgebra con soltura, responderán $-a$ a la pregunta encontrar $|a|$, si $a<0$, que es la respuesta correcta; para alguien cuyo manejo del álgebra es más limitado el encontrarse con la respuesta a como posibilidad de elección le generará un conflicto que le hará dudar en la elección de cualquiera de las dos alternativas. Conforme la habilidad algebraica del estudiante aumenta, la probabilidad de elegir la opción correcta aumentará, mientras que la de elegir la opción incorrecta disminuirá y, de hecho, el tipo de respuesta que una persona da a este tipo de pregunta, proporciona un indicio diagnóstico del estado de conocimiento del álgebra en el que la persona se encuentra.

Por otra parte, si el ítem es tal que admite dos tipos de respuesta, ambas correctas, como es el caso de aquellas preguntas que se refieren a cuestiones de gusto o de actitud, la probabilidad de elegir una u otra opción puede crecer logísticamente, de acuerdo a curvas con parámetros diferentes. Estas curvas deben ser tales que para cualquier valor de la variable independiente θ, la suma de las probabilidades de elección de cada una de las opciones sea menor o igual que uno.

Podemos empezar el planteamiento del modelo para un caso como el del primer ejemplo; es decir, aquél en el que las opciones que se presentan son una correcta y otra incorrecta. En este modelo, como ya se dijo, es de esperarse que conforme el valor de la variable a medir aumenta, la probabilidad de elegir la opción correcta aumenta y debe tender a uno. En cambio, la probabilidad de elegir la opción incorrecta debe ser grande para valores pequeños de la variable y debe tender a cero cuando ésta tiende a infinito. Es importante destacar que en la región intermedia de valores de la variable, la probabilidad de elegir una de las opciones más la probabilidad de elegir la otra es siempre menor a uno, ya que si sumaran uno bastaría con una sola curva característica para modelar esta situación. La probabilidad restante podría asignársele a la probabilidad de no elegir ninguna de las opciones presentadas, o de elegir una tercera, si la hubiera.

Si la opción correcta es fácilmente distinguible, la presencia de otras opciones generará poca duda en quien responde, es decir, la interacción entre las opciones es pequeña. Si las opciones son ambas igualmente factibles para el sujeto que responde el ítem, la interacción entre ellas debe ser grande y generar así un conflicto grande.

Utilizando un modelo de esta naturaleza se pueden diseñar ítems que permitan distinguir entre sujetos con distinto tipo de habilidad, siempre y cuando las opciones de respuesta se diseñen con base en los resultados de algún modelo cognitivo. Con un modelo simple como el que aquí se propone sería posible diseñar instrumentos que permitan determinar el intervalo de valores para la variable a medir en el que se da el conflicto y usar esta información con fines diagnósticos.

En el caso de situaciones en las que la variable no mide el nivel de desempeño, se pueden hacer consideraciones semejantes a las arriba mencionadas, aunque las curvas asociadas a las distintas opciones podrían tener formas diversas. En todo caso, la construcción de un instrumento en el que las opciones de respuesta a los distintos ítems se diseñen de acuerdo a un modelo psicológico pertinente y se analicen en base al modelo matemático que se pretende desarrollar en este trabajo, permitirá realizar un diagnóstico a través de la identificación del intervalo en el que se da el conflicto y el significado que ese conflicto tiene en términos de la variable que se desea medir.

Sección II.3 Propuesta de un modelo dinámico con interacción

A partir de los planteamientos hechos en las secciones anteriores, en esta sección se elaborará un modelo matemático de la interferencia que puede producir en un sujeto la presencia de opciones de respuesta alternativas a un ítem, partiendo de la ecuación logística expresada en una forma dinámica. Se desarrollará este modelo para el caso en el que la competencia se da entre dos opciones y posteriormente se generalizará al caso en el que haya ítems con *n* distractores.

Tomemos como punto de partida el modelo logístico de dos parámteros en su forma dinámica. En dicho modelo el cambio en la probabilidad de acertar al ítem está dado por

$$\frac{dp}{d\theta} = ap(1 - p)$$

En este modelo, si $a > 0$, la probabilidad de acertar al ítem tiende a 1 cuando θ tiende a infinito.

Ahora supongamos que se diseña un ítem en el que se puede elegir una de dos opciones de respuesta, o no responder. Sean $p_1(\theta)$ y $p_2(\theta)$ las probabilidades de elección de las opciones 1 y 2 respectivamente para un cierto valor de la variable θ. Sean, además k_1 y k_2 los valores a los que tienden cada una de estas probabilidades cuando θ tiende a infinito. Por ser $p_1(\theta)$ y $p_2(\theta)$ probabilidades de elección de las opciones de un ítem particular, debe cumplirse que $p_1 + p_2 \leq 1$, para cualquier valor de θ.

Supongamos, por otra parte, que si el ítem es de respuesta construida, es decir, que se presenta en una forma abierta y por lo mismo no hay distractores presentes explícitamente, la probabilidad de elegir una forma de respuesta acorde a la elección de cada una de las posibles opciones implícitas, está determinada por una ecuación logística. Es decir, en ausencia de la otra opción la probabilidad de responder de una u otra forma está dada por

$$\frac{dp_1(\theta)}{d\theta} = a_1 p_1 (1 - p_1 / k_1)$$

y
$$\frac{dp_2(\theta)}{d\theta} = a_2 p_2 (1 - p_2 / k_2)$$

Cuando ambas opciones están presentes es de esperarse que se genere, en quien responde, una duda o un conflicto que puede modelarse como la interacción entre las opciones presentes y que tiene un efecto sobre la variación de la probabilidad de elección de cada una de las opciones, es decir:

$$\frac{dp_1(\theta)}{d\theta} = a_1 p_1 (1 - p_1 / k_1) - m_2 p_1$$

$$\frac{dp_2(\theta)}{d\theta} = a_2 p_2 (1 - p_2 / k_2) - m_1 p_2$$

en las que m_1 y m_2 son parámetros que se refieren a la probabilidad de elegir las opciones 1 y 2 respectivamente. En la primera de estas ecuaciones puede notarse que para valores pequeños de la interacción m_2, el cambio en p_1 es aproximadamente igual a

$$a_1 p_1 (1 - p_1 / k_1)$$

y la probabilidad de elección de la primera opción, p_1 se comporta de la misma manera que en el modelo logístico sin interacción. Sin embargo, al aumentar la interacción el término $m_2 p_1$ se hace más significativo e indica cómo la presencia de la otra alternativa de respuesta afecta el crecimiento de la probabilidad de elección de la primera opción.

Podría considerarse que el parámetro de interacción m_2 fuera constante; de esta manera, el término que describe esta interacción sería proporcional a p_1 y la constante m_2 sería una medida del grado en el que la presencia de la otra opción afecta la elección de la primera. Sin embargo, un modelo de esta naturaleza no parece ser el más adecuado, debido a que la interacción de una opción sobre la otra difícilmente será constante para sujetos caracterizados mediante valores distintos de θ. El modelo es más realista si el término m_2 depende de la probabilidad de elección de la segunda opción para distintos valores de θ.

Lo mismo podría decirse para la elección de la segunda opción con respecto a la primera.

Una primera posibilidad para modelar la variación de m_2 con respecto a p_1 y de m_1 con respecto a p_2 podría ser $m_1 = p_1$ y $m_2 = p_2$; sin embargo, en general éste no es el caso, ya que sería muy improbable que una misma persona tuviera preferencias iguales para cada una de las opciones que se le presentan. En realidad, lo que es de esperarse es que cada una de las opciones atraiga la atención del sujeto en forma diferente y que esta diferencia dependa además del nivel θ en el que se encuentra el sujeto. Por ello, es más conveniente considerar que m_1 y m_2 son proporcionales a la probabilidad de elección de las opciones 1 y 2 respectivamente, es decir:

$$m_1 = b_1 p_1$$
y
$$m_2 = b_2 p_2$$

con b_1 y b_2 constantes

Las constantes b_1 y b_2 indicarían así el grado de influencia que ejerce la presencia de una opción sobre la otra. Si una opción no interacciona con la otra, en el sentido de que no causa ninguna confusión en el sujeto o que no implica ningún cambio en la toma de decisiones del sujeto que responde, b_1 y b_2 son iguales a cero. Si las opciones son muy similares desde la perspectiva del sujeto que responde, esas constantes serán cercanas a la unidad. Si, por otra parte, una de las opciones interactúa fuertemente con la otra, es decir, provoca confusión o duda en el sujeto que responde y reclama más fuertemente su atención, el coeficiente correspondiente será mayor.

La forma en la que se da la interacción entre las distintas opciones de respuesta, o distractores del ítem, depende también del tipo de pregunta. Si el ítem intenta medir el desempeño en una cierta área del conocimiento, la elección entre las distintas opciones dependerá del estado de conocimiento de la persona en el área que se pretende medir y se esperaría que la probabilidad de que se responda incorrectamente al ítem decrezca al aumentar el nivel de conocimiento del individuo, mientras que la probabilidad de responderlo correctamente crece y se acerca a k conforme aumenta el nivel de conocimiento.

Para lograr este efecto en el modelo, el coeficiente de la parte lineal del sistema de ecuaciones correspondiente a la opción incorrecta debe ser negativo, mientras el de la opción correcta es positivo. Si, en cambio, las posibilidades de elección no reflejan el grado de conocimiento en un tema, sino otro factor como gusto o actitud de los sujetos, la probabilidad de elección de las distintas opciones responde a los valores del sujeto o a sus condiciones contextuales en el momento en que toma la decisión. En este caso las curvas que modelan a cada opción pueden crecer ambas, decrecer ambas o una crecer y la otra decrecer de acuerdo a las circunstancias de lo que se desea medir, y los parámetros a_1 y a_2 pueden ser ambos positivos, ambos negativos o uno positivo y el otro negativo.

Introduciendo las constantes b_1 y b_2 que modelan la interacción, las ecuaciones del modelo pueden escribirse como:

$$\frac{dp_1}{d\theta}(\theta) = a_1 p_1 (1 - p_1 / k_1) - b_1 p_2 p_1$$

y $$\frac{dp_2}{d\theta}(\theta) = a_2 p_2 (1 - p_2 / k_2) - b_2 p_1 p_2$$

De esta manera, cuando se introduce el término de interacción en el modelo, la curva que representa la probabilidad de elección de cada uno de los distractores deja de ser logística y presenta un comportamiento no lineal.

El modelo desarrollado hasta aquí para el caso en el que dos opciones de respuesta compiten por la atención del sujeto que responde al ítem puede generalizarse, en los mismos términos que los que se han empleado hasta ahora como:

$$\frac{dp_1}{d\theta}(\theta) = a_1 p_1 (1 - p_1 / k_1) - \sum_{i=1}^{n} b_{i1} p_i p_1 \text{, para } i \neq 1$$

$$\frac{dp_2}{d\theta}(\theta) = a_2 p_2 (1 - p_2 / k_2) - \sum_{i=1}^{n} b_{i2} p_i p_2 \text{, para } i \neq 2 \quad ....$$

$$\frac{dp_n}{d\theta}(\theta) = a_n p_n (1 - p_n / k_n) - \sum_{i=1}^{n} b_{in} p_i p_n \text{, para } i \neq n$$

Estas ecuaciones permiten el análisis de ítems con cualquier número de opciones de respuesta.

Sección II.4 Interpretación de los parámetros del modelo

El modelo que se ha considerado en la sección anterior constituye una extensión del modelo logístico de dos parámetros a situaciones en las que se pretende analizar el papel de los distractores y a situaciones en las que se desea obtener mayor información diagnóstica de los sujetos a partir de las respuestas a los ítems. En esta sección se considerará el significado de los parámetros del modelo en términos de la situación que nos interesa modelar.

El modelo representa la probabilidad de elección de cada uno de los distractores del ítem tomando en consideración el hecho de que dicha elección se ve influida de alguna manera por la presencia de los otros distractores.

El modelo se presenta a través de un sistema de ecuaciones diferenciales en el que cada ecuación indica la probabilidad de elección de una opción en un ítem de opción múltiple. De la solución a estas ecuaciones se obtiene lo que podría considerarse como una curva característica para cada distractor.

Desde el punto de vista matemático los parámetros a_1 , a_2 y en general todos los parámetros a_i , para *i* desde 1 hasta *n,* en el modelo general representan la razón de crecimiento de las probabilidades de elección de la primera, segunda y demás opciones, esto es, en general del distractor *i* respectivamente, en ausencia de cualquier otra opción. Como se vió anteriormente en el caso del modelo para el ítem completo, estos parámetros están relacionados tanto con el índice de dificultad cuanto con el índice de discriminación del ítem.

En términos del modelo que se desarrolla aquí se puede considerar que estos parámetros indican o miden qué tanto los diferentes distractores demandan la atención del sujeto que responde, es decir qué tan accesibles o viables le resultan; por ello, estos parámetros pueden denominarse parámetros de accesibilidad del distractor o de la opción. Cuando a_i es grande la probabilidad de elección de la opción *i* crece rápidamente; esto puede considerarse como un indicador de que el distractor llama poderosamente la atención del sujeto que responde, o le es accesible. Si, en cambio, a_i es pequeña, la probabilidad de elegir dicha opción crece lentamente y la opción puede considerarse como menos accesible a los ojos del individuo.

Las constantes b_1, b_2 y en general b_{ij} representan la intensidad de la interacción entre las posibles parejas de opciones de respuesta del modelo. Cuando se manejan únicamente dos opciones de respuesta entre las que interesa la interacción, por comodidad se utilizarán los símbolos b_1 y b_2. En el modelo general este papel lo juegan los parámetros b_{ij}.

Como ya se mencionó, si ambas opciones son igualmente factibles para el sujeto que responde, los parámetros b_{ij} tendrán un valor cercano a uno; por el contrario, si las opciones no generan confusión alguna en el individuo que responde, los valores de los parámetros b_{ij} serán iguales a cero.

Los parámetros k_1 y k_2 en el modelo de dos opciones y k_i en general, representan los niveles máximos que pueden alcanzar las probabilidades de elección de la primera y segunda opciones de respuesta o, en general la *i*-ésima opción.

Es interesante notar que las razones a_1 / b_1 , a_2 / b_2 y en general a_i / b_{ij} representan la razón entre la accesibillidad y la competencia entre los distractores, es decir, son una medida del crecimiento neto de la probabilidad de respuesta de cada una de las opciones. Como se ha visto anteriormente, estos parámetros en el caso del modelo para un ítem, están relacionados tanto con el índice de dificultad cuanto con el índice de discriminación del ítem. En el caso del modelo para los distractores estas razones pueden considerarse una medida de qué tan atractiva es la opción correspondiente tomando en cuenta la interacción que ejercen las otras opciones.

El tipo de curva que se obtiene cuando se encuentra la solución del modelo depende también de las condiciones iniciales. Éstas juegan, en este modelo, el papel de parámetros que pueden estimarse conjuntamente con los otros parámetros del modelo. La información que las condiciones iniciales proporcionan es interesante pues permite caracterizar el nivel de desarrollo de la población con la que se utiliza el ítem o el test. Si las condiciones iniciales son tales que el valor de la probabilidad de responder correctamete al ítem es alto, se tiene un indicador de que los sujetos con los que se utiliza el ítem tienen cierto nivel de avance en el rasgo que se pretende medir con el ítem. En cambio, si dicha probabilidad es baja, se puede inferir que la muestra que responde el ítem es débil en el rasgo a medir. Cuando se consideran conjuntamente una serie de ítems en un test, los valores de las condiciones iniciales permiten localizar las áreas o rasgos a medir que requieren de mayor atención y proveen así ciertos elementos de diagnóstico.

Sección II.5 Análisis cualitativo del modelo

Una vez planteado el modelo de medida es conveniente analizar sus implicaciones desde el punto de vista matemático; es decir, describir las variables, los parámetros y analizar las posibilidades de solución.

Algunas consideraciones matemáticas

El modelo desarrollado en este trabajo involucra las derivadas de varias variables dependientes: la probabilidad de elegir cada una de las opciones de respuesta a un ítem. Un conjunto como este de ecuaciones que incluyen derivadas ordinarias de más de una variable dependiente se denominan sistemas de ecuaciones diferenciales. Puesto que las ecuaciones involucran únicamente la primera derivada de las variables dependientes se considera que el sistema es de primer orden. Por otra parte, cuando los términos que contienen a las variables dependientes, en este caso p_1 y p_2 son lineales se dice que el sistema es lineal; en otro caso el sistema es no lineal.

Por ejemplo, el sistema

$$dx/dt = 6x + y$$
$$dy/dt = 3x - 2y$$

es un sistema de ecuaciones diferenciales de primer orden lineales, mientras que el sistema

$$dx/dt = 6x + \text{sen } y$$
$$dy/dt = 3y - 2xy$$

es un sistema no lineal.

El modelo que se ha desarrollado en este capítulo para analizar la probabilidad de elección de los distintos distractores de un ítem es un ejemplo de un sistema no lineal de ecuaciones diferenciales de primer orden ya que en la expresión que aparece del lado derecho de cada igualdad se presenta un término no lineal de la forma $p_1\ p_2$.

Por otra parte la variable dependiente θ, no aparece explícitamente en las expresiones que componen el sistema de ecuaciones. En estos casos se dice que el sistema es autónomo.

Una solución a un sistema de ecuaciones diferenciales es una *n*-ada de ecuaciones paramétricas $p_1 = p_1\ (\theta\ ),\ p_2\ = p_2\ (\theta\ ),\ ...,\ p_n\ = p_n\ (\theta)$ que satisfacen las ecuaciones del sistema sobre un cierto intervalo de valores de θ por lo que las curvas solución del sistema son necesariamente funciones continuas de θ.

La solución de un sistema de ecuaciones diferenciales puede representarse de diferentes maneras. La más común en el caso de sistemas lineales es la expresión analítica de las funciones paramétricas antes mencionadas. Otra forma de representar la solución es mediante una gráfica que muestre la variación de cada una de las funciones individualmente, contra la variable independiente. Sin embargo, a menudo es más revelador hacer una gráfica en la que se muestra, en el caso de dos variables, la trayectoria de los puntos $(p_1(\theta), p_2(\theta))$ en el plano $p_1\, p_2$ conforme la variable dependiente varía en el intervalo de solución. A esta trayectoria se le conoce como trayectoria u órbita del sistema de ecuaciones diferenciales. Al plano que contiene a esta trayectoria se le conoce como plano fase o fásico del sistema.

Es conveniente citar en este punto algunos resultados de la teoría de ecuaciones diferenciales que son importantes y útiles cuando se investigan las propiedades de la solución de los sistemas autónomos:

1. Por cualquier punto del plano fase pasa únicamente una trayectoria.

2. Una trayectoria que se inicia en un punto diferente al punto de equilibrio no puede alcanzarlo para valores finitos de θ.

3. Las trayectorias en el plano fase no se cruzan, a menos que la trayectoria sea una curva cerrada, en cuyo caso se tiene una solución periódica.

Es conveniente asimismo pensar en la trayectoria u órbita como el camino que sigue una partícula que se mueve en un plano para interpretar con mayor facilidad el comportamiento del sistema, independientemente de si esta representa una situación de movimiento.

Examinando las ecuaciones diferenciales en sí mismas, se puede estudiar gráficamente el comportamiento cualitativo de las soluciones de los sistemas de ecuaciones diferenciales. Esta técnica es particularmente útil para investigar el comportamiento de sistemas no lineales autónomos en los que es muy difícil o tal vez imposible encontrar soluciones analíticas.

Dado cualquier sistema de ecuaciones diferenciales el teorema de existencia y unicidad de soluciones permite verificar si existe la solución. Este teorema dice que dado un problema de valor inicial, si las funciones que definen las ecuaciones diferenciales son continuas y sus derivadas respecto a las variables dependientes son también continuas en una región que contiene al valor inicial, entonces existe un intervalo de valores de la variable independiente alrededor del valor inicial para el que existe la solución al problema y esta solución es única. En el caso del modelo que nos concierne en este trabajo, las hipótesis del teorema de existencia y unicidad de las soluciones se cumplen y, por ello, podemos asegurar que existirá una solución al sistema de ecuaciones propuesto en el modelo y que esa solución es única dadas las condiciones iniciales.

Cuando el sistema de ecuaciones diferenciales es lineal es posible encontrar la solución en forma analítica; si no lo es, en la mayoría de los casos, debe recurrirse a los métodos cualitativos para el análisis de las soluciones y a los métodos numéricos para obtener la solución.

Analisis gráfico de las soluciones del modelo

El modelo que se derivó en la sección II.3 es un modelo no lineal. Los sistemas de ecuaciones diferenciales no lineales son, como hemos visto, difíciles de resolver analíticamente, sin embargo, las técnicas de análisis cualitativo permiten resolver algunas preguntas importantes que surgen en la situación que nos interesa estudiar, tales como ¿es posible que se de una coexistencia de elección de opciones para algún intervalo de valores de θ? o ¿es alguna de las opciones dominante sobre las otras? o ¿para qué valores de θ predomina una de las opciones sobre la otra?

Las técnicas de análisis cualitativo son muy útiles para poder responder a estas cuestiones y se pueden emplear en una gran variedad de situaciones que incluyen aquéllas en las que el sistema puede ser resuelto analíticamente. Para llevar a cabo el análisis cualitativo de las soluciones del sistema de ecuaciones diferenciales se recurre a la idea del espacio fase. En esta sección se abordará únicamente el caso del modelo para la interacción entre dos distractores por la dificultad de dibujar espacios de dimensión mayor a dos.

El análisis cualitativo de la solución de un sistema de ecuaciones diferenciales puede hacerse también a través de la graficación de las funciones solución, que en este caso serían $p_1(\theta)$ y $p_2(\theta)$.

Una curva solución al sistema de ecuaciones diferenciales es una función cuyas coordenadas son $p_1(\theta)$, $p_2(\theta)$ cuando θ varía sobre el intervalo de solución I del sistema, en el caso que nos interesa aqui del sistema que se obtuvo en la sección II.4; la curva solución se conoce asimismo como trayectoria u órbita del sistema de ecuaciones. Al plano que contiene a esta trayectoria se le llama, como ya se ha visto, plano fase o fásico del sistema y a menudo, para interpretar la información contenida en esa curva, es conveniente usar el símil de la órbita en el plano fase con la trayectoria de una partícula que se mueve en un plano coordenado.

Utilizando el plano fase se puede, mediante el análisis de las ecuaciones diferenciales en sí mismas, sin necesidad de resolverlas analíticamente, estudiar gráficamente el comportamiento cualitativo de las soluciones a los sistemas de ecuaciones diferenciales. Esta técnica es particularmente útil para estudiar sistemas autónomos no lineales para los cuales es muy difícil encontrar una solución analítica.

Entre las preguntas que se pueden plantear al hacer el análisis cualitativo destaca la búsqueda de puntos críticos o puntos de equilibrio. Utilizando la analogía de la partícula en movimiento resulta de interés encontrar los puntos en que la órbita permanece estacionaria, es decir, aquellos puntos para los que no hay movimiento ni en la dirección p_1 ni en la dirección p_2, lo que indicaría que el móvil está parado o que la partícula está detenida. Esto ocurre cuando el punto (p_1, p_2) en el plano fase es tal que $f(p_1, p_2) = 0$ y $g(p_1, p_2) = 0$ simultáneamente con f y g definidas como $\frac{dp_1}{d\theta} = f$, $\frac{dp_2}{d\theta} = g$, esto es cuando tanto $dp_1 / d\theta$ cuanto $dp_2 / d\theta$ son cero. A un punto de esta naturaleza se le denomina punto crítico, punto de equilibrio o punto fijo del sistema.

Es importante notar que cuando (p_{1_0}, p_{2_0}) es un punto de equilibrio del sistema, las ecuaciones $p_1 = p_{1_0}$ y $p_2 = p_{2_0}$ son una solución del sistema. Esta solución de estado estacionario es la única que puede pasar por el punto (p_{1_0}, p_{2_0}) en el plano fase. La trayectoria asociada a esta solución es simplemente el punto de equilibrio en sí mismo. La partícula está "en reposo" en ese punto.

Cuando una trayectoria, $p_1(\theta) = p_1$, $p_2(\theta) = p_2$ es tal que $p_1(\theta) \to p_{1_0}$ y $p_2(\theta) \to p_{2_0}$, cuando $\theta \to \infty$, decimos que la órbita se acerca o converge al punto de equilibrio. En muchos problemas es de interés preguntarse qué ocurre con la trayectoria cuando se aproxima a un punto de equilibrio. En este contexto la idea de estabilidad juega un papel primordial. Un punto de equilibrio es estable si cualquier trayectoria cercana al punto (p_{1_0}, p_{2_0}) permanece cercana a él para cualquier valor de θ. Un punto de equilibrio es asintóticamente estable si es estable y si cualquier trayectoria cercana a (p_{1_0}, p_{2_0}) se acerca a este punto cuando $\theta \to \infty$. Si el punto de equilibrio no es estable se denomina inestable.

Regresemos ahora al modelo que nos interesa en este trabajo para hacer el análisis cualitativo correspondiente. El modelo está representado por el sistema de ecuaciones diferenciales

$$\frac{dp_1}{d\theta} = a_1 p_1 (1 - \frac{p_1}{k_1}) - b_1 p_2 p_1$$
$$\frac{dp_2}{d\theta} = a_2 p_2 (1 - \frac{p_2}{k_2}) - b_2 p_1 p_2$$

en las que $p_1(\theta)$ representa la probabilidad de elección de la opción 1 de un ítem y $p_2(\theta)$ representa la probabilidad de elección de la opción 2 del mismo ítem.

Al hacer el análisis cualitativo de la solución de este sistema de ecuaciones diferenciales es importante encontrar en primer lugar los puntos de equilibrio de las órbitas y posteriormente el comportamiento de las mismas.

La condición, como se mencionó anteriormente, para tener un punto de equilibrio del sistema consiste en que ambas derivadas, $dp_1 / d\theta$ y $dp_2 / d\theta$ se hagan simultáneamente cero. en este caso esta condición significa que:

$$\frac{dp_1}{d\theta} = 0 \quad \text{cuando } p_1 = 0 \text{ ó cuando } p_2 = \frac{a_1}{b_1} - \frac{a_1}{b_1 k_1} p_1$$

$$\frac{dp_2}{d\theta} = 0 \quad \text{cuando } p_2 = 0 \text{ ó cuando } p_2 = k_2 - \frac{k_2 b_2}{a_2} p_1$$

Dependiendo de los valores de los diferentes parámetros a_1 , b_1 , k_1 , a_2 , b_2 y k_2 se pueden presentar casos con diferentes puntos de equilibrio. Es pues conveniente hacer un estudio de esos distintos casos de interés: localizar para cada uno de ellos los puntos de equilibrio y estudiar qué es lo que ocurre con la órbita en los puntos cercanos a ellos.

A continuación se hace un estudio detallado del comportamiento de las probabilidades de elección de las opciones para las distintas posibilidades de relación entre los valores de los parámetros.

Parámetros positivos

Consideremos en primer lugar la situación en la que todos los parámetros de las ecuaciones son positivos. De ser así, dependiendo de las relaciones entre los parámetros se pueden presentar diferentes casos:

CASO 1.

Supongamos que tanto a_1 cuanto a_2 son mayores que cero y que

$$\frac{a_1}{b_1} > k_2 \text{ y } \frac{a_2}{b_2} > k_1$$

En este caso los puntos de equilibrio son $(0,0)$, $(0, k_2)$, $(k_1, 0)$ y el punto de intersección de las rectas

$$p_2 = \frac{a_1}{b_1} - \frac{a_1}{b_1 k_1} p_1 \quad \text{-----(L1)}$$

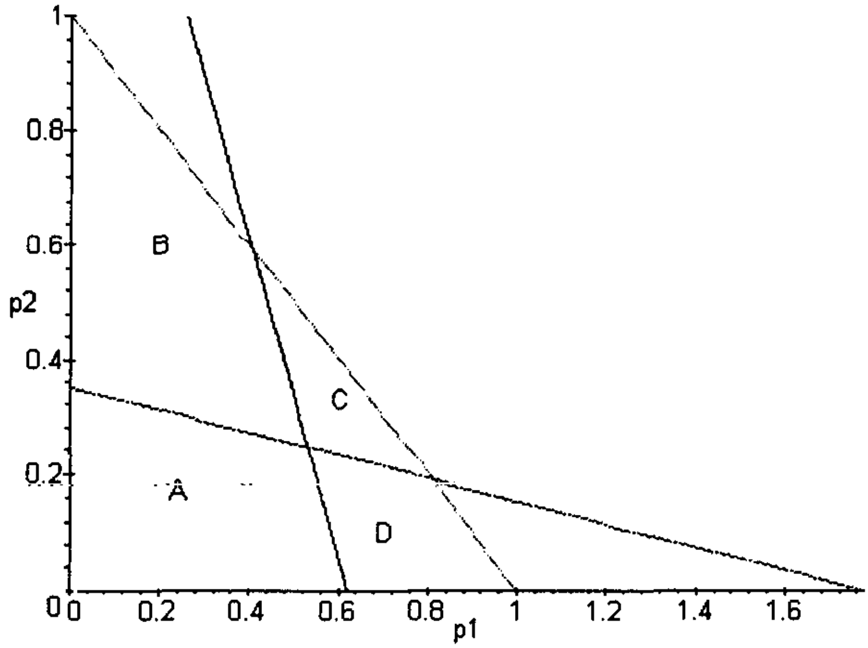
y $$p_2 = k_2 - \frac{k_2 b_2}{a_2} p_1 \quad \text{-----(L2)}$$

al que denotaremos por (p_{1I}, p_{2I}).

Si el valor inicial de los niveles de probabilidad se encuentra en cualquiera de estos valores, no habrá variación alguna del nivel de probabilidad cuando θ cambia.

La región de estados posibles en el plano fase para este sistema se muestra en la figura II.5.1a.

Figura II.5.1a: Regiones en plano fase



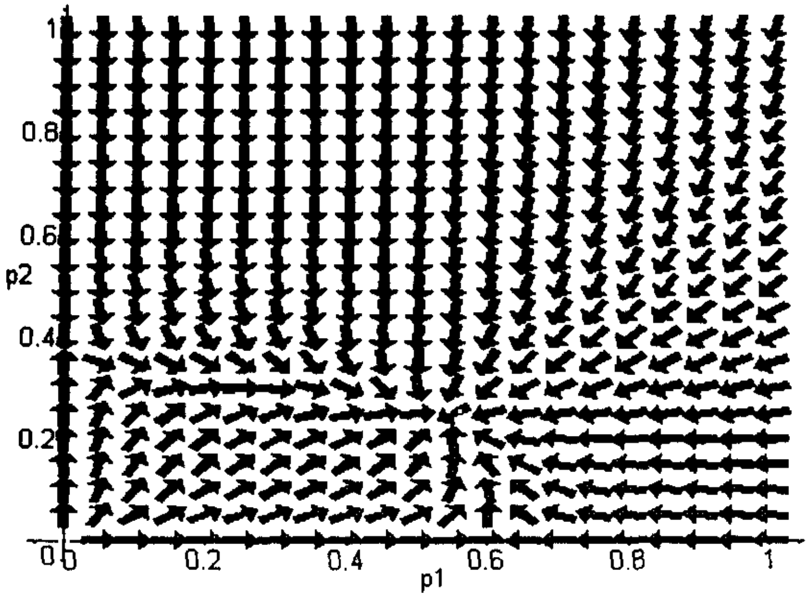
Si el estado de probabilidad se encuentra sobre la recta $p_1 = 0$ o sobre la recta $p_2 = \frac{a_1}{b_1} - \frac{a_1}{b_1 k_1} p_1$, la probabilidad de elección de la opción 1 permanece constante, por lo que la trayectoria debe cruzar estas rectas horizontalmente. Al pasar sobre las rectas $p_2 = 0$ y $p_2 = k_2 - \frac{k_2 b_2}{a_2} p_1$, en cambio, es la probabilidad de elección de la opción 2 la que no cambia, lo que implica que la trayectoria debe cruzar estas rectas verticalmente.

Para analizar aquello que ocurre cuando la trayectoria en el espacio fase se encuentra en cada una de las regiones en las que el plano queda dividido por los puntos de equilibrio y las rectas anteriores es conveniente hacer un análisis de los signos de las derivadas en cada una de ellas. Este análisis nos permite entender lo que ocurre para diferentes posibilidades de combinaciones de valores iniciales de los niveles de probabilidad, sin necesidad de calcular los parámetros ni de resolver el sistema de ecuaciones diferenciales. Es conveniente recordar que nos interesa en particular el comportamiento de las trayectorias en una vecindad de los puntos de equilibrio, especialmente nos interesa determinar si esos puntos son estables o inestables.

Cuando $\frac{dp_1}{d\theta}$ es positiva, la componente horizontal de la trayectoria aumenta (p_1 aumenta) y, por ello, el nivel de probabilidad p_1 se moverá hacia la derecha en el plano fase ; cuando $\frac{dp_1}{d\theta}$ es negativa, el movimiento del nivel de probabilidad p_1 será hacia la izquierda. De la misma manera, cuando $\frac{dp_2}{d\theta}$ es positiva, p_2 aumenta y el nivel de probabilidad p_2 se desplazará hacia arriba y cuando $\frac{dp_2}{d\theta}$ es negativo, el desplazamiento del nivel de probabilidades p_2 lo hará hacia abajo.

Las rectas que se muestran en la figura II.5.1a dividen al plano fase en regiones en las que el comportamiento de la trayectoria será diferente. Este comportamiento puede deducirse a partir del análisis de los signos de cada una de las derivadas de las variables dependientes en cada una de las regiones en las que el plano fase queda dividido. Analizando los signos de las derivadas de las funciones $p_1(\theta)$ y $p_2(\theta)$ en cada región y combinando la información obtenida, se encuentra el campo de direcciones en el espacio fase. Este se muestra en la figura II.5.1b

Figura II.5.1b: Campo de direcciones



Si analizamos el comportamiento de la trayectoria alrededor de los puntos de equilibrio encontramos lo siguiente:

Los puntos (0,0) , (0, k_2) y (k_1 , 0) son inestables. Esto significa que para condiciones iniciales cercanas a ellos, la trayectoria se alejará de ellos y se dirigirá hacia el punto (p_{1i}, p_{2i}).

El punto (p_{1i}, p_{2i}) es un punto de equilibrio asintóticamente estable, lo que significa que todas las trayectorias, independientemente de la región de la que provengan convergen hacia él.

El análisis gráfico nos permite, como ya se había mencionado con anterioridad, determinar el comportamiento de las posibles soluciones dadas distintas condiciones iniciales. Siguiendo el comportamiento de las curvas solución se puede inferir si la solución lleva a algún punto de equilibrio y si éste es estable o inestable, o en otros casos, si se aleja de él. En este caso el análisis cualitativo de las soluciones conduce a la conclusión de que, independientemente de las condiciones iniciales, siempre y cuando éstas no se encuentren exactamente sobre los puntos de equilibrio, conforme el valor de θ aumenta la trayectoria se acercará al punto (p_{1i}, p_{2i}). Es decir, la probabilidad p_{1i} de elegir la opción 1 y la probabilidad p_{2i} de elegir la opción 2 ya no variarán para valores mayores de θ.

En la figura II.5.1c se muestran las gráficas para p_1 contra θ, p_2 contra θ, y en la figura II.5.1d se muestran algunas órbitas en el espacio fase correspondientes a este caso.

Figura II.5.1c: Curvas solucion

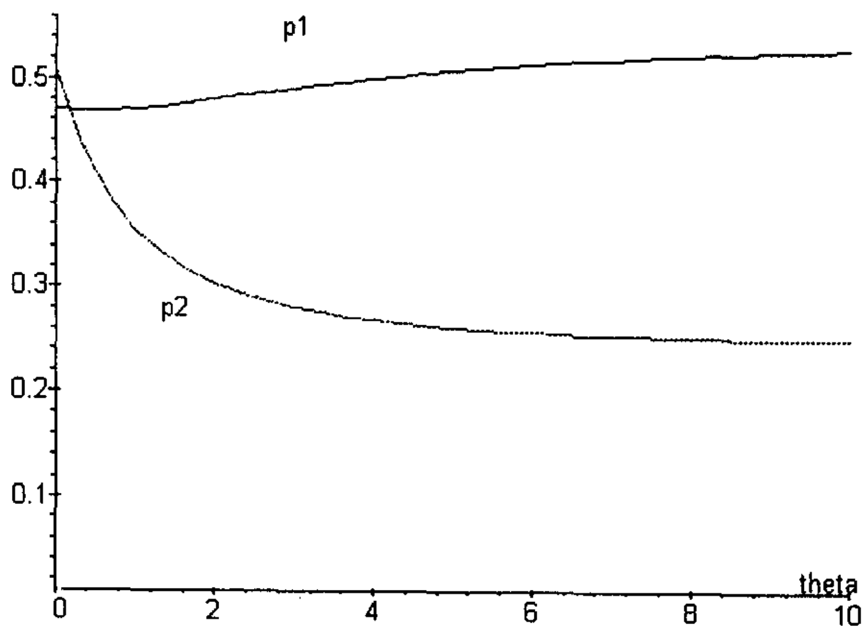
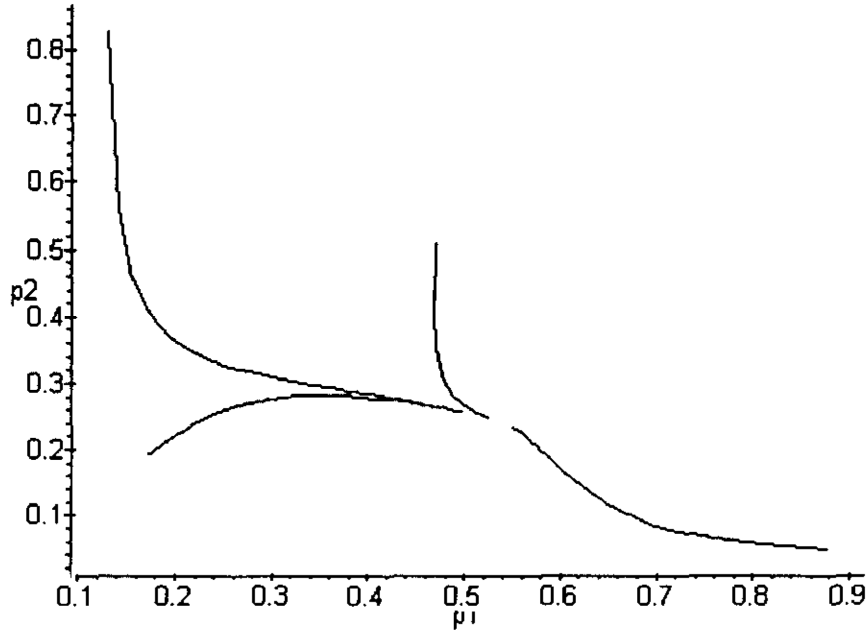


Figura II.5.1d: Trayectorias en plano fase



CASO 2.

Cuando $k_2 > \frac{a_1}{b_1}$ y $k_1 > \frac{a_2}{b_2}$ el análisis gráfico, hecho de la misma manera que en el caso anterior implica la existencia de los mismos puntos de equilibrio, es decir $(0,0)$, $(0, k_2)$, $(k_1, 0)$ y (p_{1i}, p_{1l}).

El análisis de los signos de las derivadas en cada región nos ayuda a interpretar el comportamiento de las trayectorias en la vecindad de los puntos de equilibrio. En la figura II.5.2a se muestra el campo de direcciones que se obtiene. En este caso, los puntos $(0, k_2)$ y $(k_1, 0)$ son estables, mientras que el origen y el punto de intersección de las dos rectas no lo son. En la figura II.5.2b se muestran algunas de las trayectorias compatibles con este análisis.

Figura II.5.2a: Regiones en plano fase

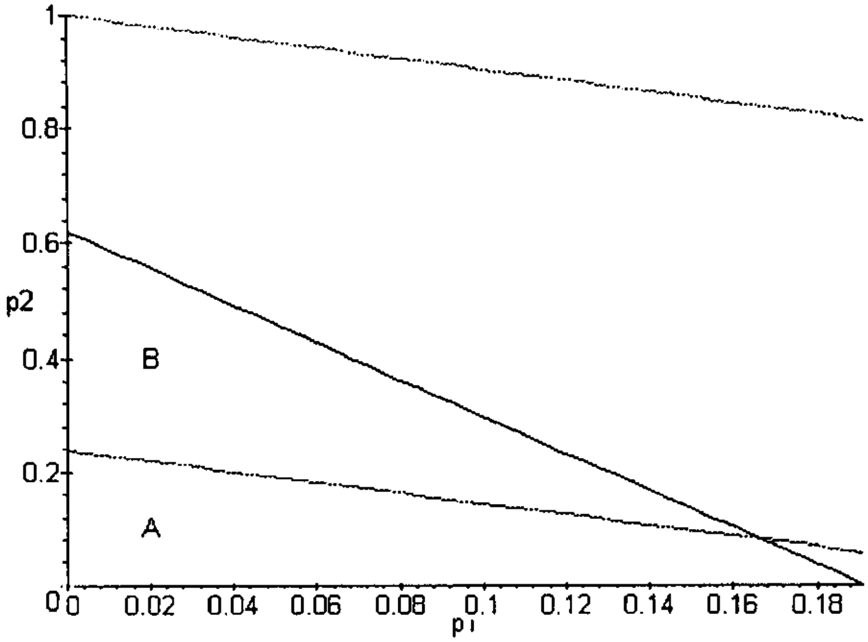
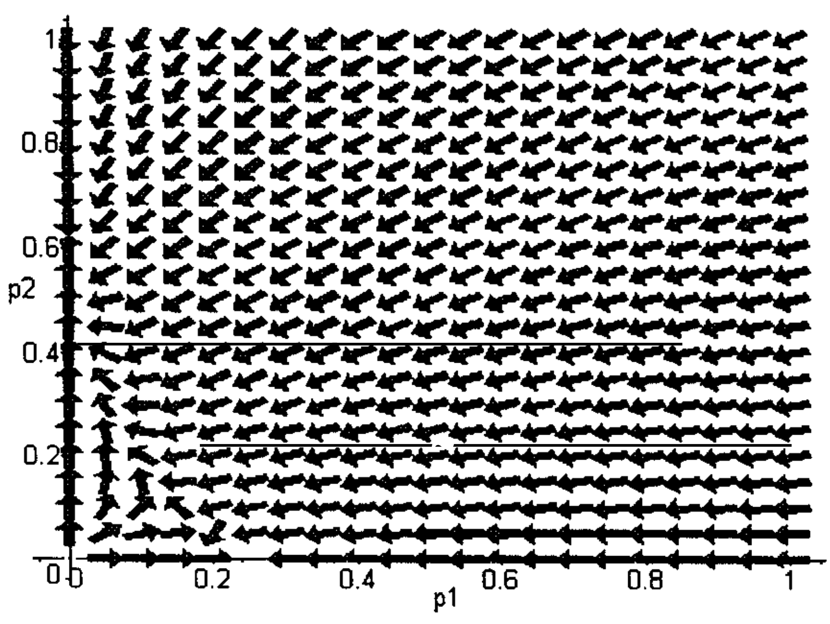


Figura II.5.2b: Campo de direcciones



En este caso, dependiendo de las condiciones iniciales se presentan dos posibilidades; la primera consiste en que al aumentar los valores de θ sea la opción 1 la que prevalezca, desapareciendo la probabilidad de elección de la opción 2 y la segunda el caso en el que prevalece la segunda opción de respuesta y desaparece la probabilidad de elegir la primera opción. En la figura II.5.2c se muestran las gráficas p_1 contra θ y de p_2 contra θ y en la figura II.5.2d se muestran algunas trayectorias en el espacio fase para este caso. Se observa aquí que la coexistencia de la elección de opciones es posible, pero no predicha por el modelo.

Figura II.5.2c: Curvas solucion

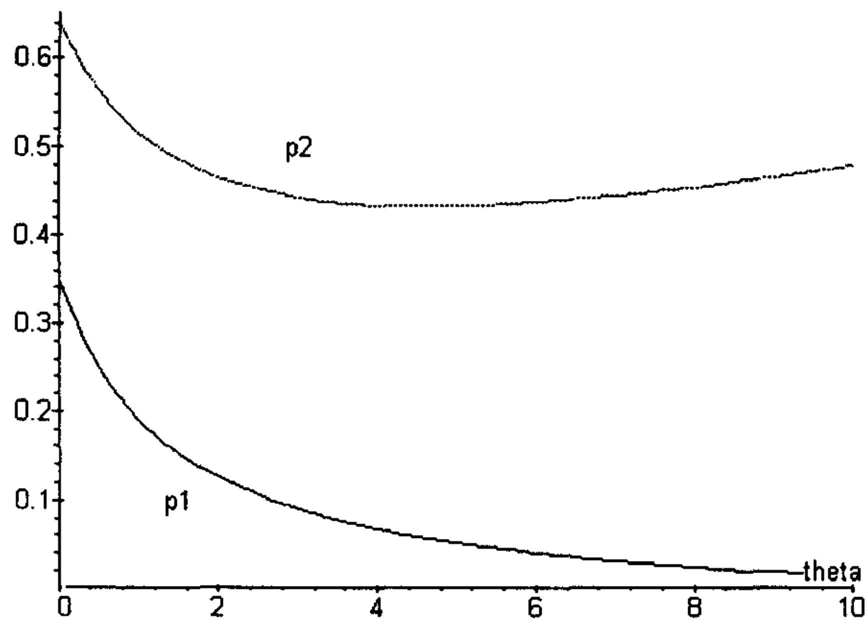
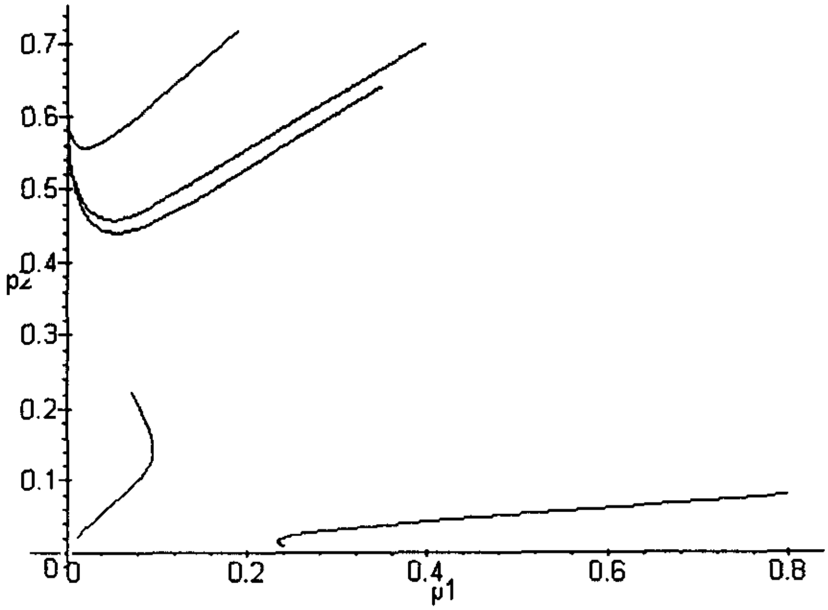


Figura II.5.2d: Trayectorias en plano fase



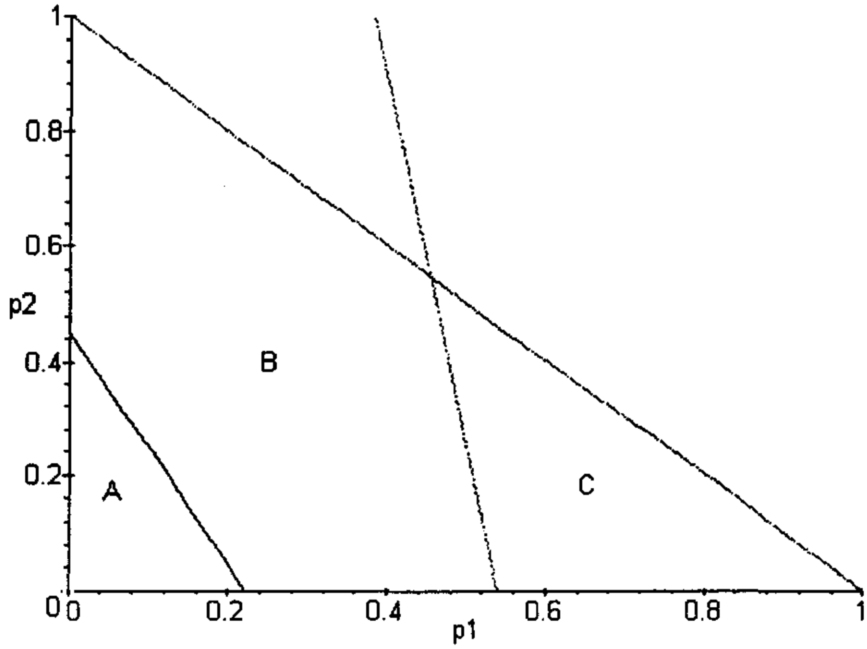
CASO 3.

Si los valores de los parámetros son tales que $k_2 < \frac{a_1}{b_1}$ y $k_1 > \frac{a_2}{b_2}$,

la recta que une al punto $(0,\ a_1\ /\ b_1)$ con $(k_1,0)$ y aquella que une a $\left(\frac{a_2}{b_2},0\right)$ con $(0,\ k_2)$ no se intersectan.

El plano fase queda entonces dividido en tres regiones que se muestran, en la figura II.5.3a y los puntos de equilibrio son únicamente (0,0), $(0,k_2)$ y $(k_1,0)$.

Figura II.5.3a: Regiones en plano fase



El análisis de los signos de las derivadas en cada región, que se muestra en la figura II.5.3b, pone en relieve que el único punto de equilibrio asintóticamente estable es el punto $(k_1, 0)$. Las gráficas correspondientes a las probabilidades de elección para las opciones en esta situación se muestran en la figura II.5.3c, mientras que la figura II.5.3d muestra algunas de las trayectorias en el espacio fase.

Figura II.5.3b: Campo de direcciones

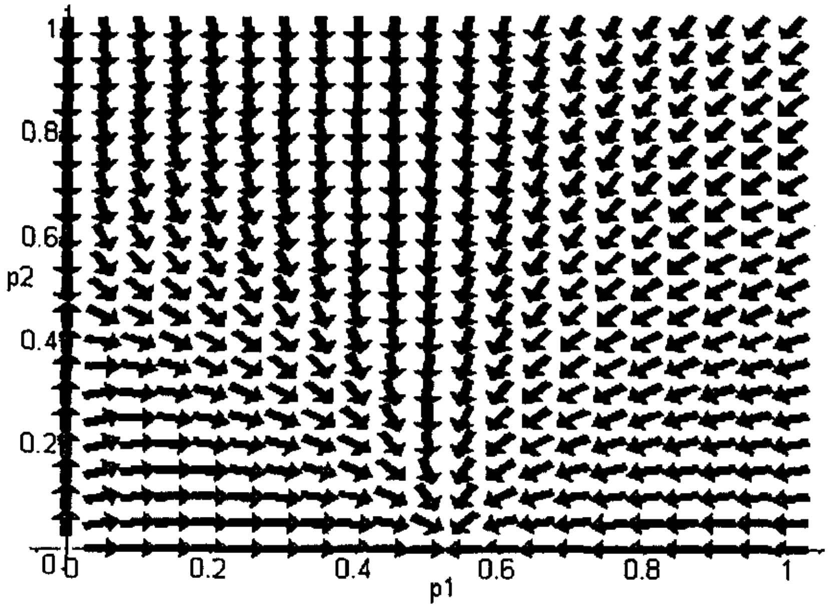


Figura II.5.6c: Curvas solucion

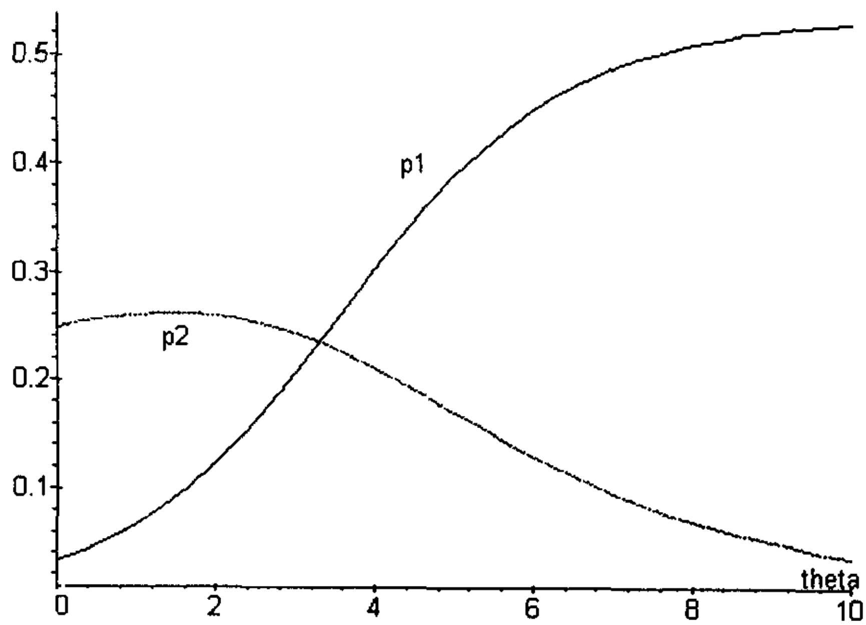
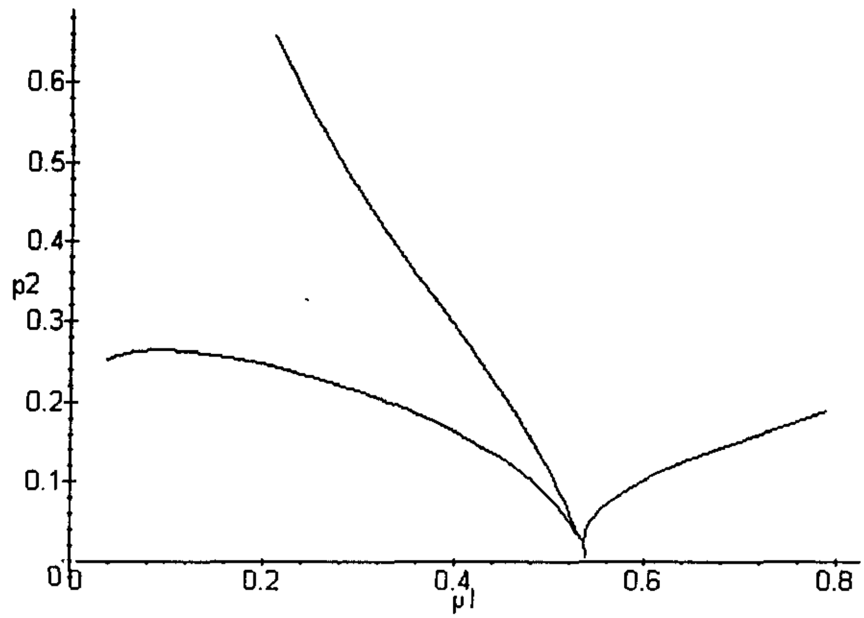


Figura II.5.3d: Trayectorias en plano fase



En esta situación, la elección que prevalece al aumentar los valores de θ es la de la primera opción, independientemente de las condiciones iniciales.

CASO 4.

Para el caso en que los valores de los parámetros cumplen las condiciones:

$$\frac{a_1}{b_1} < k_q \quad \text{y} \quad \frac{a_2}{b_2} > k_1$$

los puntos de equilibrio son nuevamente $(0, k_2)$, $(k_1, 0)$ *y* $(0,0)$. Otra vez se tiene un caso en el que no hay intersección entre las rectas mencionadas anteriormente. Las regiones permitidas para la solución se muestran en la figura II.5.4a.

El análisis de los signos de las derivadas para esta situación nos indica, como se muestra en la figura II.5.4b, que el punto $(0, k_2)$ es estable y que los puntos $(k_1, 0)$ y $(0,0)$ son inestables. Las gráficas correspondientes a las probabilidades de elección de las opciones con respecto a θ se muestran en la figura II.5.4c y algunas trayectorias en el espacio fase se muestran en la figura II.5.4d.

Figura II.5.4a: Regiones en plano fase

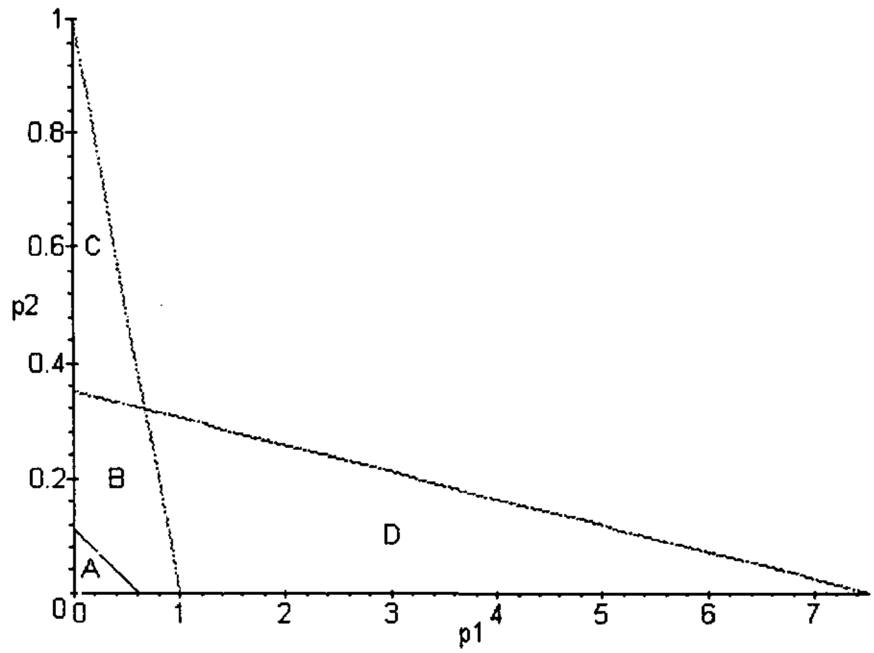


Figura II.5.4b: Campo de direcciones

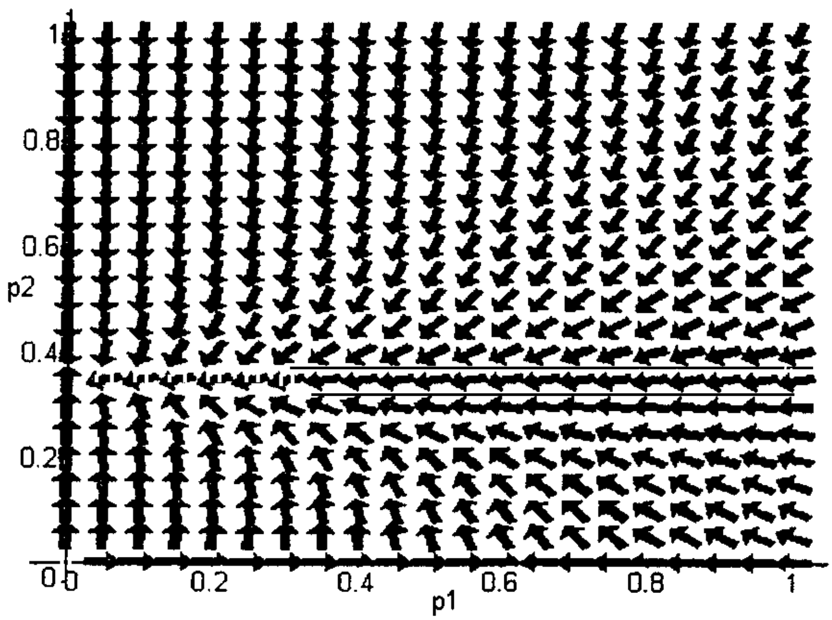


Figura II.5.4c: Curvas solucion

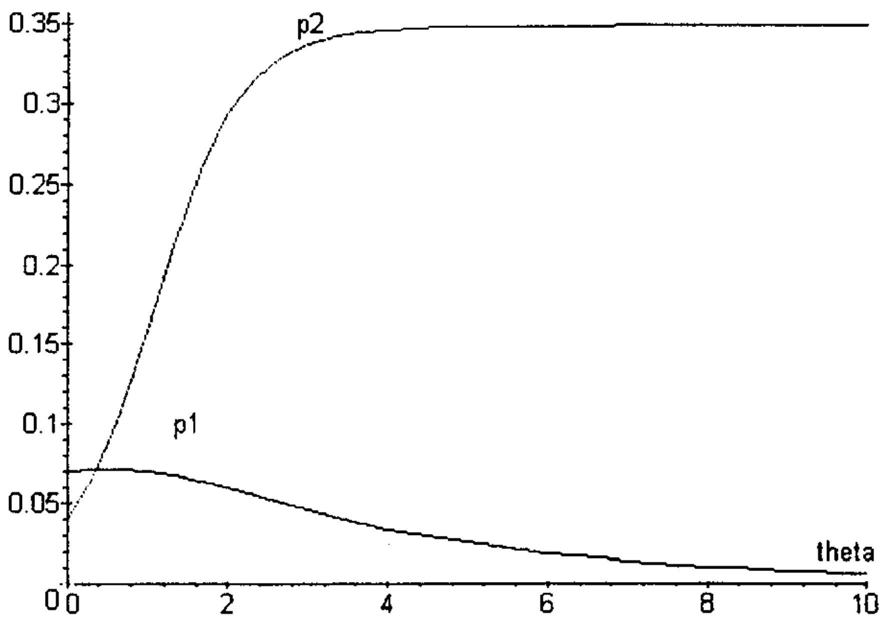
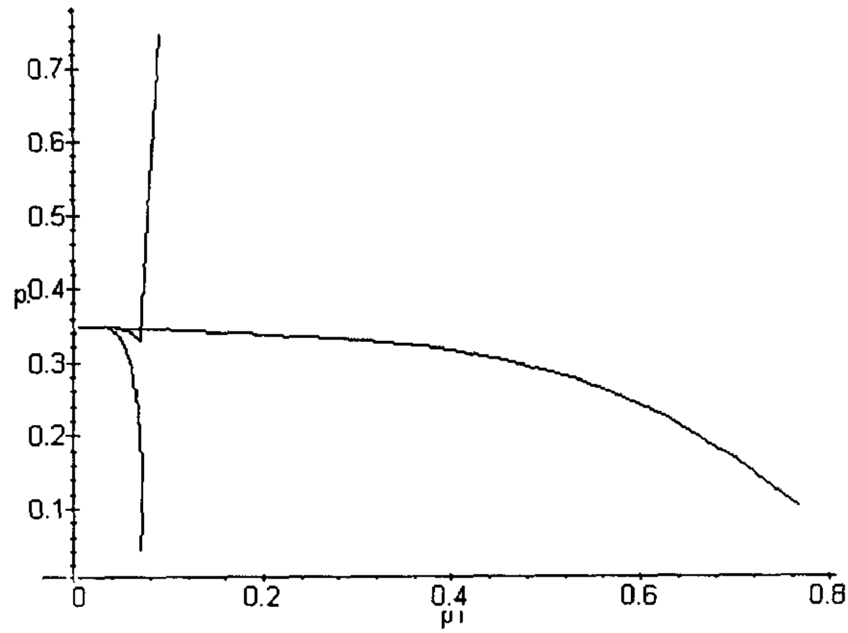


Figura II.5.4d: Trayectorias en plano fase



El análisis gráfico de este caso nos conduce a la conclusión de que, independientemente de las condiciones iniciales, la elección de la opción 2 prevalece al aumentar los valores de θ.

CASO 5.

Un último caso a considerar es aquel en el que $k_2 = \frac{a_1}{b_1}$ y $k_1 = \frac{a_2}{b_2}$.

En este caso las regiones permitidas para la solución se muestran en la figura II.5.5a. Todos los puntos sobre la recta que une a los puntos $(k_1, 0)$ y $(0, k_2)$ son puntos de equilibrio estables, incluyendo a los puntos $(k_1, 0)$ y $(0, k_2)$. El punto (0,0) también es un punto de equilibrio, pero inestable. El campo de pendientes en las distintas regiones del espacio fase se muestra en la figura II.5.5b, mientras que las gráficas correspondientes a las curvas solución y a las curvas en el espacio fase se muestran en las figura II.5.5c y II.5.5d respectivamente.

Figura II.5.5a: Regiones en plano fase

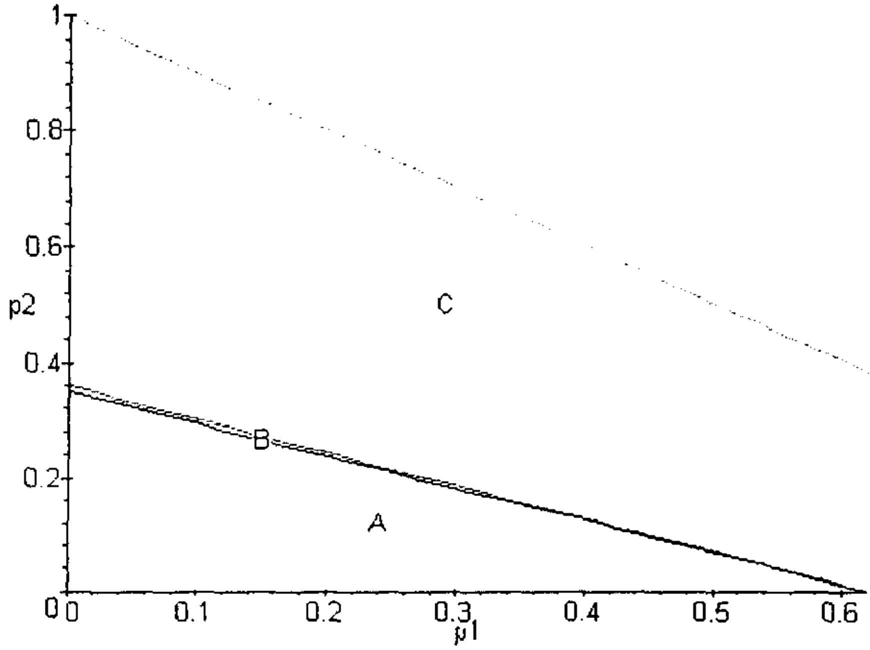


Figura II.5.5b: Campo de direcciones

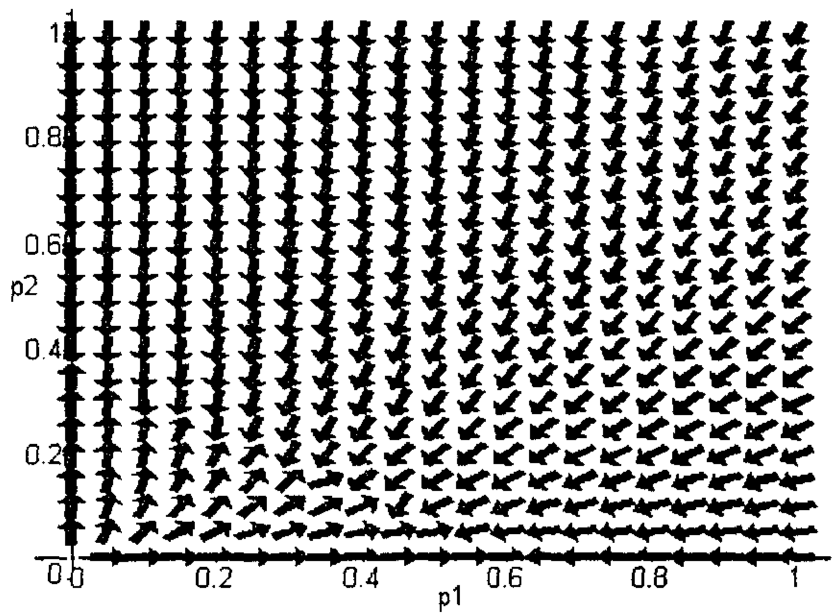


Figura II.5.5c: Curvas solucion

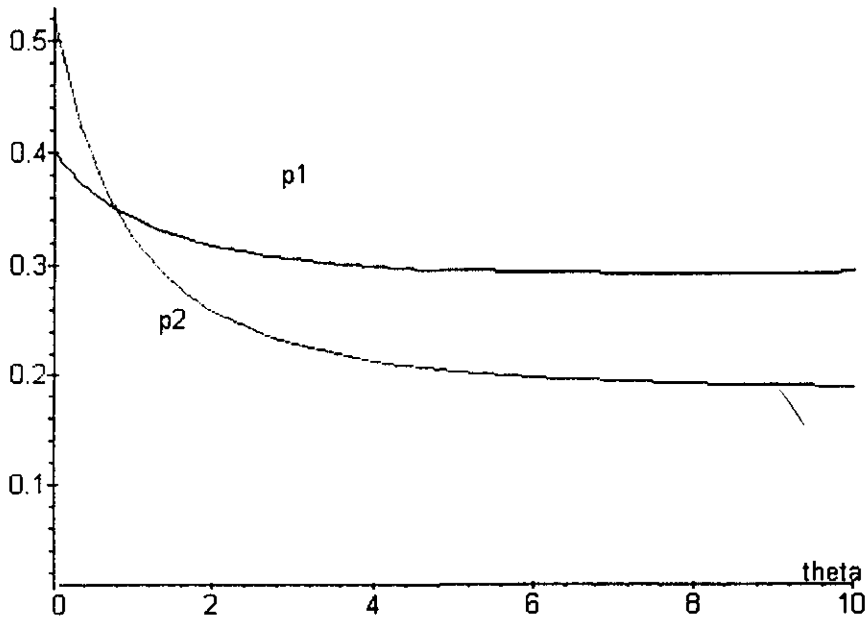
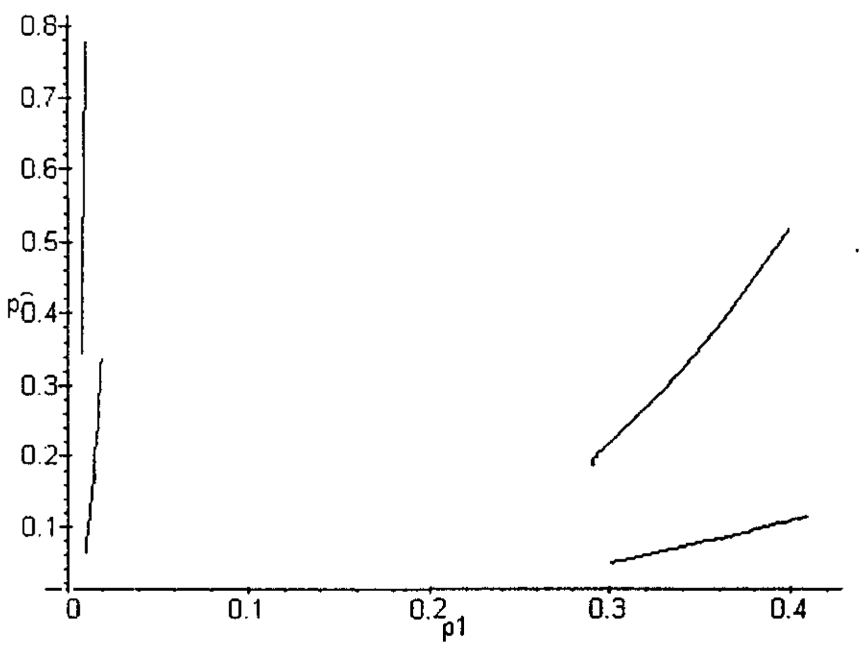


Figura II.5.5d: Trayectorias en plano fase



En este caso, la coexistencia de las dos opciones de elección para el ítem es muy posible. Los valores de probabilidad para cada una de ellas dependen de los valores de los parámetros y de las condiciones iniciales.

Hasta este punto, el análisis de los casos se ha hecho bajo el supuesto de que todos los parámetros que aparecen en el sistema de ecuaciones diferenciales son positivos

Un parámetro de accesibilidad negativo

Una posibilidad importante e interesante del modelo se presenta cuando, al considerar las curvas de probabilidad de elección de respuesta sin interacción, una de ellas es creciente y la otra es decreciente. Para efectos del modelo, esto indicaría que uno de los parámetros a_1 o a_2 es negativo, ya que son estos parámetros los que definen el crecimiento o decrecimiento de las curvas logísticas correspondientes p_1 y p_2. El caso que nos interesa analizar es aquél en el que uno de los parámetros, digamos a_1 es positivo y el otro, a_2 es negativo; así, la curva correspondiente a la probabilidad de elección de la opción 1 sin interacción será una logística creciente cuyo límite cuando $\theta \to \infty$ es k_1 y la curva correspondiente a la probabilidad de elección de la opción 2 es una curva logística decreciente que tiende a k_2 cuando $\theta \to -\infty$ y que tiende a cero cuando $\theta \to \infty$.

Nuevamente, el análisis cualitativo de los campos direccionales y de las trayectorias en el espacio fase permite predecir el comportamiento de la probabilidad de elección de cada opción para distintas condiciones iniciales y distintos valores de la variable independiente. Análogamente a la situación en la que tanto a_1 como a_2 son positivas, los puntos de equilibrio del sistema que se obtiene cuando ambas ecuaciones diferenciales del modelo se hacen cero son $(0,0)$, $(0, k_2)$, $(0, k_1)$ y el punto de intersección entre las rectas

$$p_2 = \frac{a_1}{b_2}\left(1 - \frac{p_1}{k_2}\right) \text{ y } p_2 = k_2 - \frac{b_2 k_2}{a_2} p_1,$$

es decir, (p_{1i}, p_{2i}). Los signos de las derivadas sobre los ejes y sobre las rectas son también los mismos que en la situación anterior; es decir la órbita o trayectoria cruza verticalmente sobre el eje p_1 y sobre la recta que une los puntos $(k_1,0)$ y $(0, a_1 / b_1)$. La trayectoria cruza horizontalmente sobre el eje p_2 y sobre la recta que une los puntos $(0, k_2)$ y $(a_2 / b_2, 0)$.

Es importante notar que puesto que a_2 es negativo, el término a_2 / b_2 siempre está en la región negativa del eje p_1 y siempre será menor que k_1. Por otra parte aquellos puntos para los cuales p_2 es mayor que $k_2 - \frac{b_2 k_2}{a_2} p_1$, se encuentran en una región que no puede estar incluida en las soluciones posibles, ya que en ellos la probabilidad de elección de la opción 2 sería mayor que k_2 que es el máximo permitido para p_2.

Análogamente al caso en el que a_1 y a_2 son positivos, cuando uno de los parámetros, en este caso a_2 es negativo, dependiendo de los valores de los parámetros se pueden presentar diferentes casos. Estos casos se analizan a continuación:

CASO 6.

Cuando $\frac{a_1}{b_1} > k_2$ la región factible del plano fase para las soluciones al problema que nos interesa es $0 \le p_1 \le k_1$ y p_2 mayor que cero y además menor o igual que la recta $p_2 = k_2 - \frac{b_2}{a_2} k_2 p_1$ ya que si p_2 fuese mayor que esta recta se tendría que $p_2 > k_2$ y k_2 es el máximo valor que la probabilidad p_2 puede alcanzar. Las regiones factibles para este caso y los campos de direcciones se muestran en las figuras II.5.6a y II.5.6b, respectivamente.

Figura II.5.6a: Regiones en plano fase

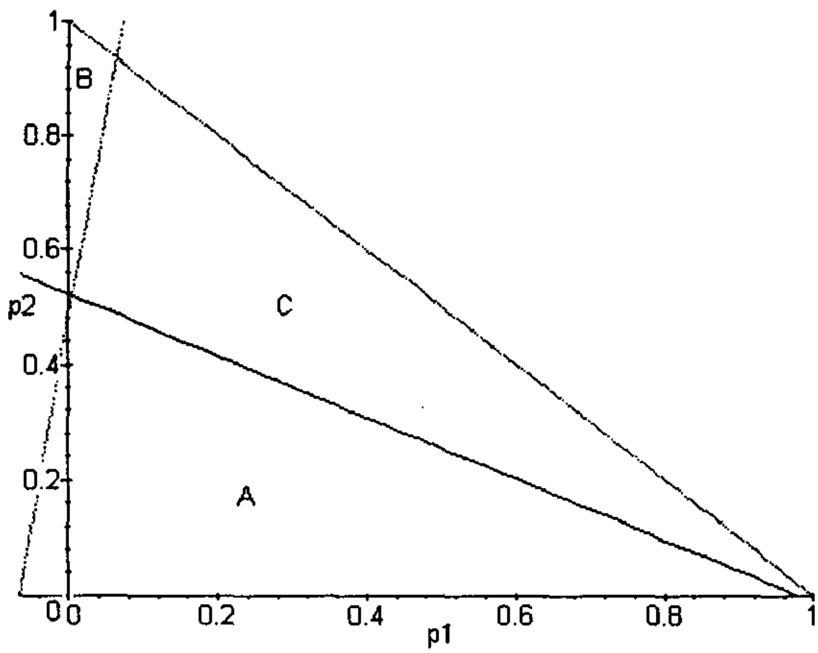
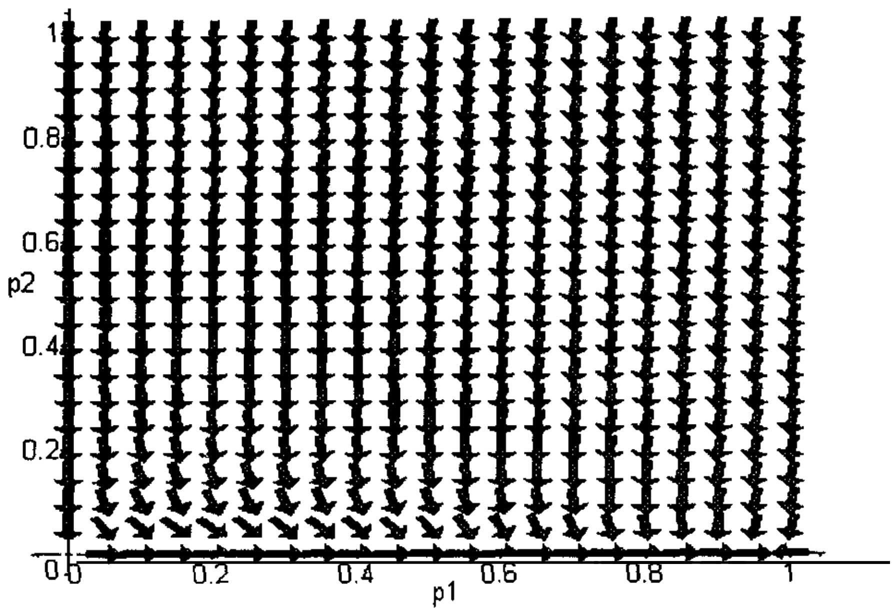


Figura II.5.6b: Campo de direcciones



Haciendo un análisis similar al que se ha venido realizando en los casos anteriores se encuentra que los puntos de equilibrio para este caso son los puntos $(0,0)$, $(0,k_2)$, $(k_1,0)$ y (p_{1i},p_{2i}); sin embargo, este último queda fuera de la región permitida para la solución. Los puntos $(0,0)$ *y* $(0,k_2)$ son inestables, mientras que el punto $(k_1,0)$ es asintóticamente estable. Esto indica que las trayectorias que se inician en cualquier punto de la región del espacio fase permitida convergerán hacia el punto $(k_1,0)$.

Las gráficas correspondientes a esta situación se muestran en las figuras II.5.6c y II.5.6d.

Figura II.5.6c: Curvas solucion

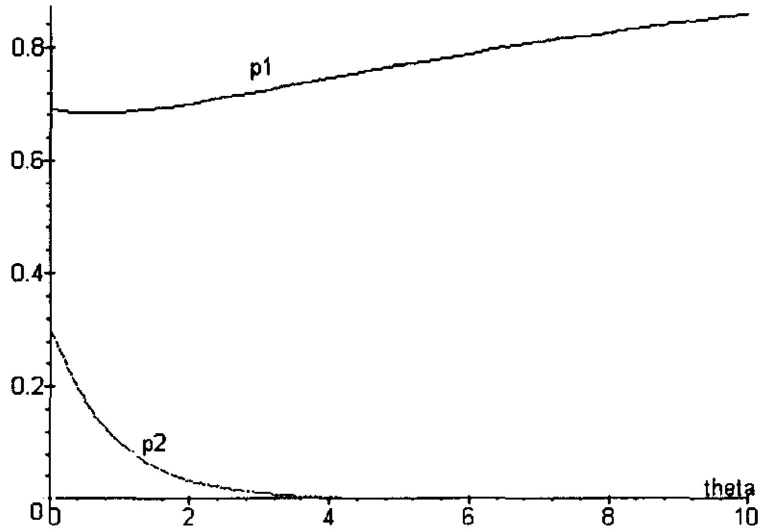
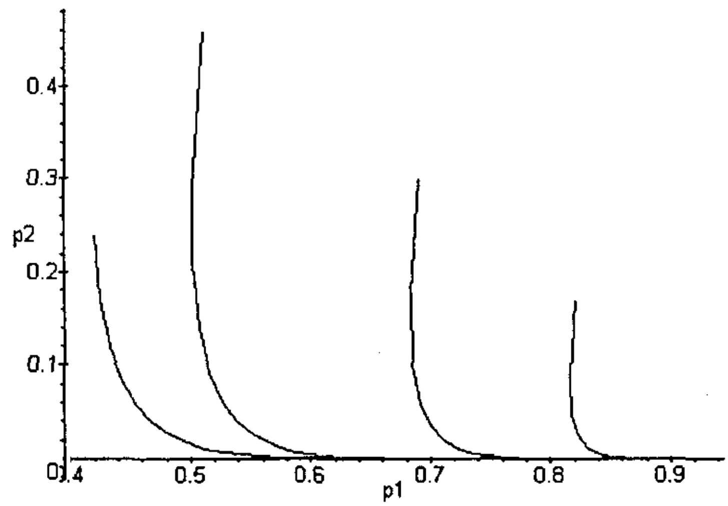


Figura II.5.6d: Trayectorias en plano fase



CASO 7.

Para aquellas situaciones en las que $k_2 > a_1 / b_1$, nuevamente el punto de intersección (p_{1I}, p_{2I}) queda fuera de la región permitida por las restricciones del modelo. Las regiones factibles y el campo de direcciones correspondientes a este caso se muestran en las figuras II.5.7a y II.5.7b.

Figura II.5.7a: Regiones en plano fase

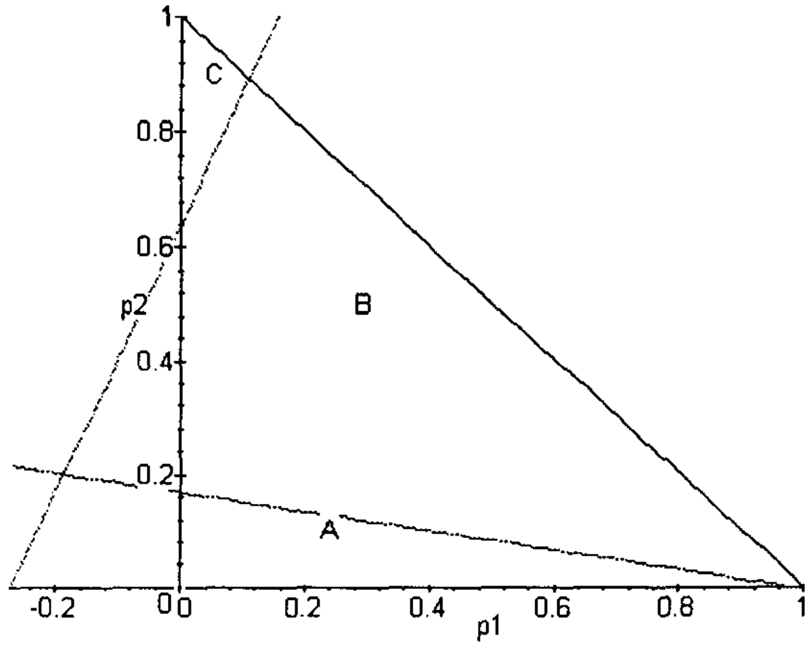
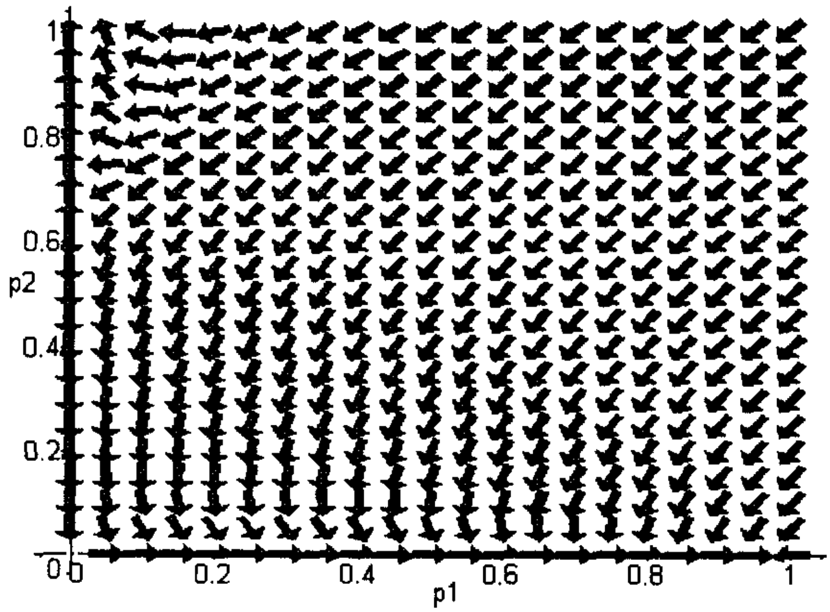


Figura II.5.7 b: Campo de direcciones



Los puntos de equilibrio en este caso son $(0,0)$, $(k_1,0)$ y $(0,k_2)$. Los puntos $(0,0)$ y $(0,k_2)$ son inestables mientras que el punto $(k_1,0)$ es asintóticamente estable.

El análisis de las trayectorias muestra que todas ellas convergen hacia el punto $(k_1,0)$ al aumentar θ.

En las figuras II.5.7c y II.5.7.d se muestran las curvas de solución y algunas trayectorias en el espacio fase que surgen en el análisis de este caso.

Figura II.5.7c: Curvas solucion

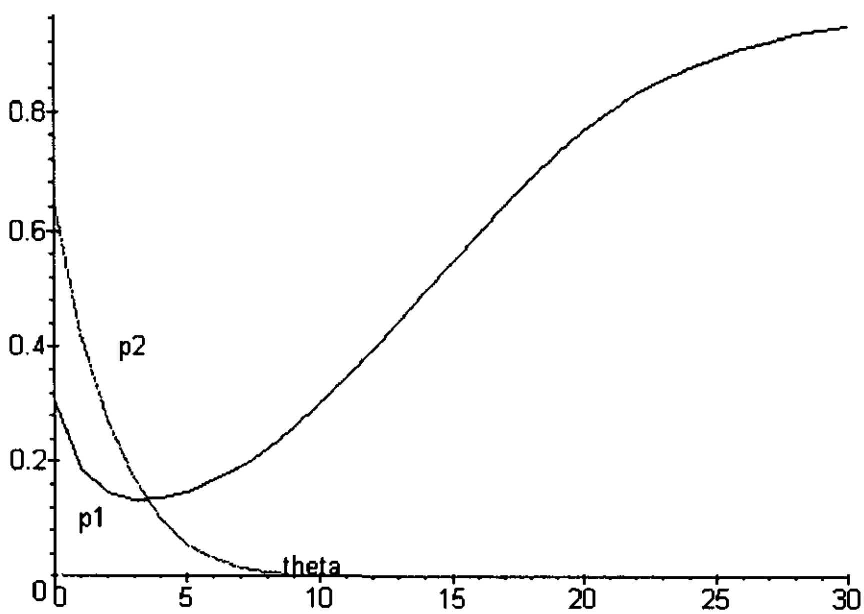
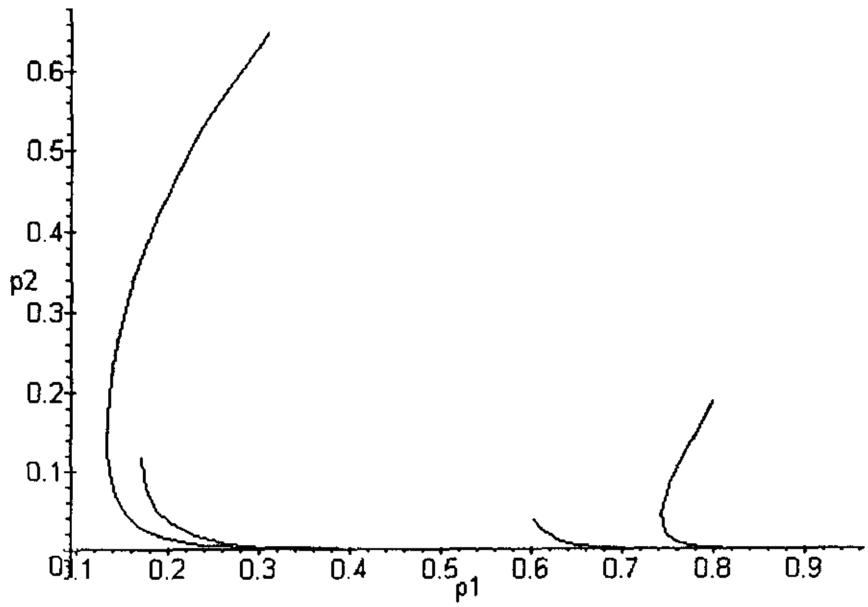


Figura II.5.7d: Trayectorias en plano fase



CASO 8.

Queda por analizar el caso en el que $\frac{a_1}{b_1} = k_2$.

En esta situación la región factible del espacio fase se limita a los puntos que son tales que cumplen con las condiciones $0 \le p_1 \le k_1$, $0 \le p_2 \le k_2$ y $p_1 + p_2 \le 1$. Nuevamente los puntos de equilibrio son los puntos $(0,0)$, $(k_1,0)$ y $(0, k_2)$ puesto que ahora el punto de intersección (p_{1I}, p_{2I}) coincide con el punto $(0, k_2)$. Las gráficas de las regiones y el campo de direcciones para este caso se muestran en las figuras II.5.8a y II.5.8b.

Figura II.5.8a: Regiones en plano fase

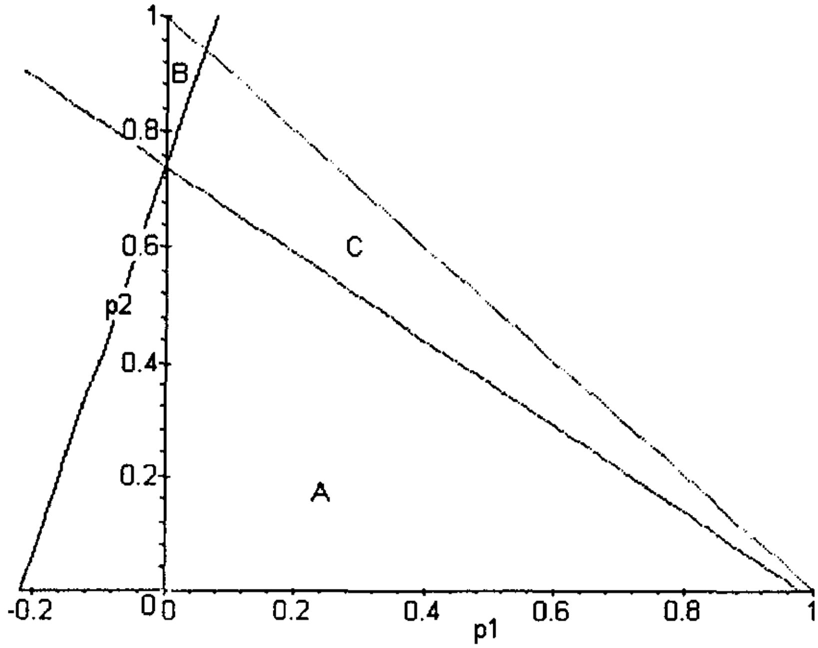
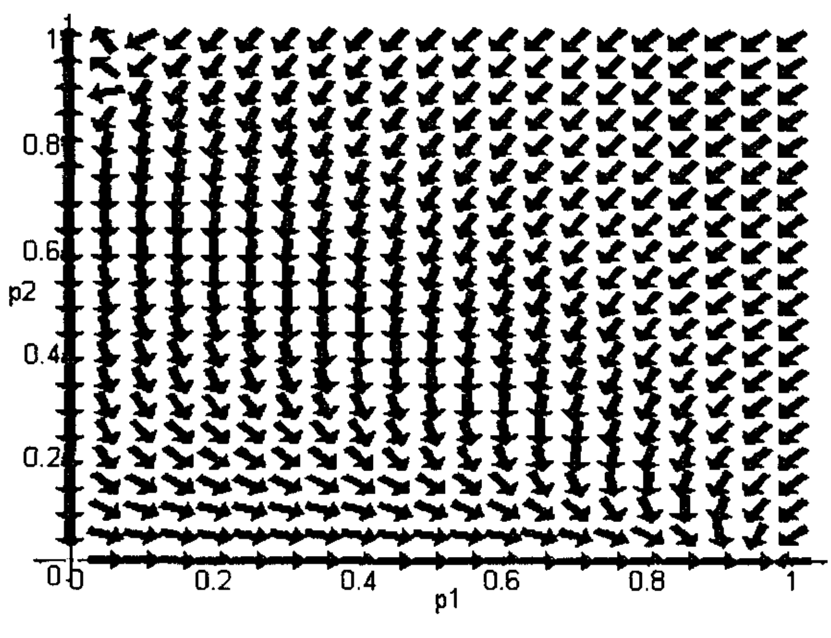


Figura II.5.8b: Campo de direcciones



El análisis de los signos de las derivadas, que se muestra en la figura II.5.8b indica que los puntos $(0,0)$ y $(0,k_2)$ son inestables, mientras que el punto $(k_1,0)$ es asintóticamente estable. Esto indica que independientemente de las condiciones iniciales, la trayectoria se acercará hacia el punto de equilibrio estable cuando crece θ.

Figura II.5.8c: Curvas solución

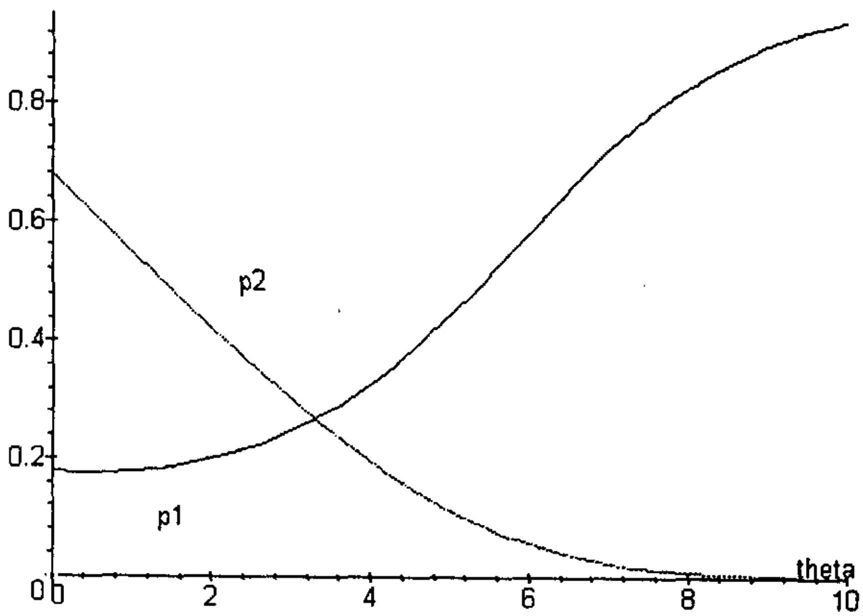
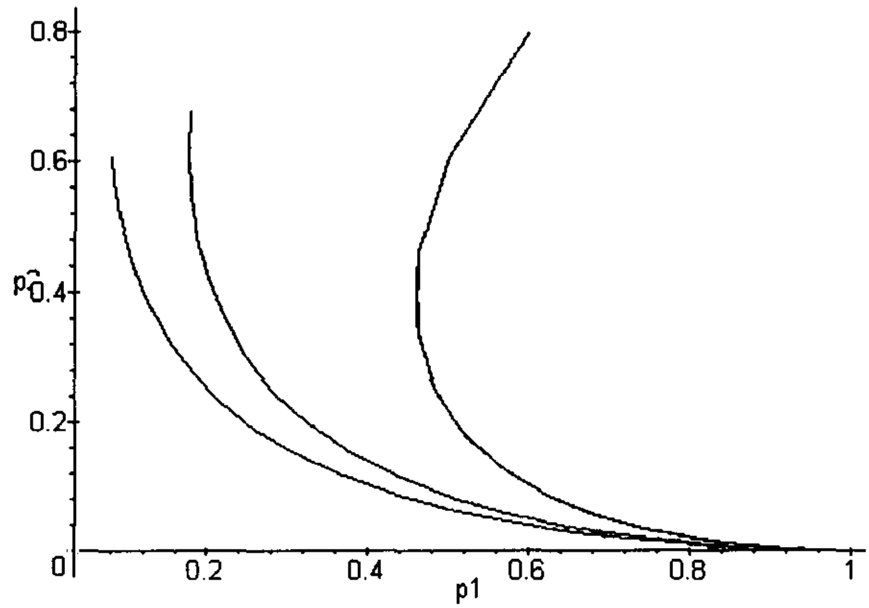


Figura II.5.8d: Trayectorias en plano fase



Un parámetro de interacción negativo

Una situación diferente se presenta cuando alguno de los parámetros b_1 o b_2 es negativo.

Supongamos que a_1, a_2 y b_1 son positivos y que b_2 es negativo. En este caso, el término correspondiente a la interacción de la opción 2 con la opción 1 en el modelo contribuye a aumentar la probabilidad de respuesta de la opción uno, en lugar de disminuirla, como sucedía en todos los casos anteriores. En términos del modelo de respuesta que aquí nos interesa, el hecho de que uno de los parámetros b sea negativo indica que la presencia de una de las opciones refuerza la elección de la otra. En otras palabras, estaríamos frente a un caso en el que el distractor no cumple su papel sino, por el contrario, actúa en apoyo de una opción específica.

El hecho de encontrar un parámetro b negativo (o ambos) al ajustar los parámetros del modelo, proporcionaría un indicador de que las opciones elegidas como distractores para el ítem no cumplen con la función para la que fueron introducidos al modelo y que habría que rediseñar las opciones de respuesta de manera que funcionen correctamente como distractores.

En esta situación se tiene que $a_2 / b_2 < 0$. Llevando a cabo el análisis cualitativo tal y como se ha hecho para todos los casos anteriores se encuentra que los puntos de equilibrio son nuevamente (0,0), $(k_1,0)$, $(0,k_2)$ y el punto de intersección (p_{1I}, p_{2I}).

Análogamente a la situación en la que uno de los parámetros a es negativo, se tiene que a_2 / b_2 está siempre en la parte negativa del eje p_1 y será siempre menor que k_1. Asimismo, los puntos para los cuales p_2 es mayor que $k_2 - \frac{b_2 k_2}{a_2} p_1$ no están permitidos ya que en ellos, la probabilidad de elección de la opción 2 sería mayor que k_2 que es el máximo permitido.

Haciendo nuevamente un análisis para las diferentes posibilidades de elección de los parametros se encuentran los casos siguientes.

CASO 9.

Si $a_1 / b_1 > k_2$, el análisis cualitativo muestra que los puntos (0,0), $(k_1,0)$ y $(0,k_2)$ son inestables. El único punto asintóticamente estable es el punto (p_{1l}, p_{2l}); sin embargo, este punto queda fuera de la región permitida para el modelo como se muestra en la figura II.5.9a. en la figura II.5.9b se muestra el campo direccional para este caso.

Figura II.5.9a: Regiones en plano fase

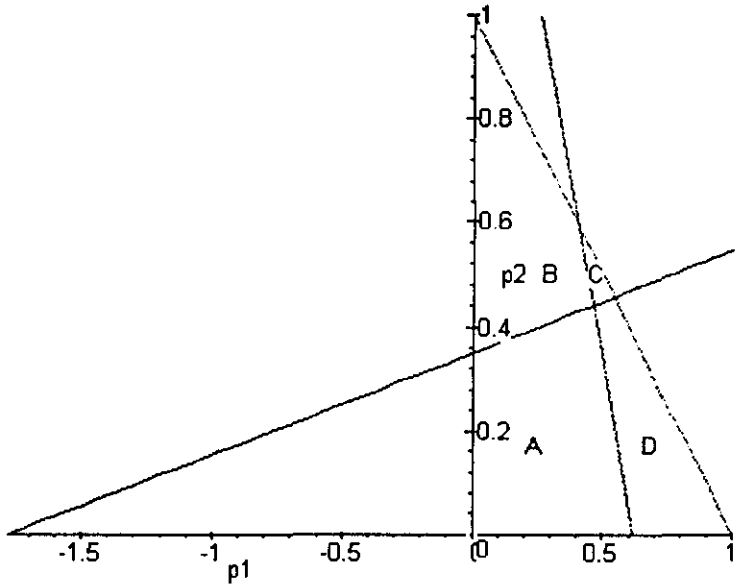
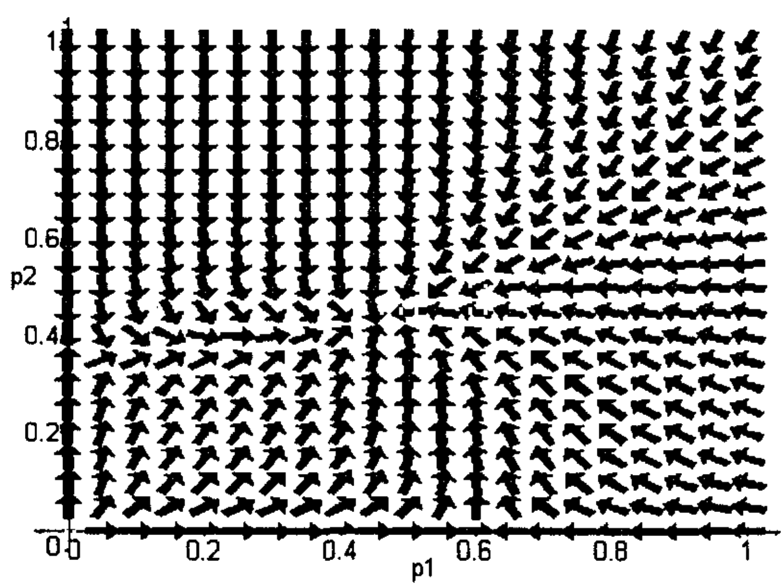


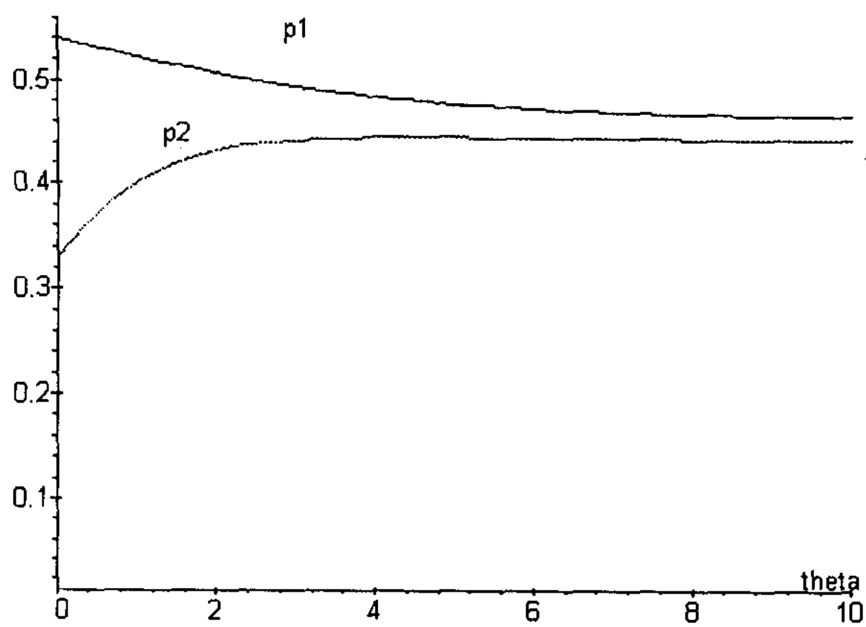
Figura II.5.9b: Campo de direcciones



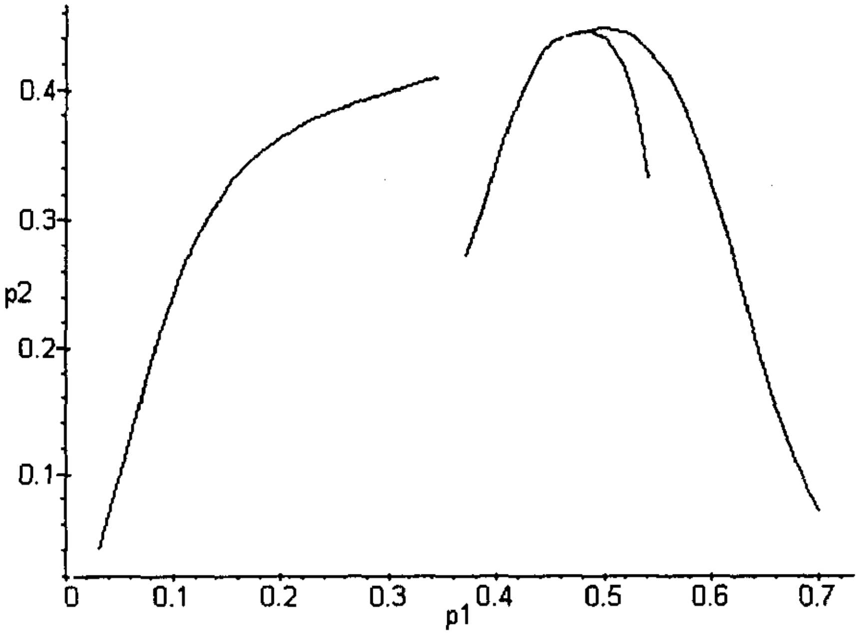
Las trayectorias que inician en la región A de la figura II.5.9a son tales que tanto la probabilidad de elección de la opción 1 como la probabilidad de elección de la opción 2 aumentan hasta que la trayectoria entra a la región B en la que la probabilidad de elección de la opción 2 crece mientras que la probabilidad de elección de la opción 1 decrece hasta alcanzar la recta $p_2 = k_2$ que es el límite de la región permitida; pero no hay puntos de equilibrio dentro de la región factible.

Las curvas solución para este caso se muestran en la figura II.5.9c y algunos ejemplos de trayectorias en el espacio fase se muestran en la figura II.5.9d.

Figura II.5.9c: Curvas solucion



FiguraII.5.9d: Trayectorias en plano fase



CASO 10.

Cuando $a_1 / b_1 < k_2$ se encuentran, al hacer el análisis cualitativo, las regiones permitidas para la solución y el campo de direcciones en cada una de ellas son las que se muestran en las figuras II.5.10a y II.5.10b respectivamente. El análisis cualitativo de las trayectorias indica que los puntos (0,0), $(k_1, 0)$ y $(0, k_2)$ son inestables mientras que el punto (p_{1l}, p_{2l}) es asintóticamente estable; sin embargo este punto queda nuevamente fuera de la región permitida para las soluciones. Esto nos indica, como se muestra en la figura II.5.10d, que muestra algunas órbitas en el espacio fase, que conforme θ aumenta, la probabilidad de elección de la opción 2 aumenta hasta salir de la región factible, mientras que la probabilidad de elección de la opción 1 desaparece.

Figura II.5.10a: Regiones en plano fase

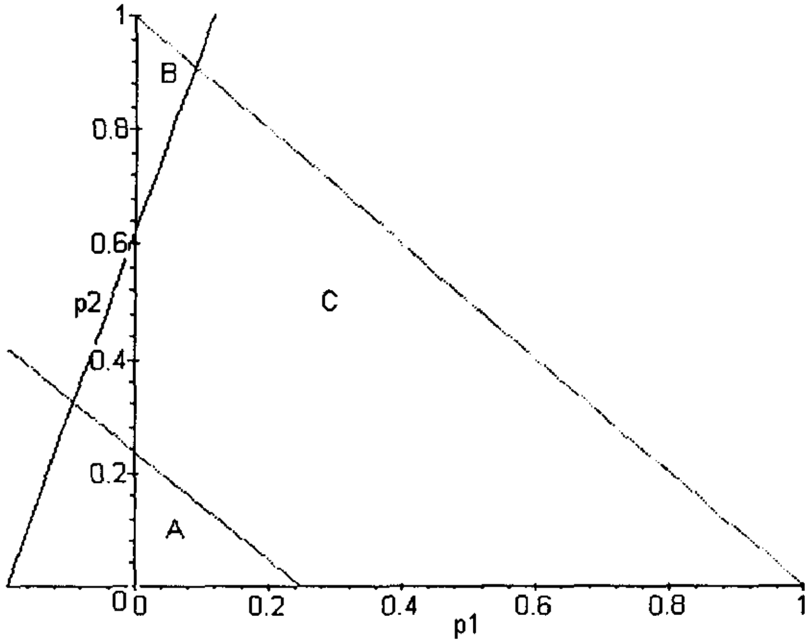
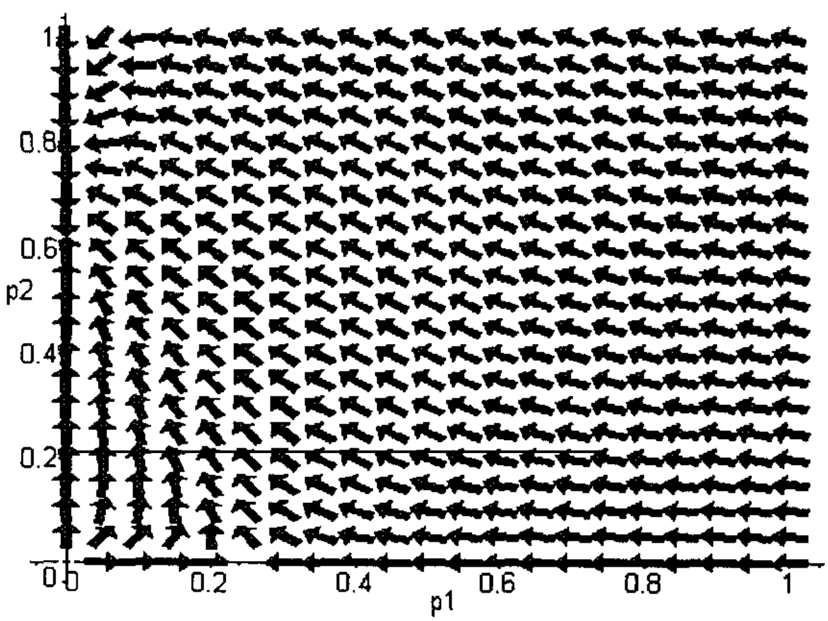


Figura II.5.10b: Campo de direcciones



En la figura II.5.10c se muestran las curvas solución para las pobabilidades de elección de cada una de las opciones en esta situación.

Figura II.5.10c: Curvas solucion

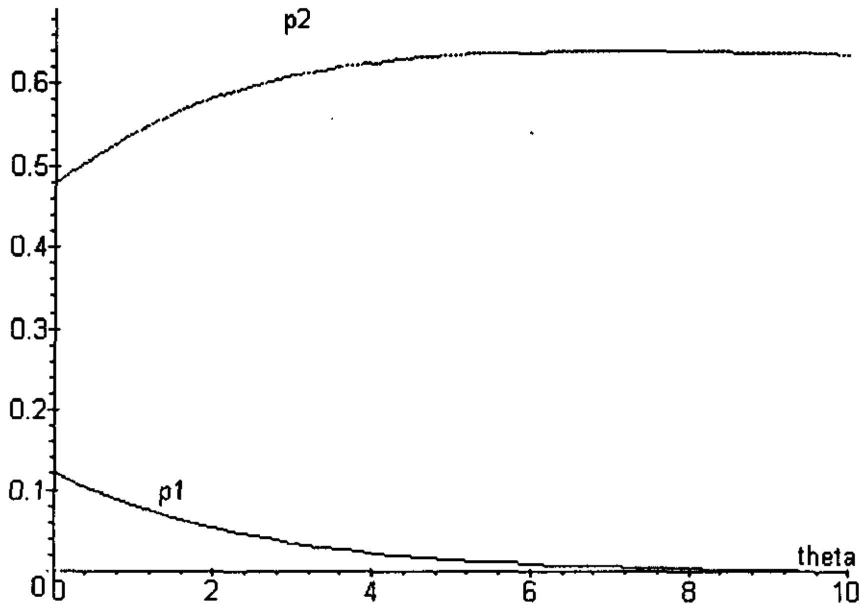
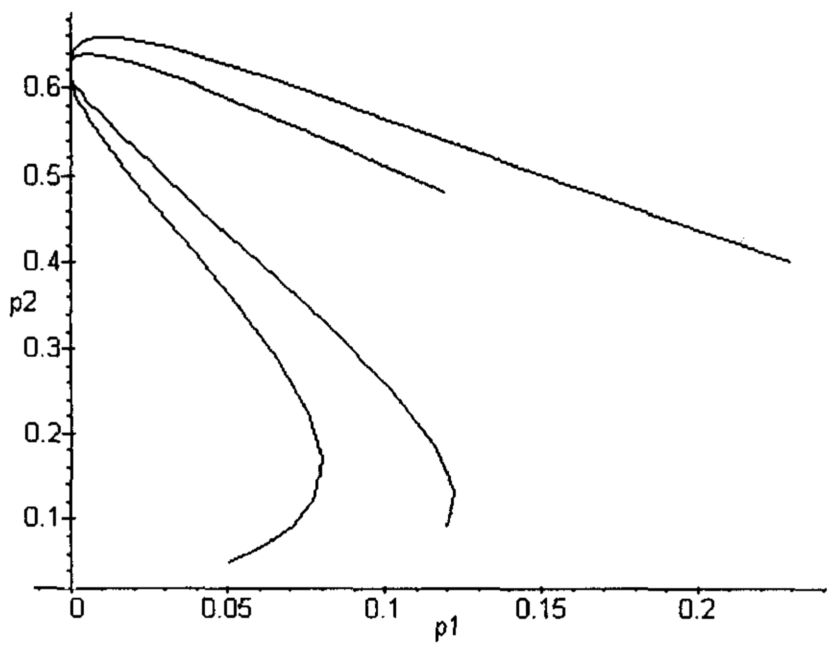


Figura II.5.10d: Trayectorias en plano fase



CASO 11.

Cuando $a_1 / b_1 = k_2$ se encuentra un comportamiento similar al del caso anterior. El único punto de equilibrio estable es el punto $(0, k_2)$. Las gráficas correspondientes al análisis cualitaivo de este caso se muestran en las figuras II.5.11a, II.5.11b, II.5.11c y II.5.11d.

Figura II.5.11a: Regiones en plano fase

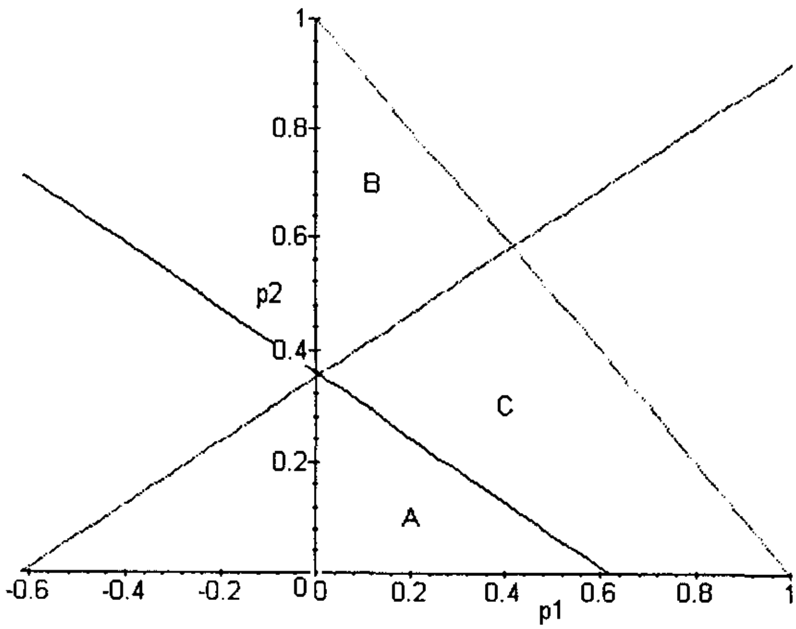


Figura II.5.11b: Campo de direcciones

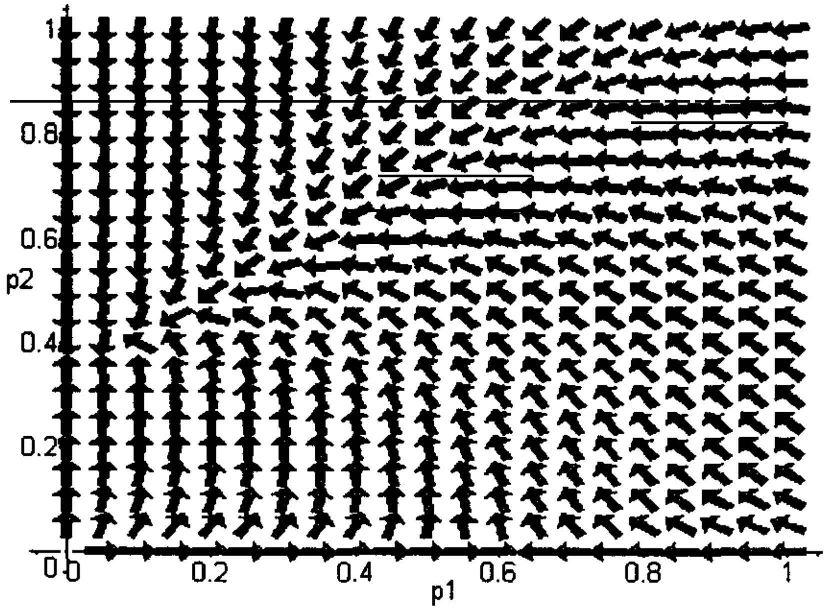


Figura II.5.11c: Curvas solucion

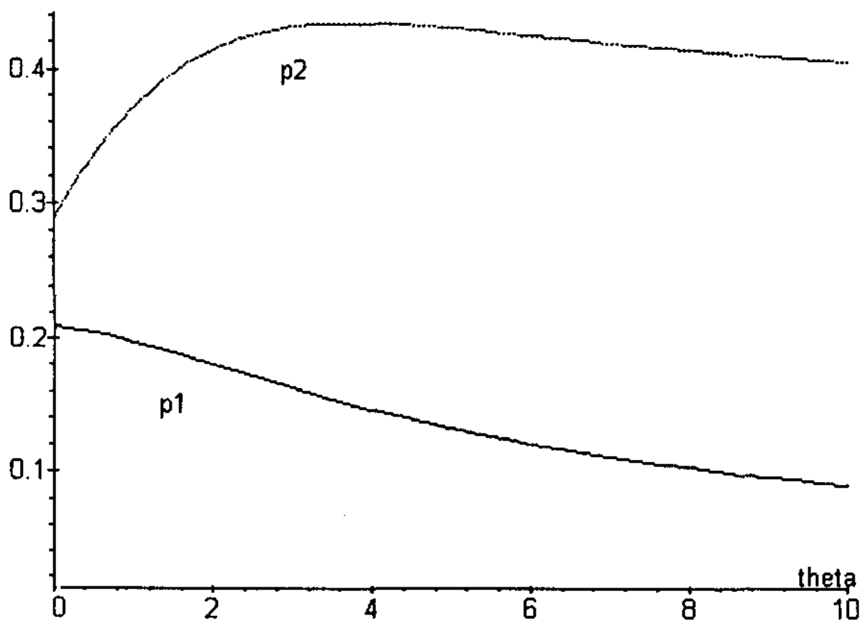
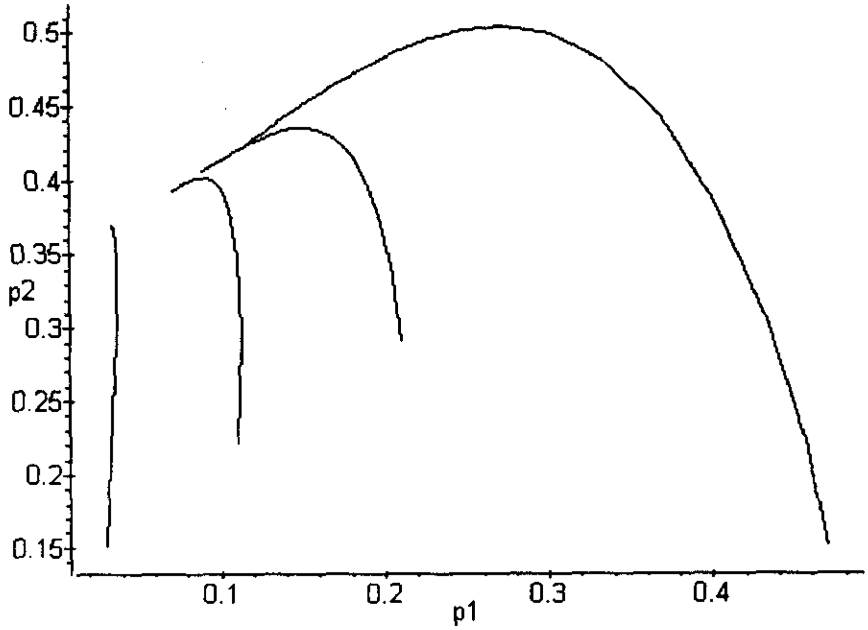


Figura II.5.11d: Trayectorias en plano fase



CASO 12.

En el caso en que los parámetros de accesibilidad tienen signos diferentes puede presentarse también alguna situación en la que uno de los parámetros de interacción sea negativo.

Supongamos que $a_1 > 0$, $a_2 < 0$, $b_1 < 0$ y $b_2 > 0$, con la restricción de que $a_1 / b_1 > k_2$. En esta situación $a/b_1 < 0$ siempre, por lo que es imposible que la restricción se cumpla. Dadas estas condiciones, no hay región factible para las soluciones del modelo y esto implica una contradicción debida al mal diseño de los distractores.

CASO 13.

Cuando los parámetros del ítem son tales que $a_1 > 0$, $a_2 < 0$, $b_1 < 0$ y $b_2 > 0$, pero con la restricción de que $a_1 / b_1 < k_2$ se tiene una situación que se cumple siempre, dado que $a_1/b_1 < 0$ y $a_1 > 0$, $a_2 / b_2 < 0$. En esta situación, las regiones permitidas para la solución son las que se muestran en la figura II.5.13a. En la figura II.5.13b se muestra el campo de direcciones correspondiente. Las trayectorias en el espacio fase, que se muestran en la figura II.5.13c, se dirigen hacia un punto de equilibrio, el punto (k_1, 0). Las curvas solución pueden, en apariencia, mostrar una situación normal y esperada (figura II.5.13d); sin embargo la probabilidad de elección de la segunda opción decrece muy rápidamente y pueden presentarse casos en los que parte de la trayectoria salga de la región permitida para la solución, lo que nuevamente indica un problema en el diseño de los distractores del ítem.

Figura II.5.13a: Regiones en plano fase

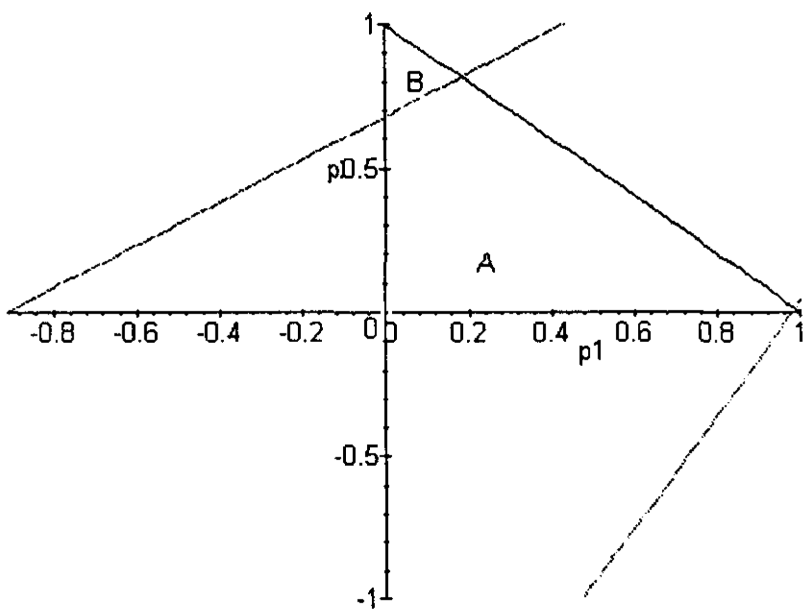


Figura II.5.13b: Campo de direcciones

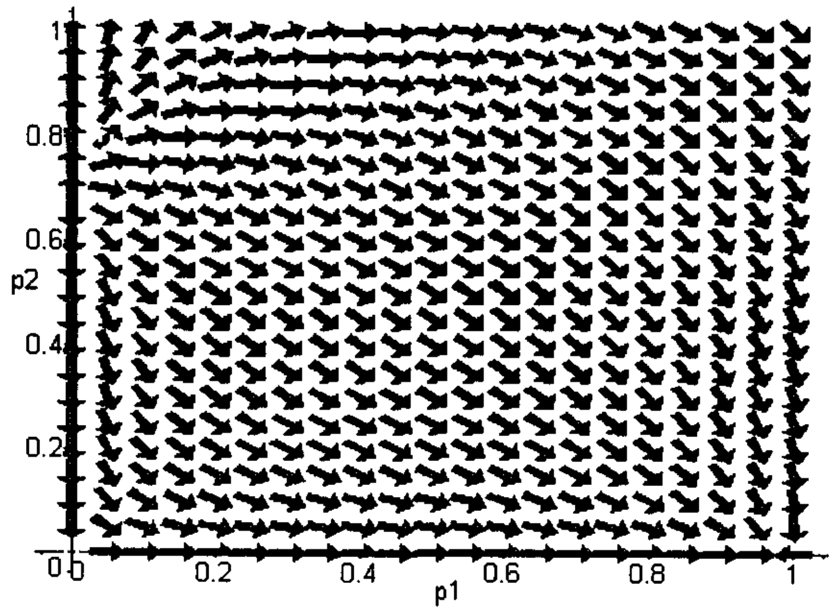


Figura II.5.13c: Curvas solucion

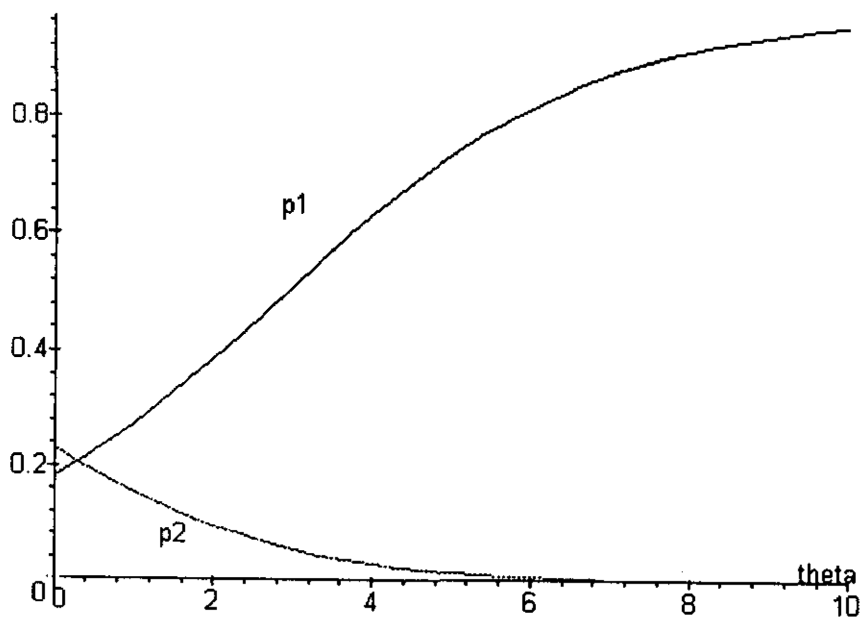
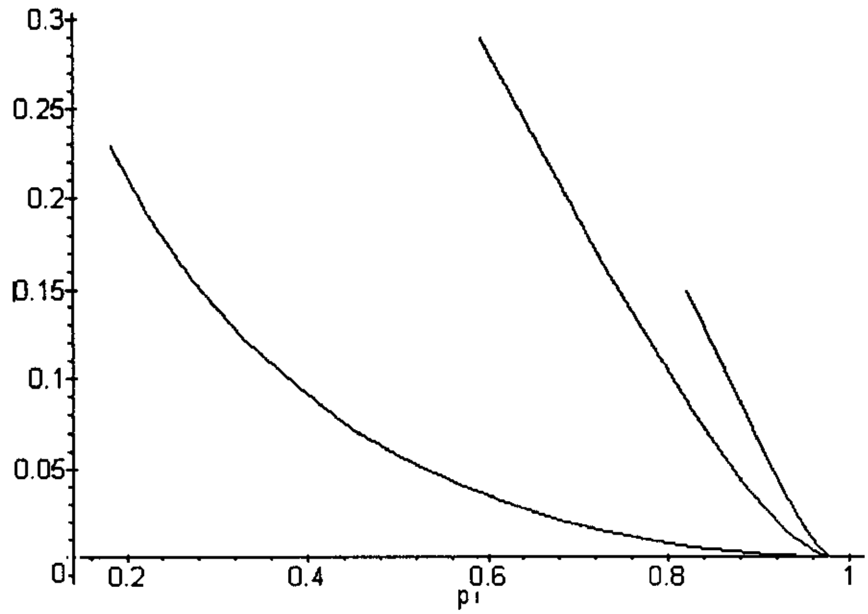


Figura II.5.13d: Trayectorias en plano fase



CASO 14.

Si los parámetros que se obtienen para el ítem son tales que $a_1 > 0$, $a_2 < 0$, $b_1 > 0$ y $b_2 < 0$, con la restricción $a_1 / b_1 > k_2$, las figuras correspondientes a las regiones permitidas en este caso del modelo y a los campos de direcciones se muestran en las figuras II.5.14a y II.5.14b. En ellas puede observarse una situación más compleja. Como se muestra en las figuras II.5.14c y II.5.14d, para algunas condiciones iniciales la elección de la primera opción se ve reforzada fuertemente por la acción del parámetro de interacción negativo; pero para la mayoría de las posibles condiciones iniciales dentro del modelo, las curvas solución se salen de la región permitida. Nuevamente, este tipo de comportamiento es un indicador de un problema de diseño de los distractores del ítem.

Figura II.5.14a: Regiones en plano fase

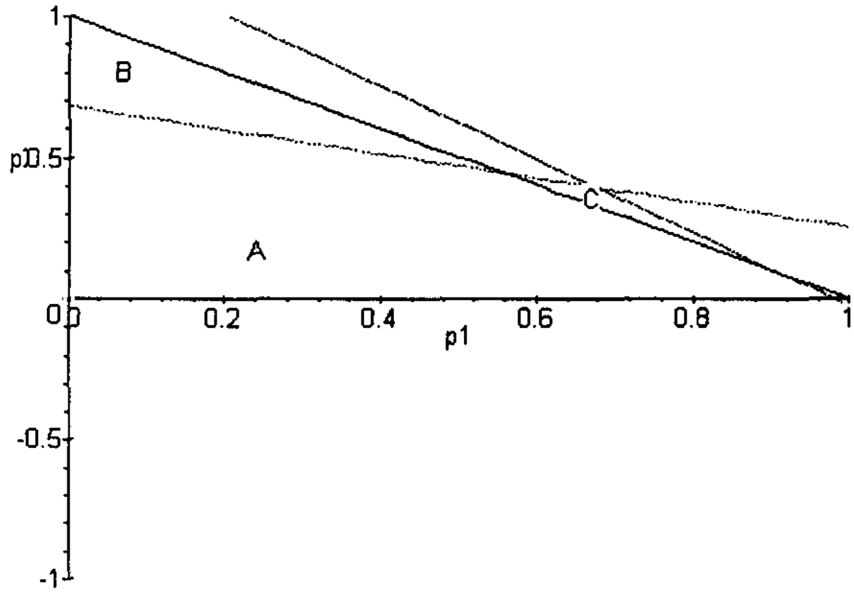


Figura II.5.14b: Campo de direcciones

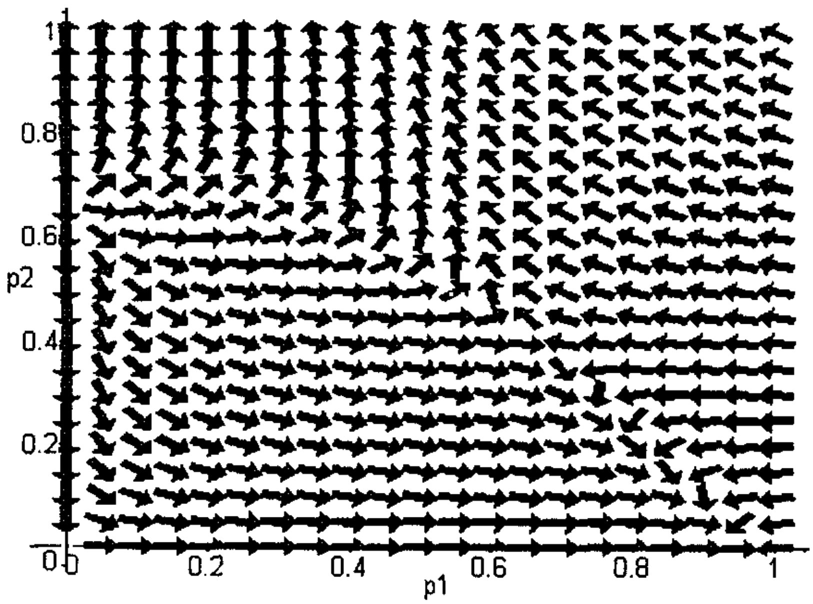


Figura II.5.14c: Curvas solucion

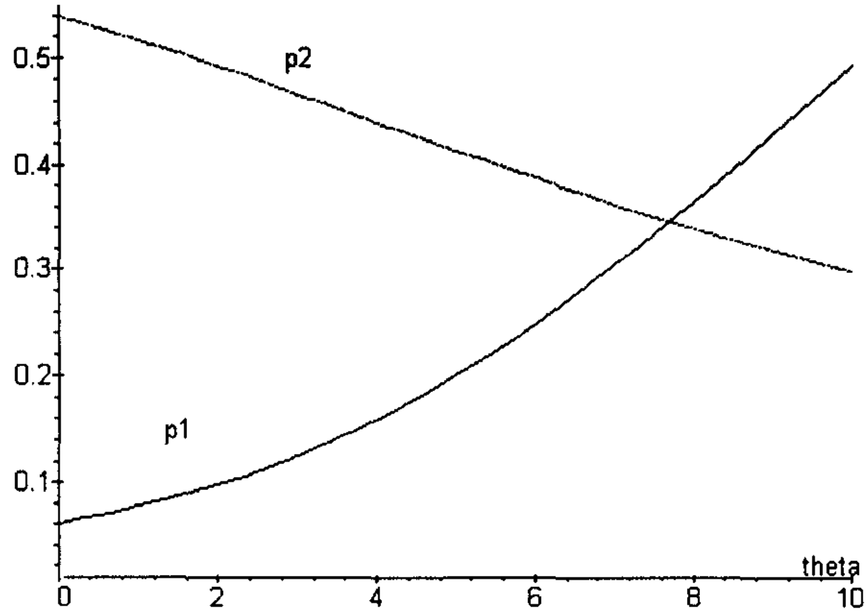
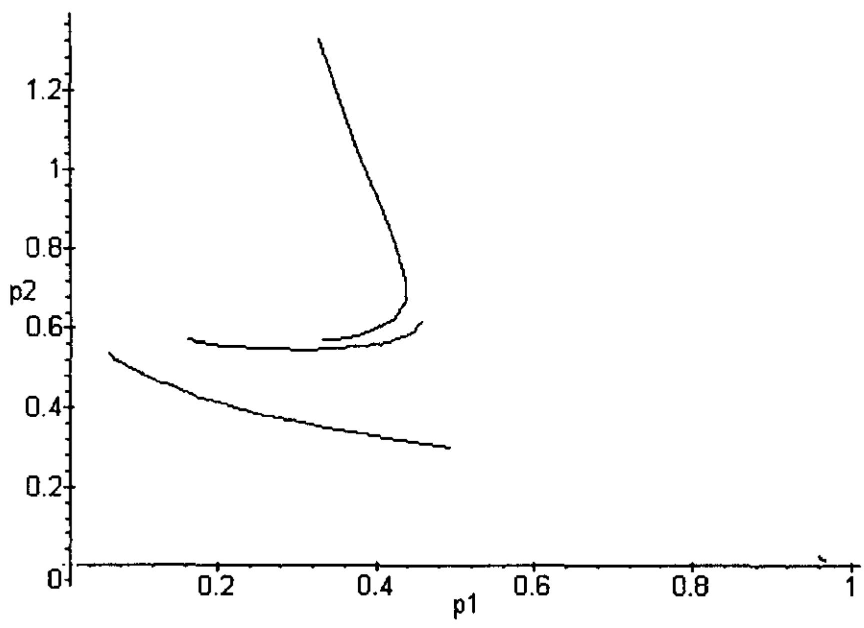


Figura II.5.14d: Trayectorias en plano fase



Consideraciones que se desprenden del análisis cualitativo

El análisis cualitativo nos muestra que el modelo propuesto puede descomponerse en tres familias: la primera para el caso en que todos los parámetros del sistema son positivos, otra para la cual uno de los parámetros de accesibilidad, que se ejemplificó con el parámetro a_2, es menor que cero y la última para el caso en el que uno de los parámetros de interacción b es negativo.

A lo largo de todo el análisis se encontró que independientemente del valor de los parámetros del sistema de ecuaciones diferenciales, los casos correspondientes a la segunda familia del modelo presentan trayectorias que convergen siempre a la solución de equilibrio $(k_1,0)$. La forma en la que esta convergencia se da depende del valor de los parámetros y de las condiciones iniciales, pero el resultado final es que conforme θ crece, la opción que se elige es la primera.

El caso de la primera familia del modelo, aquélla para la cual todos los parámetros del sistema son positivos, presenta un comportamiento más versátil, ya que, dependiendo de los valores de los parámetros, pueden predecirse casos en los que es posible obtener coexistencia de las probabilidades de elección de cada una de las opciones planteadas para el ítem y casos en los que domina la probabilidad de elección de una de las dos opciones.

En cuanto a la tercera familia del modelo, los casos posibles muestran prevalencia de la opción favorecida. Lo interesante de los resultados que se obtienen para los distintos casos de esta familia, y que es un aspecto que hace a este modelo único comparado con otros modelos de medida propuestos, radica en que el modelo proporciona una manera de determinar si las opciones de respuesta para el ítem están diseñadas de manera que funcionen como distractores adecuados.

El análisis cualitativo es, como se ha mostrado hasta aquí, una herramienta poderosa para la comprensión del comportamiento de las soluciones al modelo, pues permite predecir el comportamiento de las soluciones sin necesidad de resolverlas explícitamente.

Dado que todos los casos que se han analizado aquí puede observarse que la forma en la que las curvas se acercan o se alejan de los puntos de equilibrio depende de los parámetros específicos del modelo y de las condiciones iniciales, el modelo que se ha desarrollado en este trabajo da cuenta de una gran variedad de posibilidades para las formas de las curvas que caracterizan las probabilidades de elección de cada una de las opciones de los ítems y por ello, la posibilidad de aplicación del modelo para el análisis de los ítems de un test y para el diagnóstico se amplía considerablemente.

Las consideraciones que se han hecho hasta aquí acerca del modelo de medida son únicamente de índole matemática. En la siguiente sección se hace un análisis de las familias del modelo y de los casos que se presentan en cada una de ellas para distintos valores de los parámetros, pero en términos del modelo de medida que aquí nos interesa.

Sección II.6 Interpretación del Modelo en términos de su aplicación al diseño de ítems

El modelo de medida que se desarrolló en las secciones anteriores consiste de una familia de submodelos que surgen de las distintas posibilidades de valores que los parámetros pueden tomar. Cada uno de estos submodelos permite modelar situaciones diferentes.

El primer submodelo se presenta cuando ambos parámetros de accesibilidad a_1 y a_2 son positivos y corresponde al caso en que los ítems pueden medir preferencias o actitudes. En este caso, en ausencia de interacción entre las posibles opciones de respuesta, la probabilidad de elección de cada una de ellas crece al aumentar θ, hasta alcanzar el valor máximo que cada una de estas probabilidades puede alcanzar. Se tiene además la restricción de que para cualquier valor de θ la suma de la probabilidad de elección de la opción 1 más la probabilidad de elección de la opción 2 no debe exceder a uno y, por lo mismo, la suma de los valores máximos de cada una de ellas tampoco excede a uno.

Cuando alguno de los parámetros de accesibilidad es negativo, se tiene un modelo diferente, adecuado para el diseño de ítems que pretenden medir conocimiento, en el que, en ausencia de interacción entre las opciones de respuesta, la probabilidad de elección de una de ellas crece al aumentar θ mientras que la probabilidad de elegir la otra, aquella para la cual el parámetro a es negativo, decrece.

Si, por último, alguno de los parámetros de interacción, b_1 o b_2, es negativo, el modelo, como se ha visto en la sección anterior, implica el refuerzo a la respuesta de una de las opciones del ítem, puesto que el término de interacción tendría signo positivo, en lugar de negativo y ello contribuiría a que la probabilidad de elección correspondiente aumentara más rápidamente y, por ello, permite hacer una evaluación de los distractores, que supuestamente debieran disminuir la probabilidad de elección de la otra posible opción de respuesta, y decidir si el diseño de los distractores es el adecuado.

El primer submodelo permite representar ítems diseñados para medir actitudes, preferencias o conocimientos. En el caso de ítems que intentan medir conocimientos, las opciones de respuesta a este tipo de ítems presentan estrategias diferentes de acercamiento a un problema en particular.

Con el fin de entender, en términos del diseño de ítems, las distintas situaciones que pueden representarse con este modelo, podemos ejemplificar este submodelo mediante un ítem dentro del ámbito del conocimiento matemático que puede resolverse mediante estrategias diferentes y mediante un ítem destinado a medir preferencias. En el caso del ítem que mide conocimiento matemático la elección de la estrategia depende, por una parte, del nivel de competencia en la materia del sujeto que responde y por otra del tipo de problema presentado en el ítem. Por ello, independientemente de que el conocimiento matemático aumente, existirán problemas para los que siempre habrá cierta probabilidad de elección de cada una de las estrategias. Esto ocurre, por ejemplo, en el caso de ítems que piden la elección de la metodología a seguir en la solución algebraica de una ecuación cuadrática. Por otra parte, un ejemplo de un ítem destinado a medir preferencias puede ser alguno de los que se emplean con frecuencia en los tests vocacionales tal como preguntar qué tipo de lectura se elige con mayor frecuencia.

A continuación se analiza en términos educativos el significado de los diferentes casos que surgen cuando los parámetros toman distintos valores en este submodelo. Supongamos que el ítem en cuestión presenta una ecuación algebraica cuadrática para la cual se pide elegir la estrategia de solución y que las opciones de respuestas son: 1) estrategia de uso de fórmula, 2) estrategia de factorización, 3) sustitución de valores.

Con el fin de que el seguimiento del modelo se haga con mayor claridad, se discutirán los diferentes escenarios en el orden en que se presentaron en la sección correspondiente al análisis matemático.

Parámetros de accesibilidad positivos

CASO 1.

Este caso, en el que $a_1 / b_1 > k_2$ y $a_2 / b_2 > k_1$, puede ocurrir cuando la accesibilidad de las distintas opciones es aproximadamente del mismo orden de magnitud que los parámetros de interacción, cuando la accesibilidad de las opciones es pequeña (lo que implica que a_1 y a_2 son grandes) y la interacción correspondiente es pequeña, cuando las probabilidades máximas de elección son pequeñas, cuando las probabilidades de respuesta máxima son pequeñas o alguna combinación de estas situaciones.

La elección de estrategia de solución del ítem depende, como se ha mencionado, por una parte, del nivel de competencia del sujeto en el manejo del álgebra, pero también del tipo de problema planteado. Independientemente de que la competencia aumente, siempre existirá cierta probabilidad de elección de cada una de las estrategias, puesto que una mayor competencia implica mayor flexibilidad en la elección de la estrategia a seguir frente a un problema dado, así que es posible que exista siempre una probabilidad diferente de cero de elección para cada una de las estrategias.

Para valores pequeños de θ se tiene una cierta probabilidad de elección de estrategia. Conforme θ aumenta las probabilidades cambian hasta que se alcanza asintóticamente el punto de equilibrio. En él, independientemente del valor de θ la posibilidad de elegir cualquiera de las estrategias está dada por los valores de probabilidad de equilibrio. Las condiciones iniciales están determinadas por la naturaleza del problema a resolver. Las diferentes regiones en el plano fase muestran cómo, a medida que el conocimiento algebraico avanza se alcanza cierta flexibilidad en el uso de las estrategias.

Un análisis de las distintas regiones del plano fase permite interpretar los resultados que el modelo produce. El tipo de curva que se obtiene de la solución del modelo depende de la región del plano fase en la que se encuentran las condiciones iniciales. Recordemos que las condiciones iniciales juegan, en este modelo, el papel de parámetros. La información que las condiciones iniciales proporcionan es interesante pues permite caracterizar el nivel de desarrollo de la población con la que se utiliza el ítem o el examen. Si las condiciones iniciales son tales que el valor de la probabilidad de responder correctamente al ítem es alto, se tiene un indicador de que los sujetos con los que se utiliza el ítem tienen cierto nivel de avance en aquéllo que el ítem pretende medir. En cambio, si dicha probabilidad es baja, se puede inferir que la muestra que responde el ítem es débil en el rasgo a medir. Tomados conjuntamente una serie de ítems en un test, los valores de las condiciones iniciales permiten localizar las áreas que requieren de mayor atención y proveen nuevos elementos de diagnóstico.

En el caso que aquí nos ocupa, el tipo de curva que se obtiene se presenta en la figura II.6.1a. Esta figura muestra que si las condiciones iniciales se encuentran cerca del origen (región A de la figura II.6.1a), ambas probabilidades aumentan al crecer θ (Figura II.6.1b). Si las condiciones iniciales se encuentran en la misma región pero en puntos cercanos a $(0, k_2)$ para valores pequeños de θ el sujeto utiliza casi exclusivamente la estrategia 2 de solución al problema. Conforme θ aumenta, la probabilidad de usar la otra estrategia empieza a crecer, aunque la probabilidad de elección de la segunda sigue siendo dominante. Llega un momento, cuando la órbita cruza la recta que une los puntos $(0,\ k_2)$ y $(a_2/b_2\ ,\ 0)$, en el que se empieza a favorecer el uso de la estrategia 1, hasta que la situación se estabiliza en el punto de equilibrio. La misma situación se da en el caso de que las condiciones iniciales se encuentren cerca del punto $(k_1\ ,\ 0)$ sólo que en este caso la estrategia favorecida sería la 2 en lugar de la 1.

Figura II.6.1a: Regiones en plano fase

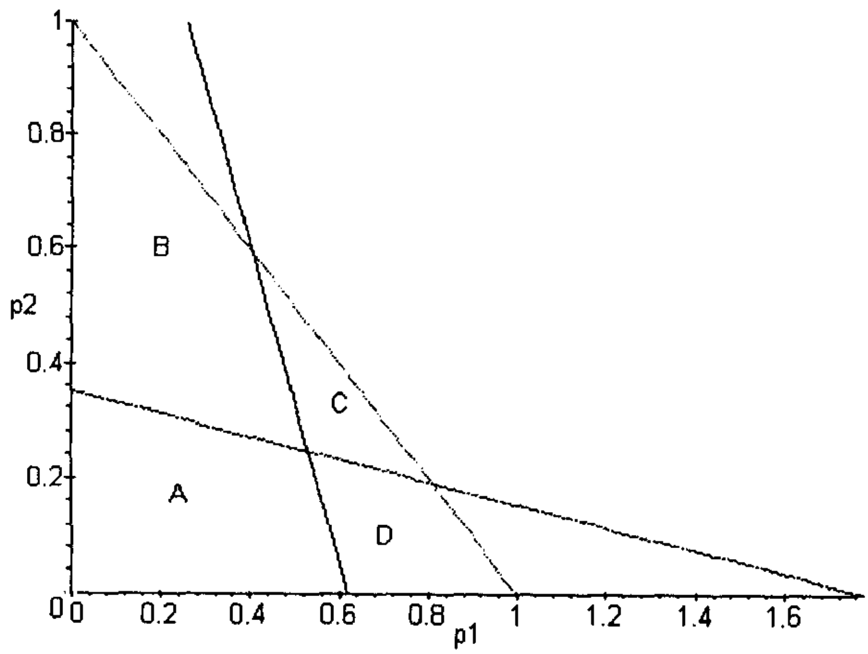
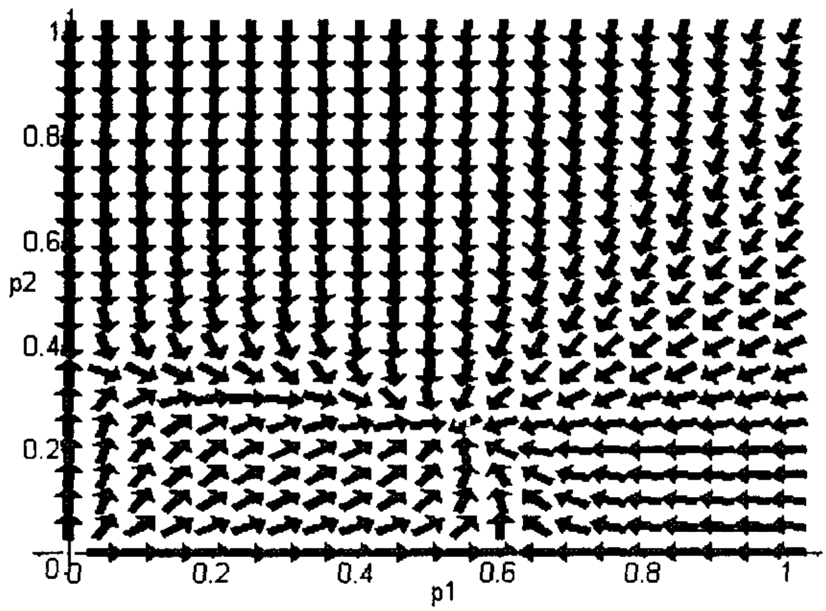


Figura II.6.1b: Campo de direcciones



Si las condiciones iniciales se encuentran dentro de la región permitida en la región D del plano fase que se muestra en la figura II.6.1a, para valores pequeños de θ se favorece la elección de la estrategia 2 y conforme θ aumenta la probabilidad de elegir la estrategia 2 disminuye y la de elegir la estrategia 1 aumenta hasta llegar al punto de equilibrio. La misma situación ocurre para condiciones iniciales en la región B de la misma figura aunque en este caso la estrategia favorecida inicialmente es la primera (Figura II.6.1b).

Si el estado inicial se encuentra dentro de la región C de la figura II.6.1a, la probabilidad de elección de ambas estrategias disminuye conforme θ aumenta, para un cierto intervalo de θ, como se muestra en la figura II.6.1b. Esta situación puede interpretarse como la existencia de casos en los que al haber un aumento en el conocimiento se dejan de usar las estrategias que se utilizaban anteriormente de manera mecánica sin entender las razones de su uso; el aumento en el conocimiento lleva entonces al punto de equilibrio en el que la elección se hace de manera razonada.

Para un ítem consistente en la elección de estrategia para resolver una ecuación cuadrática, este caso se presentaría cuando la ecuación es factorizable, pero la factorización no es fácil de distinguir.

CASO 2.

Cuando $a_1 / b_1 < k_2$ y $a_2 / b_2 < k_1$ se tiene el caso de distractores cuya accesibilidad es más o menos grande y c□yo parámetro de interacción también es grande.

En este caso es posible llegar a un punto de equilibrio como el que se analizó en el caso anterior, pero esta situación es altamente improbable e inestable (Figuras II.6.2a y II.6.2b).

Figura II.6.2a: Regiones en plano fase

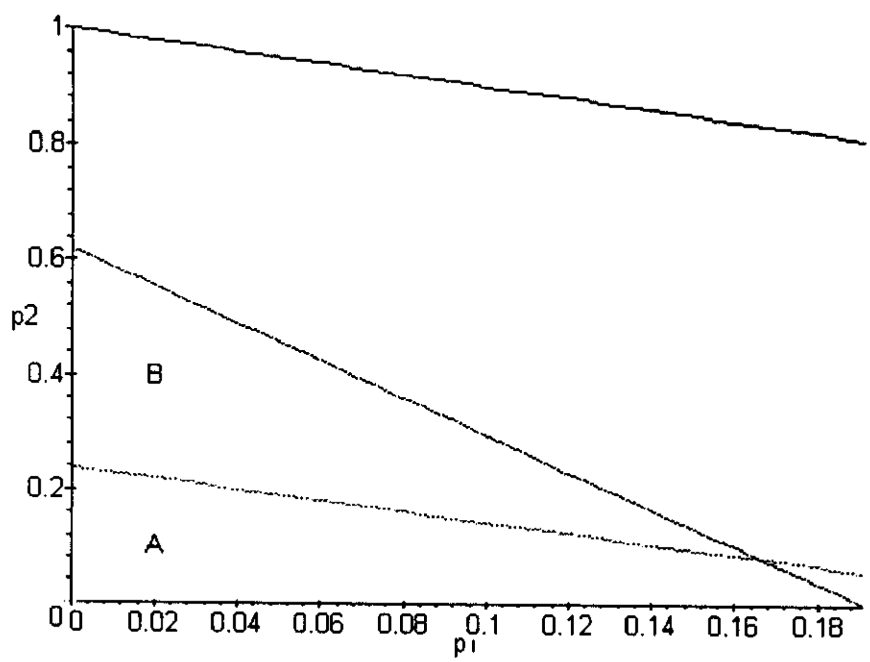
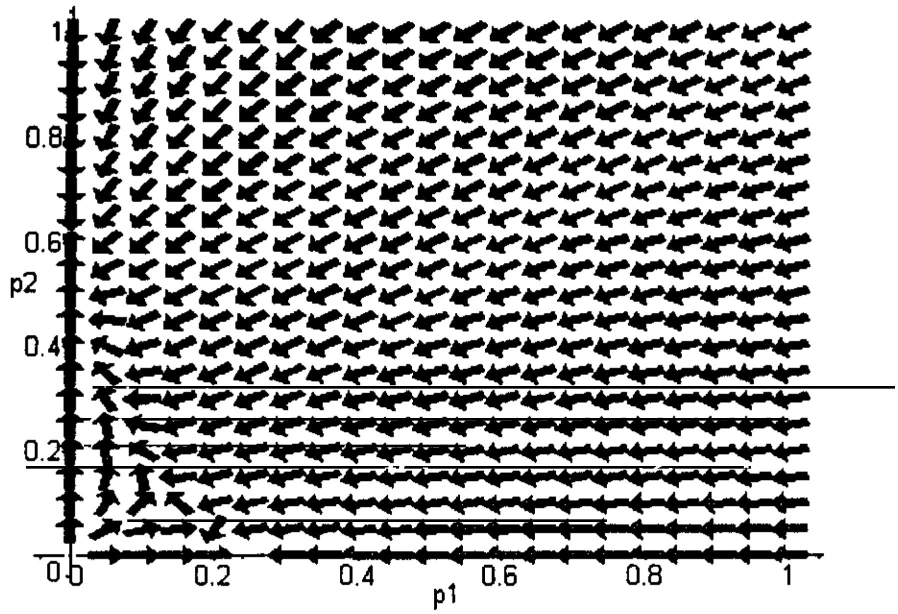


Figura II.6.2b: Campo de direcciones



Para este tipo de problema, dependiendo del valor de las condiciones iniciales, al aumentar el valor de θ la probabilidad de elección de una de las dos opciones dominará y alcanzará su máximo permitido de probabilidad, mientras que la probabilidad de elegir la otra opción se hace finalmente cero.

Si el estado inicial se encuentra en la región A, de la figura III.6.2a, ambas probabilidades aumentan conforme θ aumenta hasta que se alcanza un valor crítico en el que empieza a predominar una de las dos estrategias (Figura II.6.2b).

En el caso del ejemplo de la ecuación cuadrática esta situación puede ocurrir cuando la ecuación es factorizable, aunque la factorización no es muy simple. Para valores de θ pequeños la probabilidad de que los estudiantes utilicen la factorización como estrategia es pequeña, pero para valores mayores de θ esta estrategia será más favorecida.

CASO 3.

Si $a_1 / b_1 > k_2$ y $a_2 / b_2 < k_1$ se presenta una situación en la que la accesibilidad de la opción 1 es pequeña y la interacción causada por la otra opción es pequeña, mientras que la accesibilidad de la opción 2 es mayor y el parámetro de interacción es grande, o cuando ambas accesibilidades son del mismo orden y las interacciones también, pero estas últimas son grandes.

En este caso, independientemente de las condiciones iniciales, como puede observarse en las figuras II.6.3a y II.6.3.b, el aumento en el valor de θ indica que se han adquirido los elementos de conocimiento necesarios para elegir la primera opción que, probablemente, ofrece una estrategia más eficiente. Es interesante notar, sin embargo, que si el estado inicial se encuentra en la región A de la figura II.6.3a conforme θ aumenta hay un intervalo de θ, en la figura II.6.3b para el que ambas probabilidades aumentan, hasta que se llega a un punto en el que la probabilidad de elegir la segunda opción empieza a decrecer mientras que la probabilidad de elegir la primera sigue aumentando.

Figura II.6.3a: Regiones en plano fase

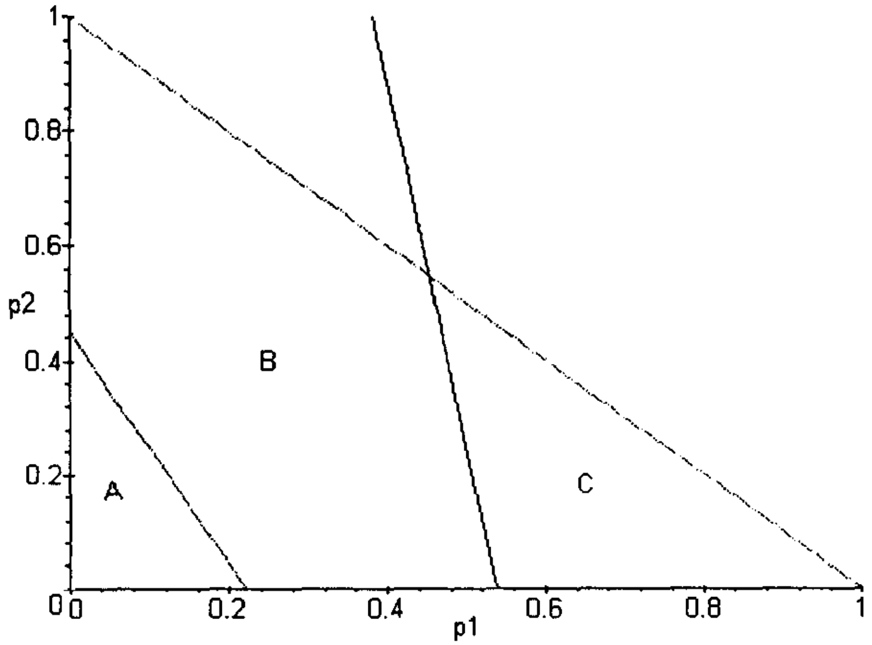
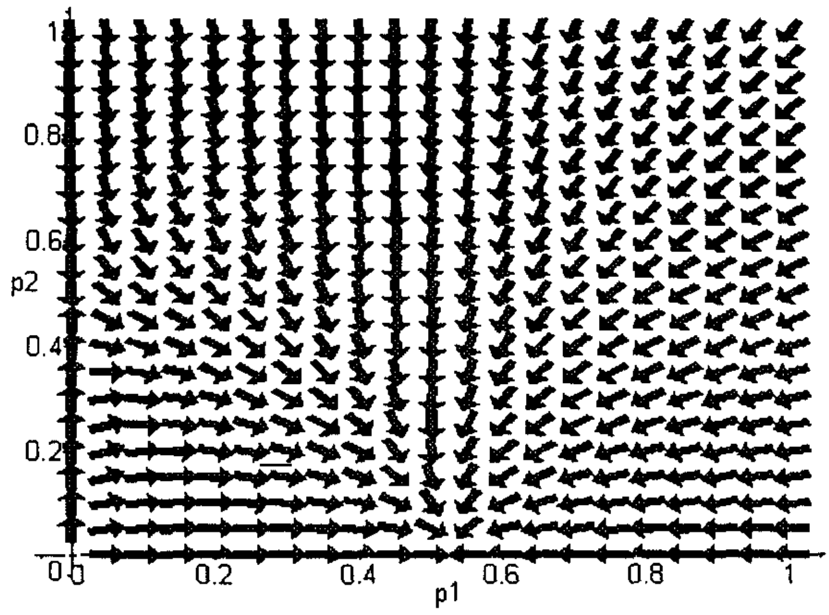


Figura II.6.3b: Campo de direcciones



En el ejemplo que se está utilizando esto podría ocurrir para ecuaciones cuya solución involucra raíces complejas, pero cuya factorización parece ser muy simple si se comete el error común de utilizar signos equivocados.

CASO 4.

Para el caso en el que $a_1 \ /b_1 < k_2$ y $a_2 \ /b_2 > k_1$ se tiene una situación análoga a la del caso anterior en la que únicamente hay que intercambiar las condiciones para las opciones. En este caso, la probabilidad de elegir la segunda opción es la que predomina, pero el comportamiento es análogo al que se discutió en el caso anterior, como se muestra en las figuras II.6.4a y II.6.4b.

Figura II.6.4a: Regiones en plano fase

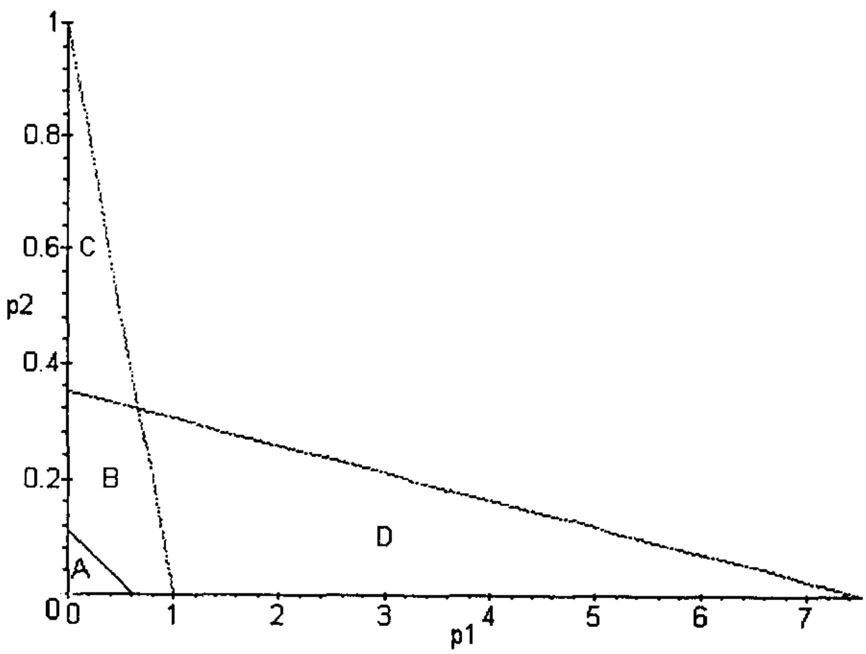
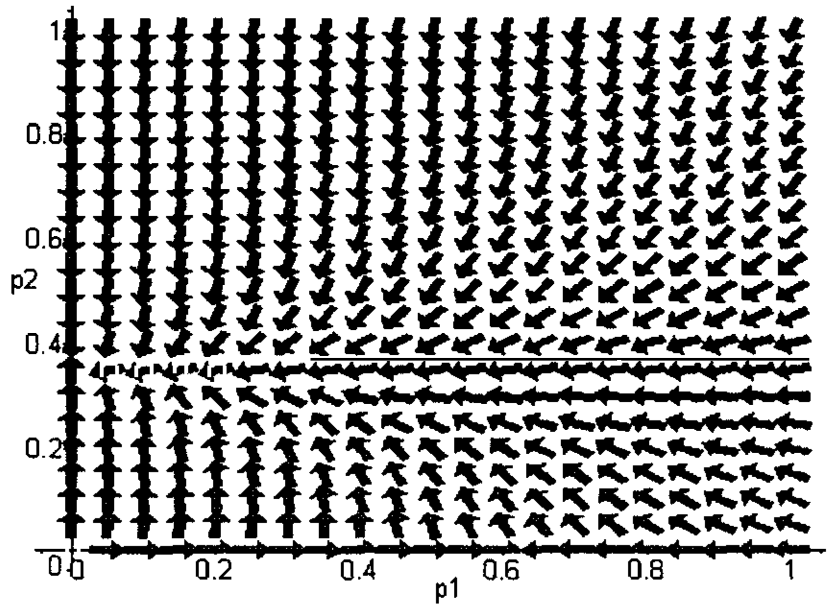


Figura II.6.4b: Campo de direcciones



El ítem que puede ilustrar este caso es aquél que presenta una ecuación cuadrática cuya factorización es muy simple. Para valores pequeños de θ predominará la elección del uso de la fórmula como estrategia de solución, pero para valores grandes de θ la estrategia de factorización será dominante.

CASO 5.

Cuando $a_1 / b_1 = k_2$ y $a_2 / b_2 = k_1$ el espacio fase queda dividido en exactamente dos regiones, como se muestra en la figura II.6.5a. Si las condiciones iniciales se encuentran dentro de la región A, conforme aumenta θ las probabilidades de elegir la opción 1 y de elegir la opción 2 aumentarían hasta alcanzar un punto de equilibrio que se muestra en la figura II.6.5b.

Figura II.6.5a: Regiones en plano fase

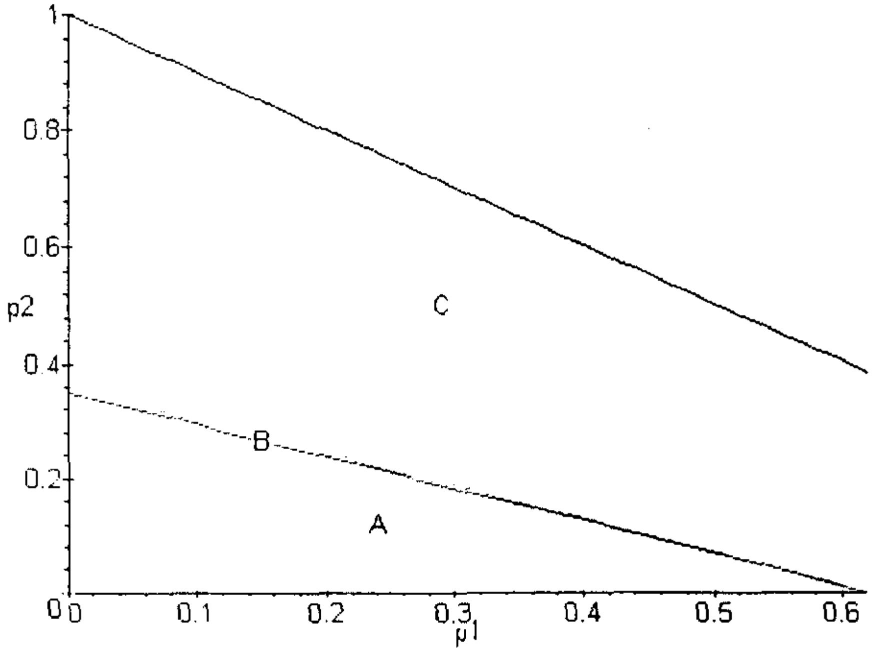
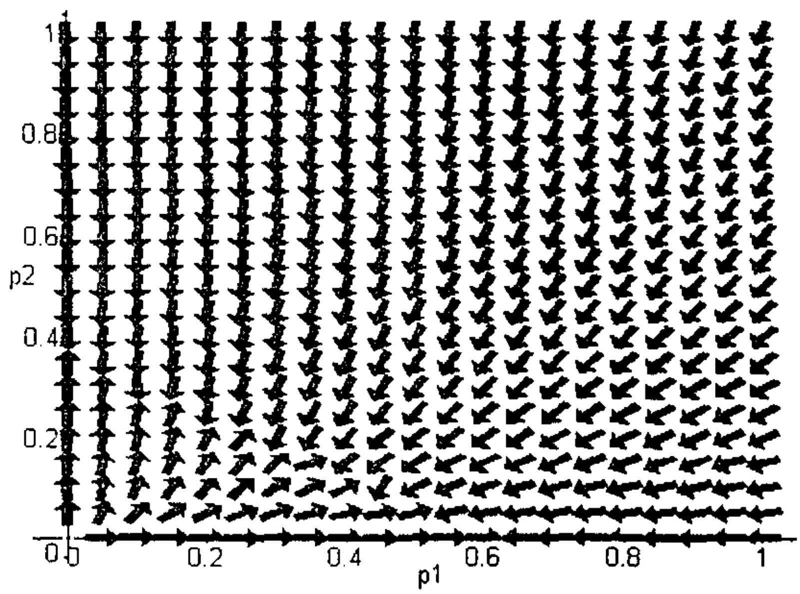


Figura II.6.5b: Campo de direcciones



Si el estado inicial se encuentra dentro de la región B de la figura II.6.5a, en cambio, ambas probabilidades decrecerán hasta alcanzar una situación de equilibrio (Figura II.6.5b) y lo que aumenta es la probabilidad de elegir una tercera opción o de no responder en caso de que no haya una tercera opción disponible.

disponible. A diferencia del caso 1, en este caso existen muchos puntos de equilibrio, cuál de ellos se alcanza cuando $\theta \rightarrow \infty$ depende de la naturaleza del problema planteado.

Esta situación se puede presentar cuando los índices de accesibilidad de las opciones son similares y los parámetros de interacción son pequeños.

En el ejemplo que se está considerando esto sucedería cuando la ecuación cuadrática puede factorizarse pero esta factorización es complicada.

Un parámetro de accesibilidad negativo

Para la segunda familia del modelo, aquélla que se presenta cuando el parámetro a_1 es negativo, la situación que se presenta es totalmente diferente. En ausencia de interacción las curvas correspondientes a las opciones de elección del ítem tienen forma logística, pero una de ellas crece y la otra decrece al aumentar el valor de la variable independiente.

Esta situación, en el ámbito de lo educativo, corresponde a un tipo de ítems en el que al aumentar el valor de la variable independiente, la probabilidad de dar una respuesta, que puede considerarse incorrecta va desapareciendo. Si el ítem en cuestión está diseñado para medir conocimiento, por ejemplo, la opción cuya logística crece representa la respuesta correcta y la opción que decrece representa una respuesta incorrecta, pero muy probable, cuando no se tiene el conocimiento suficiente.

En estas condiciones, el análisis de los distintos casos que se presentan al cambiar los valores de los parámetros, muestra que cuando $\theta \rightarrow \infty$, prevalecerá siempre la opción correcta; es decir, todas las trayectorias tienden al punto de equilibrio (k_1 , 0) y, además, se esperaría que k_1 tomara el valor de 1.

El análisis de los casos 6, 7 y 8 muestra un comportamiento similar, pero es interesante notar que en los casos 6 y 7 se tiene una región factible de solución al problema en la que, para valores de θ dentro de un intervalo eópecífico, ambas probabilidades de elección de opciones disminuyen.

CASO 6

El caso en el que $a_1 / b_1 > k_2$ puede ocurrir cuando la opción correcta es muy accesible independientemente de la competencia de la otra opción, cuando las opciones de respuesta son medianamente accesibles para el sujeto o cuando la opción correcta es poco accesible pero la interacción no juega un papel importante.

En este caso el punto (k_1 ,0) es asintóticamente estable, es decir, la probabilidad de elección de la opción correcta aumenta hasta llegar a su máximo valor, mientras que la probabilidad de elección de la opción incorrecta tiende a cero. Las condiciones iniciales dependen de la naturaleza del problema a resolver.

Las diferentes regiones del espacio fase que se muestran en la figura II.6.6a, determinan la naturaleza de la curva que caracteriza a cada opción. Un análisis de estas distintas regiones muestra que, si las condiciones iniciales se encuentran en la región A, la probabilidad de elección de la opción correcta crece monotónicamente hasta llegar a su valor máximo k_1 mientras que la probabilidad de elección de la opción correcta crece monotónicamente hasta llegar a su valor máximo k_1 mientras que la probabilidad de elección de la opción incorrecta decrece monotónicamente hacia cero (Figura II.6.6b).

Figura II.6.6a: Regiones en plano fase

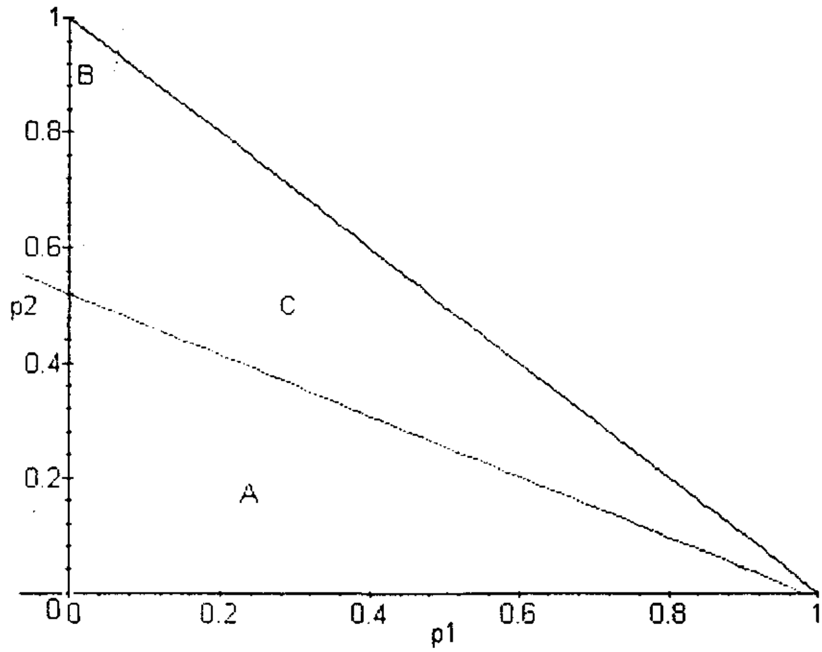
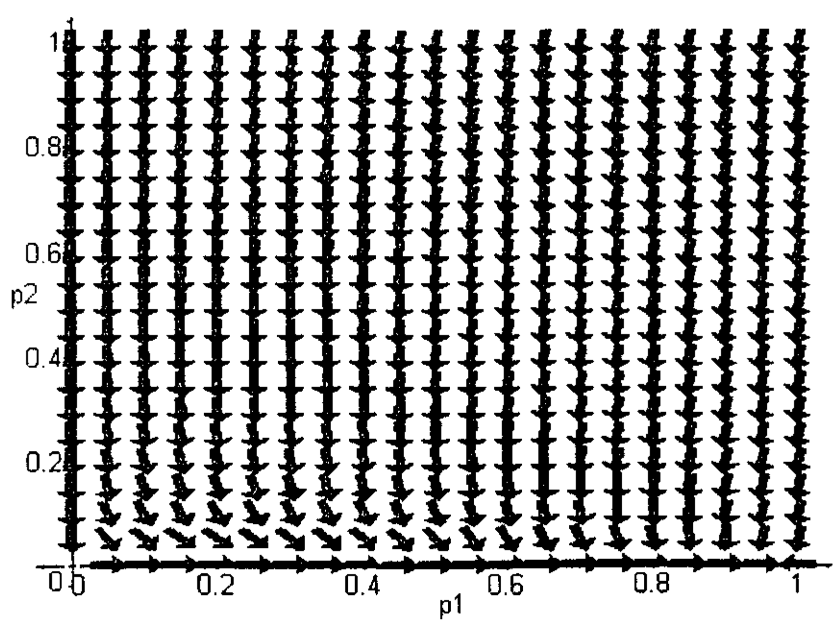


Figura II.6.6b: Campo de direcciones



Cuando las condiciones iniciales se encuentran en la región B del espacio fase que se muestra en la figura II.6.6a, se tiene que para un intervalo de valores de θ, tanto la probabilidad de elección de la opción correcta cuanto la probabilidad de elección de la opción incorrecta decrecen; en ese intervalo de valores de θ, crece la probabilidad de elección de la tercera opción o, en su caso, de no responder. Esto puede ocurrir, probablemente porque la incertidumbre producida por las alternativas produce en el alumno que responde un estado de estupor; se puede identificar, además un punto en la que la duda del alumno es máxima.

Posteriormente la probabilidad de elección de la opción correcta vuelve a crecer hasta alcanzar su valor máximo k_1 , mientras que la probabilidad de elección de la opción incorrecta sigue decreciendo (Figura II.6.6b).

El decrecimiento de la probabilidad de elección de la opción correcta cuando aumenta θ puede indicar que los sujetos para los que el valor de θ se encuentra en ese intervalo, dudan o están confundidos y optan, ya sea por la elección de otra opción o por no responder. Desde el punto de vista cognitivo este intervalo de valores de θ puede asociarse a una etapa de transición entre niveles diferentes de conocimiento en la adquisición de nuevos conocimientos y a una etapa de desequilibrio cognitivo.

Este caso puede ejemplificarse mediante un ítem en el que se pide la solución a un problema algebraico simple, en el que la respuesta correcta puede identificarse rápidamente por prueba y error aún cuando los distractores estén bien diseñados. Otro ejemplo sería un ítem en el que todas las opciones que se presentan son consideradas difíciles por el sujeto que responde.

CASO 7

Los ítems que corresponden a este caso son aquellos para los cuales $a_1/b_1 < k_2$ que corresponden a ítems en los que la opción correcta es accesible y por ello la interacción es pequeña, cuando la opción correcta es medianamente accesible y cuando la interacción entre las opciones es mediana y en el caso en el que la opción correcta es poco accesible y la interacción grande.

Las curvas correspondientes a la probabilidad de elección de las diversas opciones en el espacio fase se comportan, como se muestra en las figuras II.6.7a y II.7.b, de una manera similar a las curvas que aparecen en el caso 6, pero en el caso de aquellos ítems que cumplen con las características mencionadas para este caso, la situación en la que ambas probabilidades de elección para las dos opciones disminuyen para un cierto intervalo de valores de θ puede presentarse con mayor frecuencia, pues en este caso la región factible del espacio fase es mayor que en el caso 6. En esas condiciones, hay una mayor posibilidad de que exista un intervalo de valores de θ en el que la probabilidad de elección de las alternativas 1 y 2 decrece y, en cambio, crece la probabilidad de elección de la tercera opción o, en su caso, de no responder. Esto puede ocurrir, como ya se mencionó debido a la incertidumbre producida por las alternativas que induce en el alumno que responde un estado de estupor. En este caso también se puede identificar un punto en el que la duda del alumno es máxima.

Figura II.6.7a: Regiones en plano fase

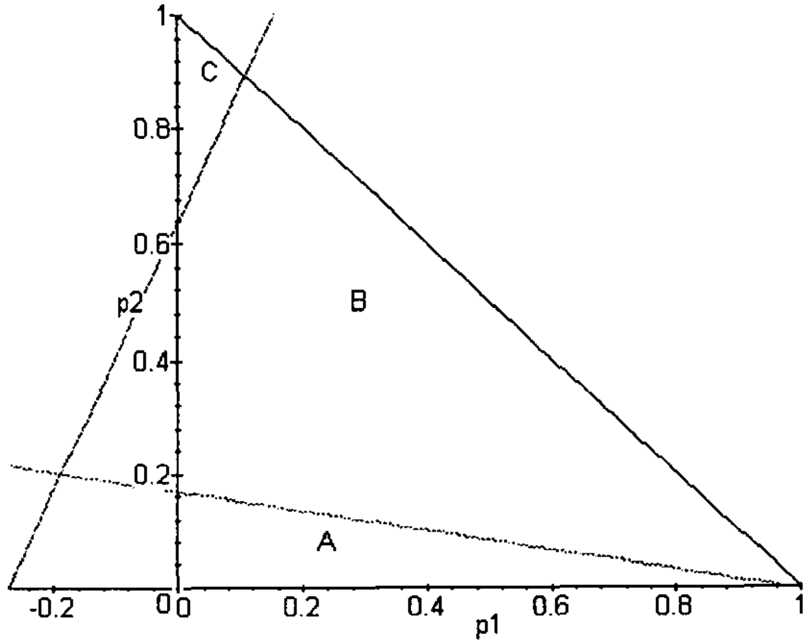
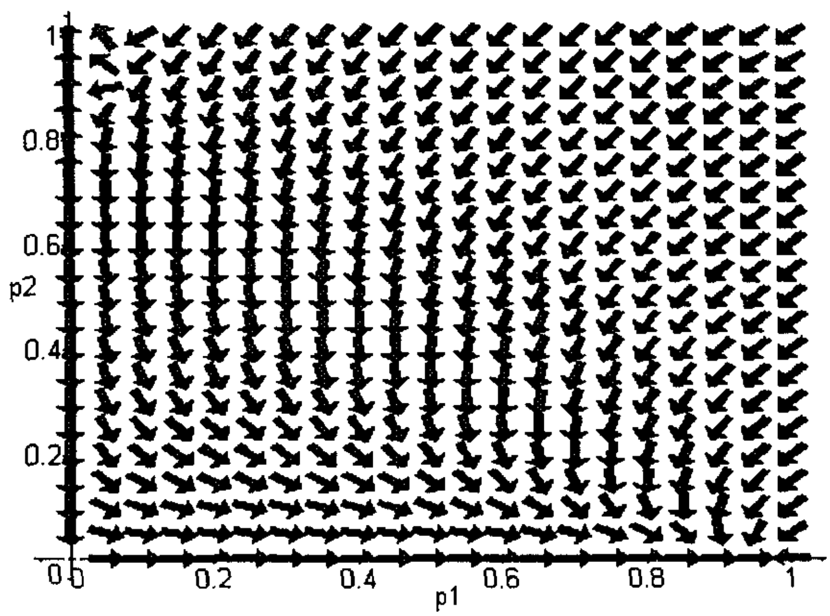


Figura II.6.8b: Campo de direcciones



Este caso puede ilustrarse por ejemplo, con un ítem en el que se pide encontrar la ordenada al origen de una línea recta y en el distractor se incluye como respuesta la abscisa por la que la recta cruza al eje horizontal. En este caso la opción correcta es medianamente accesible, pero el distractor juega un papel importante.

CASO 8

Este caso se presenta cuando $a_1 / b_1 = k_2$. El comportamiento de las curvas correspondientes a este caso es similar al de los casos 6 y 7 que se han discutido con anterioridad y se muestra en las figuras II.6.8a y II.6.8b.

Figura II.6.8a: Regiones en plano fase

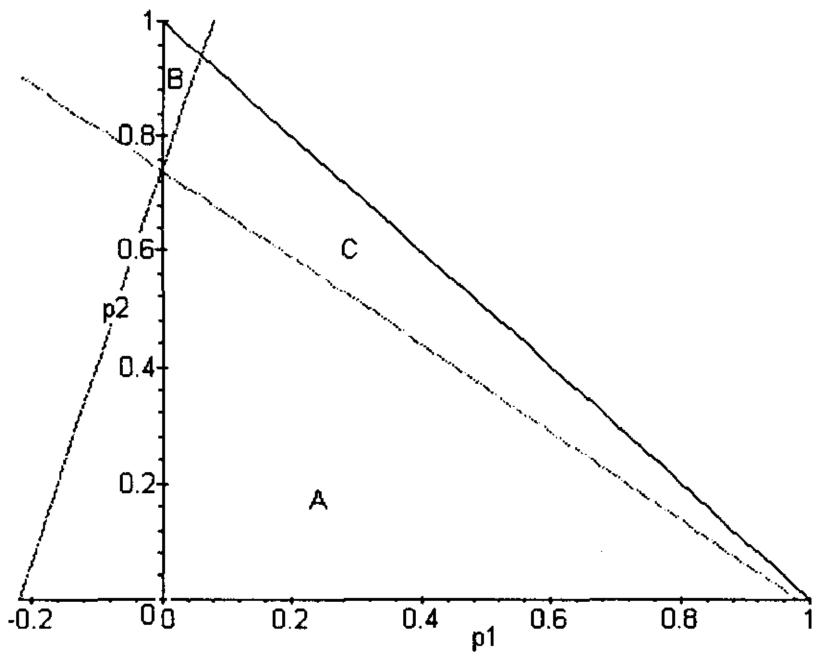
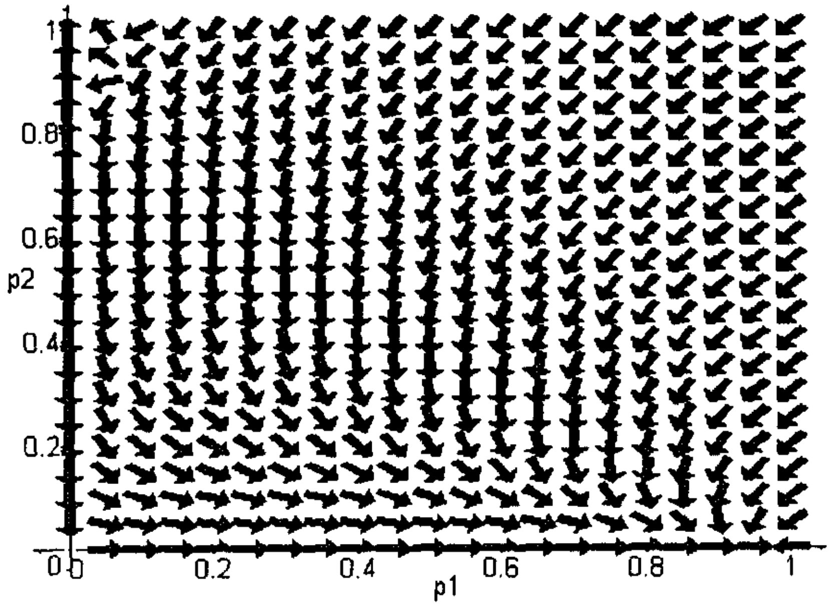


Figura II.6.8b: Campo de direcciones



Los ítems que corresponden a este caso son aquellos para los cuales la atractividad de la primera opción es exactamente igual al máximo de probabilidad de elección de la segunda opción.

Un parámetro de interacción negativo

Por último, y como ya se ha mencionado, la familia que se obtiene cuando alguno de los parámetros de interacción es negativo tiene un valor importante en el diseño de los ítems, más que en cuanto al diagnóstico o la clasificación en términos educativos.

Los resultados que se obtienen indican la prevalencia de la opción de respuesta que se esta reforzando a través del diseño, como era de esperarse; el equilibrio en un punto fuera de la región de permitida o las curvas solución salen de la región permitida sin alcanzar un equilibrio. En términos del modelo, esta familia se presenta únicamente en aquellos casos en los que los distractores no se han formulado correctamente y por ello es útil sobre todo en la etapa de validación de los ítems.

Posibilidades de Diagnóstico del modelo.

El análisis cualitativo de las soluciones al modelo que se presenta en este capítulo muestra comportamientos interesantes de las trayectorias de solución que, al interpretarse desde una perspectiva educativa y con ayuda de un marco conceptual cognitivo, permiten detectar las zonas del espacio fase en el que hay cambios de comportamiento de la trayectoria y localizar el intervalo de valores de θ en los que esto sucede. Una vez que se cuenta con esta información es posible diseñar ítems con las características deseables acordes a lo que se quiere medir.

Cuando el comportamiento de las curvas que describen la probabilidad de elección de una opción de respuesta a un ítem se analiza a la luz de una teoría cognitiva, es posible localizar los intervalos de la variable independiente θ para los cuales se las curvas tienen formas interesantes o diferentes. De esta manera se obtiene como resultado una subdivisión de los valores de la variable independiente en regiones para las cuales el comportamiento de los sujetos es diferente y esta información puede usarse con fines de diagnóstico o de clasificación.

Sección II.7 Estimación de los Parámetros

Una vez desarrollado el modelo se presenta el problema de la estimación de los parámetros que involucra. La cuestión a tratar es la forma de encontrar los valores de las constantes y los valores de la variable independiente que aparecen en el modelo a partir de las mediciones hechas con sujetos. Surgen en este punto algunas dificultades. Por una parte es necesario saber si los datos contienen la información que se está buscando; es decir si los parámetros están determinados en forma única por la función que describe al modelo. Por otra parte los términos no lineales incluidos en el modelo y su naturaleza dinámica hacen que los procedimientos de estimación se compliquen. Finalmente, el valor de la variable independiente θ que describe a cada sujeto no se conoce y debe estimarse, también, a partir de los datos obtenidos de la aplicación de los ítems a una muestra representativa de sujetos.

El modelo desarrollado en este trabajo es un ejemplo de un problema inverso puesto que involucra la identificación de k parámetros a partir de observaciones acerca de la forma en la que el sistema evoluciona. Como es sabido, la mayor parte de los problemas inversos no satisfacen los postulados de Hadamard acerca de los problemas bien planteados: i) Pueden no tener una solución en un sentido estricto; ii) las soluciones pueden no ser únicas o depender de forma no continua de los datos. Este tipo de problemas conocidos como "problemas mal planteados", involucran severas dificultades numéricas y requieren del uso de métodos de regularización para su solución. Entre los efectos que los problemas mal planteados pueden tener están las amplificaciones de errores de alta frecuencia, la restauración de la estabilidad por el uso de información a priori, la propagación del error y la pérdida de información. Todos estos problemas se tienen que tomar en cuenta cuando se resuelve un problema inverso cuya solución conduce a un problema mal planteado. Sin embargo, en una situación concreta es siempre posible imponer algunas condiciones que hacen posible la regularización del problema y que permiten evitar las dificultades antes mencionadas.

Si bien no existe una teoría razonablemente unificada que permita el tratamiento de los problemas inversos no lineales mal planteados, existen métodos y algoritmos adecuados a la situación particular con la que el investigador se enfrenta.

En el problema específico que se plantea en este trabajo, el modelo que nos concierne puede dividirse en dos partes: una parte lineal y una parte no lineal. En casos como este se puede intentar la estimación de los parámetros en dos fases: En la primera se toma únicamnte la parte lineal de problema, que se corresponde, como se ha demostrado anteriormente, con un modelo de la teoría de respuesta a los ítems de dos parámetros, y en la segunda se incorpora la parte no lineal y se recurre a los métodos que se han desarrollado para trabajar con problemas inversos.

Estas dos fases pueden iterarse hasta encontrar la solución óptima en términos de los datos obtenidos de la aplicación de un test a una muestra convenientemente grande de sujetos.

Estimación de la parte lineal del modelo

En la primera fase del procedimiento el problema básico es la estimación de los parámetros de accesibilidad de los ítems a partir de los parámetros de dificultad y de discriminación obtenidos del modelo logístico de dos parámetros y la estimación de la habilidad de los sujetos de la muestra bajo este mismo modelo.

En los modelos de la teoría de respuesta a los ítems, la estimación de los parámetros se hace siguiendo el criterio de que los parámetros seleccionados sean tales que maximicen la probabilidad de que ocurran los datos que se han obtenido en la aplicación a la muestra. El método que se emplea más a menudo es el "método de máxima verosimilitud". La estimación se va haciendo, en la mayor parte de los casos, por aproximaciones sucesivas mediante iteraciones hasta obtener los valores para los parámetros que se consideren satisfactorios en términos del criterio de minimización de error elegido.

El procedimiento seguido tradicionalmente para identificar los parámetros y la habilidad en esta primera fase involucra nuevamente la necesidad de partir el procedimiento en dos partes que se iteran hasta obtener resultados satisfactorios. En la primera parte, se suponen conocidos los valores de los parámetros del ítem y se estima la habilidad. En la segunda se mantienen constantes las habilidades de los sujetos y se estiman los parámetros de los ítems.

Puesto que se conoce la forma logística de la curva característica del ítem y ésta proporciona la probabilidad de que los sujetos con determinado valor en θ acierten el ítem. La respuesta a un ítem para un determinado nivel de θ es una prueba de Bernoulli, por lo que, si u_i es la variable que indica la respuesta al ítem y esta puede tomar los valores 1 ó 0 dependiendo de si el sujeto acierta o no en su respuesta al ítem, entonces se tiene que:

$$P(u_i /\theta) = P(u_i =1/\theta)^{u_i} P(u_i =0/\theta)^{(1-u_i)} = P_i^{u_i}(1 - P_i)^{(1-u_i)} = P_i^{u_i} Q_i^{(1-u_i)},$$

en la que

$$Q_i = 1 - P_i .$$

Si un examinado con habilidad θ responde a *n* ítems, la probabilidad conjunta de las respuestas u_i, para *i* desde 1 hasta *n*, y se cumple el supuesto de independencia local, las respuestas a los *n* ítems son independientes de manera que:

$$P(u_i / i=1,2,\ldots,n/\theta) \quad \Pi_i \ (P_i \ (\theta))^{U_i} (Q_i \ (\theta))^{(1-U_i)}$$

que es la probabilidad conjunta de respuestas a los *n* ítems.

Cuando las respuestas se observan, la expresión anterior deja de ser una afirmación acerca de la probabilidad y se convierte en una función conocida como la función de verosimilitud denotada por

$$L(u_1, \ldots, u_n / \theta) = \Pi_i \ (P_i \ (\theta))^{U_i} (Q_i \ (\theta))^{(1-U_i)}$$

En esta función los valores de *P* y de *Q* están dados por el modelo logístico de dos parámetros.

Utilizando la función de verosimilitud como una función de criterio. El valor de θ que maximiza esta función puede tomarse como un estimador del valor de θ, el que mejor ajusta dados los datos de la muestra.

La estimación se lleva a cabo en dos pasos: primero se suponen valores para los parámetros del modelo y se estiman las puntuaciones de los sujetos mediante la maximización de la función *L* para cada valor de θ, en el cual se hacen estimaciones sucesivas de θ por el método de Newton - Raphson, hasta que se logra la convergencia deseada, es decir que el valor obtenido en una iteración no difiera significativamente del obtenido en la iteración anterior. (Andersen, 1973; Lord, 1975 y 1980; Haberman, 1979; Swaminathan y Gifford, 1983).

Dado que este método de estimación de parámetros es utilizado frecuentemente, existen varios paquetes computacionales que permiten llevar a cabo la estimación.

En el caso del modelo que nos concierne, la estimación de los parámetros se debe hacer para cada una de las opciones del ítem. Los parámetros de accesibilidad, que se obtienen a partir de los parámetros de dificultad y de discrtiminación de un modelo logístico de dos parámetros, pueden tomar, a diferencia de éste, valores positivos o negativos. Por otra parte, el valor máximo de la probabilidad de elección de cada opción no es necesariamente uno. Los programas convencionales permiten únicamente la estimación de parámetros positivos y toman como valor máximo de la probabilidad de elección uno, por lo que, para que dichos programas sean aplicables a este modelo en particular, sería necesario hacerles algunos cambios.

Una vez estimados los parámetros de "discriminación" y de "dificultad" de las opciones se puede utilizar la definición del parámetro de accesibilidad para calcularlo. En el caso del parámetro de máxima probabilidad de elección de las opciones, k, es posible introducir algún criterio con el cual se fijen sus valores. Este criterio puede determinarse a partir de datos provenientes de la aplicación de los ítems a poblaciones específicas de sujetos expertos y novatos.

Estimación de la parte no lineal

La siguiente etapa en la estimación de los parámetros del modelo consiste en la introducción del modelo dinámico que contiene los parámetros de interacción, es decir, en la introducción de la parte no lineal del problema.

El ajuste de los parámetros de esta parte se hace utilizando el modelo dinámico de elección de respuesta a las opciones del ítem en el que los parámetros de

en estas situaciones deben ser capaces de resolver la situación de una manera estable usando una cantidad razonable de recursos computacionales. Una familia de métodos de optimización que han probado ser exitosos para la solución de problemas del tipo del que interesa aquí son los métodos quasi-Newton (Burden y Faires, 1985) y los llamados métodos de Newton truncado y técnicas de gradiente conjugado para lograr una convergencia más rápida hacia la solución del problema y para ahorrar en términos de costo de cómputo (Engl, 1994; Zhu, Byrd, Lu y Nocedal, 1994; Dembo y Steinway, 1983; Gómez y Morales, 1997).

Estos métodos retienen la mayor parte de las propiedades de convergencia rápida del método original de Newton y contienen recursos para afrontar los problemas de estabilización cuando se manejan en conjunto con métodos de gradiente conjugado.

El procedimiento que se sigue cuando se utilizan estos métodos es el siguiente: Dado un punto en el espacio de parámetros, se calcula la solución al modelo dinámico y con ella la funcion objetivo *F* regularizada. Con la técnica de gradiente conjugado se busca cuál es la dirección en la que se obtiene un mayor descenso de la función *F*. Se calculan los valores de los parámetros en un punto cercano al anterior, moviéndose en la dirección seleccionada y nuevamente se resuelve el modelo y se calcula *F*. Este procedimiento se repite hasta llegar al valor de convergencia deseado.

La estimación de los parámetros del modelo que se desarrolló en este trabajo puede llevarse a cabo utilizando los procedimientos aquí mencionados e iterando tantas veces como sea necesario entre la fase de estimación de la parte lineal del modelo y la estimación de la parte no lineal del modelo, hasta lograr la convergencia deseada en la función *F*.

En el siguiente capítulo se presentará la aplicación de un procedimiento de esta naturaleza para la estimación de los parámetros en una situación particular. Se utiliza el modelo para analizar los resultados obtenidos de la aplicación de un cuestionario que pretende medir los conocimientos de los sujetos respecto al concepto de variable en el álgebra elemental y se demuestra así la viabilidad del modelo.

Con el fin de poner a prueba el modelo en esa situación particular, se hará en primer lugar la descripción de una teoría cognitiva que sustenta el diseño del cuestionario y el análisis del mismo en términos de la teoría. Finalmente se analizarán los datos obtenidos de la aplicación del modelo en términos de los parámetros obtenidos para cada uno de los ítems del test. ◆

Referencias

ANDERSEN, E. B. (1973). *Conditional inference and multiple choice* questionnaires. British Journal of Mathematical and Statistical Psychology, 26, 31-44.

BANKS, H. y KUNISH, K. (1989). Parameter estimation techniques for distributed systems. Birkhauser, Boston.

BURDEN, R. L. Y FAIRES, J. D. (1985). Numerical Analysis, 3rd Edition. Boston. Prindle, Weber y Schmidt.

DEMBO, R.S. y STEINWAY, T. (1983) Truncated Newton methods for large scale unconstrained optimization. Math Progf, 26, 190-212.

ENGL, H. W., (1994). Inverse problems. Notas para curso. Comunicación personal.

FISHER, G. H. (1973). The linear logistic test model as an instrument in educational research. Acta Psychologica, 37, 359 –374.

FISHER, G.H. (1995). Some neglected problems in IRT. Psychometrika, 60, 459-485.

GLASER, R., LESGOLD, A. Y LAJOIE, S. (1987). Towards a cognitive theory for the measurement of achievement. En R. RONNING, J. GLOVER, J. C. CONOLEY y J. WITT (Eds.) The influence of cognitive psychology on testing and measurement, Vol. 3, 41-85. Hillsdale, NJ, Erlbaum.

GOMEZ, S y MORALES, J.L. (1997). A truncated method for solving parameter identification problems. Documento de trabajo. Depto. de Matemáticas, ITAM, México.

HABERMAN, S.J. (1979). Analysis of qualitative data: New developments, Vol. 2. New York: Academic Press.

LORD, F.M. (1975). Formula scoring and number right scoring. Journal of Educational Measurement, 12, 7-12.

LORD, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale N.J. Lawrence Erlbaum.

MISLEVY, R. J y WILSON, M. (1996). Marginal maximun likelihood estimation for a psychometric model of discontinous development. Psychometrika 61, 41-71.

PAULSON, J. (1985). Latent cass representation of systematic patternsin test responses (ONR Technical Report) Portland, OR: Portland State University.

SWAMINATHAN, H. y GIFFORD, J.A. (1983). Estimation of parameters in the three latent trait model. En D. Weiss (Ed.), New horizons in testing. New York: Academic Press.

TATSUOKA, K. K. (1983). Rule-Space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20, 345-354.

ZHU, C., BYRD, H., LU, P. y NOCEDAL, J. (1994) LBFGS-B: Fortram subroutines for large-scale bound constrained optimization, Report NAMII, EECS, Department, Northwestern University, USA. 16, 345-364.

CAPÍTULO III

APLICACIÓN DEL MODELO DE MEDIDA A SITUACIONES EDUCATIVAS

Para ser consistentes con una visión del aprendizaje como un proceso activo de construcción de conocimientos, los tests requieren centrarse en conceptos clave en un área particular del mismo y tomar en consideración la variedad de tipos y niveles de comprensión que los estudiantes pueden tener de esos conceptos. En la evaluación de la información que se obtiene de la aplicación de estos tests no basta con el conocimiento del número de respuestas correctas e incorrectas de cada alumno, ya que a partir de esa información no es fácil inferir los niveles de conocimiento de los alumnos, ni obtener datos que permitan establecer un diagnóstico.

En los tests tradicionales referidos a los conocimientos de matemáticas, está implícita una visión del estudiante como un sujeto pasivo que absorbe el conocimiento que se le comunica. Muchos de estos tests requieren poco más que la habilidad de recordar fórmulas y de hacer las sustituciones adecuadas para llegar a una respuesta correcta. En este tipo de tests el diagnóstico se vuelve simplemente un asunto de identificación de los espacios en los que los estudiantes han sido incapaces de obtener la respuesta correcta, infiriéndose que es en ellos donde hace falta conocimiento y donde se requiere una enseñanza remedial de la cual no puede especificarse mucho más.

Abramovitz y Ramsay (1992) mostraron que en muchas ocasiones la respuesta a los ítems es más compleja de los que suele esperarse. Mediante técnicas de análisis categórico y modelos de spline, encontraron que la mayor parte de los examinados responden equivocadamente ítems fáciles y que, en muchos casos, el tratamiento de los ítems como binarios hace que se pierda información interesante y útil contenida en la elección de las opciones alternativas.

En las últimas décadas, por otra parte, se ha avanzado notablemente en la comprensión de la forma en la que los estudiantes aprenden. En particular hay mayor consciencia de la naturaleza activa y constructiva de la adquisición del conocimiento y de la importancia de las concepciones previas de los sujetos, así como de las representaciones personales del conocimiento en las distintas áreas temáticas. En los modelos constructivistas del aprendizaje, la diferencia entre un sujeto menos avanzado y otro más avanzado no radica en la diferencia de conocimiento factual, ni entre lo que cada uno puede recordar en un momento dado (aunque éste sigue siendo un aspecto importante de un desempeño competente), sino en la diferencia en el tipo de concepciones, relaciones y comprensión que los estudiantes muestran cuando enfrentan un problema y en las estrategias que usan para resolverlo.

El desempeño de los estudiantes novatos puede entenderse en términos de los modelos inapropiados o ineficientes que han construido. Sin embargo, las *concepciones inadecuadas de los estudiantes pueden pasar desapercibidas* si las respuestas correctas a un test dependen únicamente del conocimiento superficial de leyes, fórmulas y algoritmos de manipulación.

Los puntos de vista constructivistas consideran que, sin importar qué tan avanzado está un estudiante, tiene cierto conocimiento o comprensión de los temas en cuestión; información de la cual echa mano ante un problema específico. Más que ser incorrectas, las respuestas pueden mostrar una *comprensión parcial de la materia que se usa consistentemente en la solución de* problemas.

Una implicación de estas teorías en la evaluación y diagnóstico del aprendizaje de los alumnos es que los modelos cognitivos de los estudiantes deben tomarse en consideración en el diseño de las pruebas y en al análisis de los resultados obtenidos de la aplicación de las mismas. En un gran número de áreas el nivel de desempeño se define y se mide mejor en términos de los niveles de comprensión de los conceptos clave que poseen los alumnos, más que en términos de los hechos y los procedimientos que ellos pueden recordar.

En el modelo que nos interesa aplicar en el presente trabajo, se intenta interpretar los resultados de la aplicación de los tests en términos de un modelo matemático que incluye parámetros que pueden asociarse a elementos específicos de una teoría de aprendizaje. A partir de esos datos es posible hacer inferencias acerca del nivele de aprendizaje en el que los estudiantes se encuentran en relación con conceptos clave en un área particular del conocimiento.

El modelo planteado en el capítulo anterior extiende los modelos de medida logísticos, que son los que se emplean más frecuentemente y abre la posibilidad de incorporar aspectos cognitivos de las distintas áreas del conocimiento. En este capítulo, se presenta una teoría cognitiva que, utilizada en conjunto con el modelo propuesto en el capítulo anterior, puede ayudar a lograr un mejor diagnóstico del nivel de comprensión de los estudiantes de un concepto a partir de sus respuestas a ítems específicos. La aplicación que se hace en el caso del presente trabajo se enmarca dentro del campo del aprendizaje de las matemáticas. En particular se hace la aplicación utilizando ítems relacionados con el aprendizaje del concepto de variable en el álgebra elemental.

La conjunción del modelo de medida con la teoría cognitiva debe comenzar haciendo una enumeración explícita de los procesos mediante los cuales los sujetos eligen diferentes alternativas de respuesta al ítem y de la forma en la que estos procesos se relacionan con los parámetros de los ítems. Por esta razón el presente capítulo se inicia con una revisión de un modelo de diagnóstico basado en la teoría de Piaget, que sirve como base de un marco teórico más explícito referido al aprendizaje de los conceptos matemáticos.

Se presentará a continuación una discusión acerca de los problemas principales que los estudiantes manifiestan cuando se enfrentan al concepto de variable y algunos resultados relevantes de la literatura en didáctica respecto a dicho concepto, para dar paso a la presentación de un conjunto de preguntas que conforman un cuestionario de diagnóstico.

Posteriormente, se ejemplificará la aplicación del modelo mediante el análisis de dichos ítems, identificando los resultados con los diversos casos del modelo presentados en la sección II.5 y se justificará la relación entre los procesos cognitivos de los sujetos que responden al ítem y los parámetros del modelo en términos del marco teórico. De esta manera es posible dar sentido a las respuestas de los estudiantes en términos de los procesos cognitivos involucrados y hacer del test una herramienta más efectiva de diagnóstico.

Sección III.1 Diagnóstico: una perspectiva teórica

En el marco general de la teoría educativa, el diagnóstico pretende conocer a la persona o grupo, aclarar sus distintos atributos y formular indicadores de su comportamiento ante las distintas situaciones que enfrenta. Este conocimiento se toma como punto de partida en el diseño de estrategias para ayudar a la persona o grupo a avanzar.

Suele definirse el diagnóstico tomando en consideración el estado del conocimiento de un sujeto en un momento dado de un proceso de instrucción para tomar decisiones. Este tipo de definición es estática y no toma en consideración que el aprendizaje está en constante evolución. Una definición más dinámica del diagnóstico incluye la naturaleza cambiante de los procesos de adquisición de conocimiento y los considera como procesos en los que la valoración pretende tomar en cuenta la interacción entre las variable más importantes. Una posible definición de diagnóstico es la siguiente: "El diagnóstico es un proceso técnico de identificación y valoración de las variables más relevantes y específicas que inciden en un alumno, grupo o situación escolar, para comprender y explicar su interacción, garantizando una adecuada intervención o toma de decisiones". (García Nieto, 1987, 1994). Esta definición parte de la concepción del diagnóstico como un proceso de naturaleza dinámica y científica, contra la visión del diagnóstico como una intervención incidental en un problema educativo.

El conocimiento que provee el diagnóstico se alcanza por una parte mediante la acumulación de datos, y por otra, por los medios que permiten el acopio de tales datos (García Lozano, 1966 en Martínez, 1990). Esto implica que para poderlo llevar a cabo es importante delimitar las características del objeto de estudio para determinar con claridad las distintas posibilidades de acciones diagnósticas (García Lozano, 1966 en Martínez, A. 1990). Así, el proceso diagnóstico puede considerarse como el conjunto sistemático de indagaciones para conocer un hecho educativo con el fin de proponer sugerencias.

Aceptada la definición de García Nieto del diagnóstico como un proceso técnico de identificación y valoración de las variables más relevantes y específicas que inciden en una situación escolar, para comprender y explicar su interacción garantizando una adecuada intervención o toma de decisiones, nos interesa situar esta definición en el contexto del problema del aprendizaje de las matemáticas.

Un problema conocido, que es preocupación a nivel mundial, es la dificultad de los estudiantes para lograr un aprendizaje sólido y significativo de las matemáticas. Debido a la naturaleza propia de esta disciplina, conforme se avanza, el conocimiento de los nuevos conceptos requiere de un manejo adecuado de las técnicas y conceptos anteriores a ellos; esto hace que pocas personas lleguen siquiera a intentar el acercamiento a las matemáticas de nivel superior.

El diseño de instrumentos y de marcos adecuados de análisis para el diagnóstico del aprendizaje de los conceptos matemáticos es indispensable para recabar la mayor cantidad de información posible acerca del tipo de imagen que los alumnos tienen de los conceptos de una disciplina dada y para orientar las decisiones de los pasos a seguir para que continúen avanzando en su nivel de comprensión o para corregir sus posibles errores.

Un diagnóstico de la naturaleza que se propone en la definición anterior requiere de la aclaración de las variables más relevantes de dicho aprendizaje. Esta delimitación de variables depende del marco teórico elegido como referencia para el desarrollo del modelo que sustenta el análisis diagnóstico. Distintas teorías acerca del aprendizaje conducen a modelos de diagnóstico con diferentes variables.

En particular la teoría de Piaget sobre la construcción del conocimiento proporciona, como se verá más adelante, elementos que permiten reconocer las variables que interaccionan de manera más prominente en el ámbito de las matemáticas. La epistemología piagetiana centra la atención en los procesos y en los mecanismos generales de la producción del conocimiento por parte del individuo. Estos son aplicables al aprendizaje de las matemáticas y por ello pensamos que esta epistemología aunada a una aplicación pertinente y concreta al aprendizaje de las matemáticas, brinda los elementos requeridos para desarrollar un marco teórico útil para la evaluación y el diagnóstico del aprendizaje de las matemáticas. Un marco teórico de esta naturaleza debe permitir el diseño de formas de análisis y de intervención orientadas a hacer este aprendizaje más eficaz.

La epistemología de Piaget, ha tenido un fuerte impacto en la psicología y en la investigación sobre enseñanza de las ciencias y de las matemáticas. Los acercamientos constructivistas derivados de esta epistemología se han convertido en un paradigma dominante que sustenta los marcos teóricos de numerosas investigaciones en enseñanza de las matemáticas a todos los niveles, ya que proporcionan elementos claros, manejables y sustentados empíricamente, que permiten explicar muchos de los fenómenos y de los problemas que se presentan en las situaciones de aprendizaje.

La obra de Piaget es muy abundante y de difícil lectura, pero la profundización en ella proporciona un gran número de elementos que permiten al investigador educativo explorar desde diversas perspectivas las situaciones de enseñanza y de aprendizaje: desde la comprensión y explicación de los obstáculos epistemológicos y cognitivos de los alumnos, hasta la contextualización y la solución de problemas complejos. Los modelos constructivistas intentan comprender la naturaleza del aprendizaje de los conceptos y de su tematización en estructuras conceptuales a través de la idea de que el conocimiento no se adquiere en forma lineal por la acumulación de información. El conocimiento crece de manera compleja, mediante reestructuraciones que producen etapas bien diferenciadas en las que el papel de las acciones del sujeto sobre los objetos de conocimiento y de la abstracción son preponderantes.

La posibilidad de reconocer estas etapas o estadios en la construcción de conceptos determinados, así como de las condiciones y mecanismos que pueden hacer posible el paso de uno a otro, mediante la reorganización de las estructuras mentales, se ha convertido en un reto para los investigadores en educación. Esta referencia a la existencia de las estadios identificables le ha brindado a las ciencias de la educación un marco teórico que permite el desarrollo de métodos y de instrumentos de estudio.

En el caso particular del diagnóstico, la idea de estadios en la construcción del conocimiento permite el diseño de instrumentos. Estos instrumentos no son fáciles de diseñar pues deben construirse de tal manera que permitan la exploración de la forma como la persona afronta una situación problemática desde diferentes perspectivas, de tal manera que puedan dilucidarse con mayor claridad los componentes esenciales de su acercamiento. Sin embargo, el diseño de este tipo de instrumentos es prometedor en un gran número de campos del conocimiento.

Consideraremos a continuación los elementos primordiales de la teoría de Piaget de la que se tomarán los elementos que se considera pertinentes en el diagnóstico del aprendizaje de las matemáticas.

Elementos básicos de la teoría de Piaget

La evolución del conocimiento

La teoría de Piaget sobre el desarrollo del pensamiento parte de la suposición de que, contrariamente a las hipótesis del empirismo, el sujeto no observa pasivamente la realidad que le rodea haciendo una copia mental de ella en la que se destacan sus características principales, sino más bien la mente interpreta y construye activamente, de manera dialéctica en la interacción del sujeto con el objeto de conocimiento, una representación de la realidad que se inicia con la organización actual de su conocimiento. El sujeto transforma su representación de la realidad de acuerdo a la forma como se organiza el conocimiento para poder aceptarla. Este proceso de transformación tiene dos características: la resistencia al cambio y la necesidad del mismo. Uno lleva a la estabilidad del conocimiento y el otro a su crecimiento. Ambos procesos pueden operar simultáneamente (Labinowitz, 1987).

El desarrollo del conocimiento se lleva a cabo mediante dos procesos adaptativos: asimilación y acomodación. En el proceso de asimilación se incorporan las percepciones de nuevas experiencias dentro de la estructura mental actual del sujeto y puesto que hay resistencia al cambio, las percepciones pueden ser tergiversadas para ajustarse al marco conceptual existente. Por otro lado el sujeto modifica y enriquece las estructuras conceptuales como resultado de nuevas interacciones con los objetos de conocimiento que demandan cambios en la forma de percibir las experiencias y, por ello, la estructura mental del sujeto se ve obligada a reestructurarse. Cuando las estructuras existentes se modifican, esa modificación puede involucrar la reorganización de estructuras existentes o la elaboración de algunas nuevas que permiten la inclusión de la nueva información. A este proceso Piaget le llama acomodación . El proceso de acomodación es aquél que lleva a la modificación de las estructuras mentales del sujeto de una forma determinante.

La compensación entre estos dos procesos conduce progresivamente al sujeto a niveles superiores de entendimiento. Piaget le llama a esta compensación intelectual activa equilibrio y al proceso por el cual se da, equilibración. El logro de una equilibración requiere del paso por un estado de desequilibración, que incluye el conflicto interno entre interpretaciones opuestas. Este estado es necesario para que se de la evolución en la explicación de los fenómenos que enfrenta el sujeto.

Así, el cambio en el conocimiento se va dando con la participación activa del sujeto, participación que puede ser física o social, en la interacción continua entre la mente y la realidad, en forma de pasajes entre estados estables en los que hay resistencia al cambio. El paso de un estado a otro implica una reorganización no lineal de las concepciones del individuo.

La Interacción dialéctica Sujeto - Objeto

Cuando Piaget habla de la construcción del conocimiento toma una posición concreta acerca de la interacción entre el sujeto que conoce y el objeto de conocimiento. La mente del sujeto posee estructuras que influyen en la forma en que se acerca a la realidad y por ende en la forma en que se construye el conocimiento. La interacción sujeto - objeto es para Piaget una interacción dialéctica. El enfrentamiento con el objeto hace que el sujeto se plantee nuevas preguntas acerca de éste que tratará de responder usando sus estructuras mentales. Cuando éstas no son adecuadas para la solución del problema, se desequilibran y a través de la acción o construcción activa del sujeto encuentran un nuevo equilibrio en que las estructuras cognoscitivas se reacomodan.

Mediante la repetición de este proceso de equilibración - desequilibración cada vez a mayores niveles de abstracción, el sujeto va construyendo el conocimiento.

El deseo de aprender, a la luz de esta teoría se despierta a través de la observación de hechos, entendiéndose por hechos cualquier evento que pueda considerarse como un observable interpretado por el observador. Sin embargo, los hechos por sí mismos no conducen a la construcción del conocimiento, se requiere de la participación activa del sujeto en la generación de relaciones entre estos hechos para que se produzca la inquietud por entender la realidad y con ella la evolución del conocimiento.

Las funciones adaptativas

Para Piaget el conocimiento es el resultado de la adaptación del sujeto a su medio (físico, social y cultural); esta adaptación se produce mediante la puesta en marcha de dos funciones, mutuamente complementarias, que prolongan las fuciones adaptativas del individuo. Estas funciones son la asimilación, es decir, la integración de información en las estructuras cognitivas y la acomodación, el ajuste requerido para manejar obstáculos que se presentan en la construcción del conocimiento, de las que ya hemos hablado anteriormente. Mediante la asimilación y la acomodación se producen cambios estructurales en el conocimiento.

Las funciones de asimilación y acomodación son el centro de la interpretación de los fenómenos intelectuales propuestos por Piaget, por ello ahondaremos un poco más en ellas. Psicológicamente hay asimilación cada vez que el individuo incorpora a sus estructuras cognitivas personales el dato de la experiencia. Asimilar un objeto o una situación es actuar sobre él para transformarlo en sus propiedades o en sus relaciones. La actividad de transformación que interviene en el proceso de asimilación es esencial, es la coordinación de acciones la que constituirá el marco o esquema al que se incorporarán ulteriormente objetos o acontecimientos nuevos. El esquema es la estructura de una acción, que cuando se fija se convierte en "repetible" y por lo tanto en "aplicable", por asimilación a diferentes situaciones. La asimilación llega a atribuir una estructura de significación a las acciones.

Piaget distingue tres tipos de asimilación (Piaget, 1977): la asimilación reproductora, que consiste en la repetición simple de una acción que asegura su fijación; la asimilación recognitiva, es decir la discriminación de los objetos que pueden ser asimilados a un esquema particular; y la asimilación generalizadora, que es la más fecunda pues conduce a ampliar el terreno de un esquema dado y por lo mismo a ampliar la clase de objetos que pueden ser asimilados.

La acomodación consiste en un ajuste cada vez más fino de los esquemas de las acciones para adaptarlos mejor a las características particulares de los objetos. Estos cambios no los sufre pasivamente el sujeto. La acomodación es una actividad orientada; no hay acomodación si el individuo se conforma con comprobar la inadecuación de esos esquemas y se desinteresa enseguida del objeto que ha provocado su acción inicial.

Acomodar es afinar y modificar las estructuras mentales cuando los marcos de pensamiento, al enfrentarse a una situación imprevista, se revelan inoperantes. Es el fracaso de la asimilación el que conduce a la acomodación pero ésta sólo se da gracias a una iniciativa del sujeto y supone además un esfuerzo.

La actividad cognitiva se caracteriza por el equilibrio entre la asimilación y la acomodación. Este equilibrio no se concibe como un estado estático de reposo debido a la anulación de fuerzas que actúan en sentido contrario, sino más bien como una compensación dinámica que es el fruto de la actividad del sujeto en respuesta a las perturbaciones externas. (Hatwell, 1992).

Piaget insiste en el hecho de que no hay acomodaciones sino a continuación de una asimilación o de una tentativa de asimilación. Es decir, las conductas cognoscitivas tienen su origen en el mismo sujeto que actúa sobre el medio para transformarlo al mismo tiempo que reacciona a sus incitaciones. No hay posibilidad de lectura de la experiencia sino hasta que el sujeto ha construido los marcos lógicos o esquemas indispensables para la asimilación de ese dato.

Las asimilaciones y acomodaciones integran un estado autorregulatorio de transformaciones continuas entre condiciones de equilibrio, desbalance y nuevo equilibrio. La creación de cadenas de acciones autorreguladoras ayuda a la comprensión de nuevos hechos o a la reinterpretación de hechos conocidos anteriormente (Ginsburg y Opper, 1979).

Los Instrumentos del Conocimiento

El acercamiento epistemológico de Piaget a la evolución del conocimiento expresado en su obra Psicogénesis e Historia de la Ciencia, realizada conjuntamente con García (Piaget y García, 1982) contiene conceptos que ayudan a profundizar en los procesos de evolución del conocimiento. En ella se introducen tres categorías: instrumentos, procesos y mecanismos del conocimiento.

Al entrar en funcionamiento la asimilación y la acomodación se generan instrumentos cognoscitivos que permiten al sujeto la organización de los datos de la realidad. Estos instrumentos son las abstracciones y las generalizaciones. Es a través de abstracciones y generalizaciones que se aplican a los contenidos de los conceptos y a la forma de representarlos como el sujeto da un sentido a sus acciones. Las abstracciones son resúmenes concisos de aquello que se comparte en el significado de un concepto bajo diferentes contextos. Las generalizaciones en cambio extienden el rango de aplicabilidad de las definiciones de los conceptos.

Piaget propone distintas formas de abstracción. La abstracción empírica que se refiere a los objetos exteriores al sujeto, en los cuales se comprueban ciertas propiedades para extraerlas y analizarlas independientemente. La abstracción reflexiva, que tiene un papel preponderante en la construcción de los conceptos de la física y, sobre todo, en la construcción de los conceptos de la matemática, se refiere a las acciones y operaciones del sujeto y a los esquemas que éstas le conducen a construir. Esta abstracción de acuerdo a Piaget y García es reflexiva en dos sentidos inseparables. Por una parte, en el sentido de que hace pasar lo que es abstraído de un nivel inferior a uno superior (por ejemplo de la acción a la representación) y, por otra, una reflexión en el sentido mental que permite una reorganización sobre el nuevo plano de lo que ha sido extraído del plano precedente. En el caso de las matemáticas en las que el sujeto elabora a la vez formas y contenidos, la abstracción reflexiva es la única que opera, por ello volveremos con más detalle a ella.

Mediante las distintas formas de abstracción se llega a la formulación matemática de las relaciones. Al hablar de esta formulación, Piaget no se refiere a escribirlas en forma de ecuaciones o en el lenguaje formal de la matemática. Se refiere a su estructuración lógica.

A las diferentes formas de abstracción les corresponden diferentes formas de generalización. Cuando se refieren únicamente a datos empíricos, las generalizaciones tienen una naturaleza extensional y constituyen pasajes de algunos a todos o bien de leyes particulares a leyes más generales. La abstracción reflexiva permite en cambio generalizaciones completivas y constructivas que constituyen síntesis nuevas en el seno de las cuales las leyes particulares adquieren nuevas significaciones.

De acuerdo a Piaget y a García, estas nociones conducen a una interpretación específica de las matemáticas. En lugar de considerarlas como un sistema de deducciones referidas a entidades dadas desde el principio, proponer que ellas son extraídas de las acciones u operaciones del sujeto significa que toda acción del sujeto está coordinada con otras acciones, que de estas acciones se extraen formas que pueden ser desprendidas de sus contenidos y que estas formas se coordinan a su vez y dan nacimiento, por reflexión, a las operaciones fundamentales que constituyen el punto de partida de las estructuras lógico algebraicas (Piaget y García, 1982).

Las generalizaciones completivas consisten en retomar las partes a partir del todo, enriqueciéndolas en este proceso y tornándolas en instrumentos de nuevo conocimiento. Cuando los sujetos usan estos instrumentos al formular preguntas o plantear hipótesis y al confrontar observaciones e interpretaciones con las predicciones y deducciones esperadas en el marco de su modelo conceptual, se pueden esperar dos tipos de resultados: a) la formulación de correspondencias entre conceptos hechas en términos de comparaciones y b) la construcción de nuevos conceptos mediante la transformación de los conocidos.

La abstracción reflexiva

La abstracción reflexiva es un concepto introducido por Piaget para describir las estructuras lógico matemáticas de un individuo durante la evolución de su sistema cognitivo. Piaget hizo dos observaciones de importancia acerca de esta noción: la primera se refiere a que la abstracción reflexiva no tiene un comienzo absoluto sino que está presente desde las primeras coordinaciones de las estructuras senso-motoras (Beth y Piaget, 1966), la segunda es que continúa a través del desarrollo de las capacidades matemáticas más avanzadas hasta el punto que la historia entera del desarrollo de las matemáticas, desde la antigüedad hasta nuestros días, se puede considerar como un ejemplo del proceso de abstracción reflexiva (Piaget, 1977). Piaget consideraba a este proceso la fuerza conductora de las reconstrucciones involucradas en el paso a través de los distintos niveles operativos.

De hecho, Piaget distingue tres tipos diferentes de abstracción (Piaget y García. 1982). La empírica se deriva del conocimiento de las propiedades de los objetos. Tiene que ver con experiencias que le parecen externas al sujeto. El conocimiento de estas propiedades es, sin embargo, interna y es el resultado de las construcciones hechas internamente por el sujeto. De acuerdo a Piaget este tipo de abstracción permite la identificación de las propiedades comunes a distintos objetos, extensiones y generalizaciones.

La abstracción semiempírca es intermedia entre la empírica y la reflexiva. Tiene que ver con los objetos pero introduce relaciones entre objetos que están asociadas a las acciones del sujeto.

La abstracción reflexiva se obtiene a partir de lo que Piaget llama coordinaciones generales de las acciones (Piaget 1970, 1977) y como tal su fuente es el sujeto y es completamente interna. Este tipo de abstracción es constructiva; conduce a un tipo muy diferente de generalizaciones y resulta en síntesis de las relaciones previamente construidas que adquieren nuevo significado (Piaget y García, 1982).

Los tres tipos de abstracción no son completamente independientes. Las acciones que conducen a las dos últimas se realizan sobre objetos cuyas propiedades conoce el sujeto a través de la abstracción empírica, y ésta, a su vez, sólo es posible a través de esquemas de asimilación que se construyen mediante abstracción reflexiva.

Esta interacción mutua se puede resumir de la siguiente manera. La abstracción empírica y la semiempírica permiten al sujeto extraer conocimiento de los objetos, ejerciendo acciones sobre ellos. La abstracción reflexiva permite la interiorización y la coordinación de estas acciones para formar nuevas acciones y, finalmente, construir nuevos objetos de conocimiento. Este proceso se repite sobre los nuevos objetos y así, sucesivamente, el sujeto construye conocimiento.

De acuerdo a Piaget el desarrollo de las estructuras cognitivas se debe a la abstracción reflexiva . El mismo interpretó el resultado de muchos experimentos con niños en términos de abstracción reflexiva (Piaget 1977), pero su papel no se restringe al desarrollo intelectual de los niños, su perspectiva incluye las construcciones matemáticas (Beth y Piaget, 1966) y lo considera el método a partir del cual se derivan todas las estructuras lógico matemáticas (Piaget, 1975, 1980). Para apoyar su posición, Piaget trató de explicar varios conceptos matemáticos avanzados en términos de construcciones que resultan de este proceso como el de grupos , la teoría de categorías y el de función. En forma más general, Piaget consideraba que es la abstracción reflexiva en su forma más avanzada la que conduce al tipo de pensamiento matemático por el cual la forma o el proceso se separa del contenido y que los mismos procesos se convierten, en la mente del matemático, en objetos de contenido (Piaget, 1973). La abstracción reflexiva aparece así como una descripción del mecanismo de desarrollo del pensamiento en general.

La primera parte del mecanismo de abstracción reflexiva consiste en extraer propiedades de acciones físicas o mentales a un nivel particular de pensamiento. Aquello que se abstrae se proyecta a un plano más elevado de pensamiento donde se repite el proceso y lo que es más importante, donde se construyen nuevas combinaciones por la conjunción de abstracciones.

Los Procesos del Conocimiento

Los instrumentos del conocimiento dan lugar a los procesos cognitivos. Los procesos cognitivos tienen que ver con la búsqueda de aquellos invariantes que se usan en la definición de los conceptos. La producción o modificación de los conceptos se puede generar por los siguientes procesos:

El análisis de las razones que justifican las abstracciones y las generalizaciones que es el más importante de todos. En matemáticas se pueden distinguir demostraciones que únicamente verifican un teorema y las que además proveen sus razones. Este proceso pone en relieve el papel del sujeto en el conocimiento: es el camino que corresponde a situar todo evento real entre un conjunto de posibles y una necesidad concebida como el único posible actualizado. Ni lo posible ni lo necesario son observables; ambos son el producto de actividades inferenciales del sujeto, por ello la formulación de inferencias o deducción es otro proceso de conocimiento.

Otro proceso importante es la proposición de integraciones y diferenciaciones que conducen al equilibrio dinámico de las asimilaciones y las acomodaciones. Esta proposición se expresa en las relaciones complejas del sujeto que se aproxima sin cesar al objeto y de un objeto que retrocede a medida que los descubrimientos de nuevas propiedades presentan problemas nuevos. En el caso de las matemáticas las integraciones de nuevas estructuras de conjunto van acompañadas de diferenciaciones retroactivas que llegan a introducir nuevas diferenciaciones en el seno de los subsistemas.

Por último la construcción de tematizaciones, es decir, de arreglos lógicos de conceptos en teorías coherentes, dan lugar a un nuevo proceso que es la prolongación de las abstracciones reflexivas a las que ya nos hemos referido.

La transformación de las relaciones y la búsqueda de invariantes ante esas transformaciones permiten para Piaget la definición de los conceptos. Las transformaciones son a su vez fuente de nuevas transformaciones y es a través de ellas que se llega a la tematización o integración de los conocimientos. Se llega así a la equilibración entre diferenciación o búsqueda de los elementos que constituyen los conceptos y la integración de los mismos y de ahí a la explicitación de los conceptos formalizados mediante relaciones de transformación.

Las transformaciones derivadas de explicaciones particulares y locales se van vinculando entre sí para formar estructuras que vuelven a aplicarse a otro nivel iniciándose así nuevamente el proceso de aprendizaje.

El propósito al cual debe aspirar la integración de estos dominios es la promoción de lo que Piaget llama autorregulación, es decir, el proceso de formación de nuevos patrones de razonamiento o lo que es lo mismo, de formas consistentes de enfrentarse a una clase de problemas similares.

La autorregulación involucra el análisis de la situación, la consideración de soluciones tentativas, la evaluación de su efectividad y, si no son efectivas, la construcción de nuevas formas de acercarse al problema.

Los Mecanismos del Conocimiento

Los mecanismos del conocimiento son las operaciones relevantes en la producción de patrones de razonamiento y en la consolidación de estrategias para la solución de problemas, la construcción de modelos y la toma de decisiones.

Estos mecanismos aparecen sin cesar y operan a tres niveles: el *intra-* para el análisis de los objetos, el *inter-* para la transformación de los conceptos y el *trans-* para la construcción de estructuras. Los niveles de esta triada son de carácter funcional, no estructural y describen los aspectos dinámicos del crecimiento del conocimiento.

El nivel *intra-* corresponde al análisis de los eventos u objetos en términos de sus propiedades. Las explicaciones a este nivel son locales y particulares. Las relaciones entre objetos de conocimiento no contienen ningún elemento de necesidad y su forma es muy cercana a la de una simple generalización.

Las relaciones que caracterizan al nivel *intra-* son relaciones internas que consisten en formas aisladas; los elementos internos de esas relaciones no se combinan unos con otros. Este nivel puede ilustrarse con el descubrimiento de alguna acción y el análisis únicamente de sus propiedades o implicaciones internas.

Las limitaciones en este nivel son la ausencia de coordinaciones de unas propiedades con otras para formar un agrupamiento y la existencia de errores que podrán ser corregidos subsecuentemente, así como brechas en las consecuencias que los sujetos son capaces de deducir.

El nivel *intra-* consiste entonces de una concentración en acciones u operaciones repetibles, sin que exista la capacidad de insertarlas en un sistema de condiciones o consecuencias que extendería sus aplicaciones y las incluiría en un sistema de transformaciones, a pesar de que existe ya el inicio de las transformaciones, aunque aplicadas únicamente a los objetos modificados por las acciones involucradas; las transformaciones se enfocan a las acciones particulares que quedan aisladas y son analizadas únicamente en términos de sus propiedades individuales sin tomar en consideración otras acciones u objetos.

En síntesis, las operaciones al nivel *intra-* conducen a definiciones cerradas. Las propiedades de los objetos a este nivel están limitadas a explicaciones locales y particulares.

Las transformaciones derivadas de las relaciones entre objetos ocurren en el nivel *inter-*. En este nivel las instancias diferentes se comparan unas con otras y se establecen correspondencias entre ellas que conducen a la construcción de relaciones y transformaciones.

El nivel *inter-* se caracteriza por correspondencias entre acciones y objetos de conocimiento que en el nivel *intra-* permanecían aisladas y por el descubrimiento de los invariantes requeridos por esas transformaciones. En el nivel *inter-* se encuentra la capacidad de deducir, a partir de una operación inicial, que ha sido comprendida por otras que son implicadas por ella o por coordinarla con acciones u operaciones similares. Esta capacidad permite la formación de agrupamientos de sistemas que incluyen una nueva transformación. A este nivel, sin embargo, no pueden coordinarse en una misma clase elementos lejanos uno del otro. En el nivel *inter-* sólo se pueden hacer coordinaciones paso a paso y por ello cercanamente relacionadas con procesos discursivos.

Mediante correspondencias se llega a la idea de transformación. Una vez que se conoce la transformación se deducen nuevas correspondencias. Las transformaciones a su vez dejan invariantes. Cuando se encuentra una transformación pero no se explica y cuando se tiene un invariante de la transformación se está en el *inter-*.

La explicación de la transformación mediante su incorporación en estructuras de las que las transformaciones son un caso particular es una característica del nivel *trans-*. La comprensión de esas transformaciones en términos de estructuras cognitivas generales se logra en este nivel .

En el nivel *trans-* se logra hacer una síntesis de las transformaciones, una vez que éstas se han dominado y se han generalizado. De esta manera se forman estructuras de conocimiento más complejas. En el nivel *trans-*, las totalidades que anteriormente eran inaccesibles se tornan posibles mediante la adquisición de propiedades sistémicas. Este nivel se caracteriza por la evolución de estructuras cuyas relaciones internas corresponden a las transformaciones que provienen del nivel *inter-*, pero estas transformaciones, en el nivel *trans-*, adquieren una generalización más global en la que se lleva a cabo una síntesis de transformaciones que antes se veían como diferentes. A este nivel se produce una reorganización del conocimiento adquirido en los niveles precedentes y que en cierta forma los incluye, aunque a un nivel de comprensión más elevado.

El equilibrio que se logra mediante los mecanismos *intra-* y los *inter-* son fuentes de múltiples desequilibrios. Las formas de equilibrio dinámico más completas no se logran mas que con mecanismos que implican operaciones en el nivel *trans-* y de equilibración. Estas estructuras se tornan estables en función de conexiones entre transformaciones y a través de los intercambios con el exterior.

Los mecanismos de construcción del conocimiento consisten en el paso entre cada par de estos niveles por una parte y por otra se encuentra el mecanismo general de equilibración. En el paso de una etapa a otra el desequilibrio juega un papel importante pues es éste el que conduce a la síntesis indispensable entre asimilación y acomodación.

Es de suma importancia resaltar, para terminar este breve resumen de la teoría de la génesis del conocimiento elaborada por Piaget y García en la obra antes mencionada, que el paso de un estado a otro de conocimiento obedece siempre a estos mismos mecanismos generales.

A partir de la aplicación de los instrumentos, procesos y mecanismos de conocimiento se pueden distinguir cuatro momentos centrales para el desarrollo de conceptos matemáticos, y particularmente importantes para nuestros propósitos (Dubinsky, 1991):

- La interiorización que se refiere al listado de una sucesión de acciones materiales a un sistema de operaciones interiorizadas.

- La coordinación general de acciones que implica la combinación de dos o más procesos para construir uno nuevo.

- La tematización o encapsulación por la cual las acciones y las operaciones se convierten en estructuras que a su vez se convierten en objetos de la teoría, de forma que ésta se va conformando de estructuras cada vez más fuertes.

- La generalización completiva a través de la cual el sujeto aplica un esquema existente a una colección más amplia de fenómenos.

A estos cuatro tipos, Dubinsky, en la obra entes mencionadas, añade:

- La reversibilidad del proceso original que no sólo implica la posibilidad de revertir un proceso de pensamiento sino su utilización como un medio de construcción de nuevos procesos.

Sección III.2 Elementos relevantes para el aprendizaje de las matemáticas

La teoría psicogenética de Piaget nos permite acercarnos a los procesos de enseñanza y aprendizaje con la idea de enseñar al estudiante a tomar una actitud crítica y activa ante su propio aprendizaje. La principal idea de la epistemología de Piaget nos dice que la construcción del conocimiento sólo se logra mediante la participación activa del sujeto en interacción con el objeto de conocimiento.

El aprendizaje de las matemáticas requiere en particular que el estudiante sea capaz de discriminar, interpretar y organizar datos, relaciones , conceptos y teorías y de articularlos de una forma que posibilte su uso para la solución de problemas de una manera eficiente.

El reconocimiento de los hechos, las acciones sobre los objetos, su interpretación y las reflexiones acerca de las relaciones entre los objetos dan origen a los conceptos o a una disciplina y mantienen un deseo por aprender. En el caso de las matemáticas estos hechos pueden provenir directamente de la experiencia con la realidad, pero, como el mismo Piaget lo dice, pueden ser de naturaleza más abstracta y provenir del ámbito de la matemática misma; por ello son, en ocasiones, difíciles para los estudiantes, particularmente conforme el nivel de abstracción avanza y la cantidad de prerrequisitos conceptuales aumenta.

El significado de las palabras, la definición de funciones y de las operaciones en un campo como el de las matemáticas es de importancia fundamental. El lenguaje de la disciplina modula la adquisición del conocimiento y permite superar el analfabetismo matemático que se presenta no solamente en su forma más burda como la incapacidad de reconocer símbolos, palabras o frases, sino también de una manera más sofisticada en el sentido de que estos símbolos, palabras y frases se usan con una idea pobre o fallida de lo que significan.

Las teorías matemáticas surgen a través de los años a partir de debates, controversias y cambios de dirección. La evolución del conocimiento matemático en el sujeto requiere de que él mismo experimente también controversias y cambios de rumbo que se verán reflejadas en sus procesos de equilibración y desequilibración y en el progreso de sus conceptualizaciones a través de sucesivos procesos de asimilación y acumulación.

El uso de modelos matemáticos y la aplicación de los conceptos y teorías a la solución de problemas específicos dentro del dominio de la propia matemática, o fuera de ella, requieren de diferenciaciones y de integraciones de los conceptos a niveles de abstracción cada vez más elevados y de tematizaciones amplias en las que las relaciones entre todos esos conceptos se entiendan claramente.

Procesos, Instrumentos y Mecanismos en el Conocimiento Matemático

El conocimiento de los conceptos matemáticos evoluciona a partir de la reflexión sobre las acciones que se ejercen sobre los objetos del conocimiento. Estas acciones pueden ser de índole física o mental. Es a través de la reflexión sobre estas acciones que los conceptos se enriquecen, y que se establecen relaciones entre ellos. En muchas ocasiones los cursos de matemáticas se quedan solamente a un nivel en el que se efectúan algunas acciones sobre objetos matemáticos sin permitir que el estudiante reflexione, plantee preguntas, establezca relaciones y las exprese de diversas maneras, cada vez más precisas, hasta llegar a las definiciones formales. Asimismo, las formas de evaluación, muchas veces, se limitan a hacer preguntas únicamente a ese nivel de acciones memorizadas.

La asimilación de los materiales y el proceso de autorregulación en matemáticas requieren de un manejo particular del lenguaje que ayuda al establecimiento de relaciones entre los distintos conceptos. Al generarse estas relaciones entre conceptos, éstos se convierten en algo de naturaleza más dinámica o procesal.

Debido a que la construcción de relaciones lógicas viene de la experiencia y de la acción dialéctica del sujeto sobre los objetos de conocimiento en términos de inferencias, análisis de datos y oportunidades de aplicación, es importante partir de aquéllas relaciones que ya han sido construidas para, mediante procesos de transformación, intentar la identificación de invariantes que conduzcan a definiciones de objetos matemáticos precisos. La formulación matemática de las relaciones permite que el proceso de transformación sea más eficiente y duradero.

Las transformaciones de las diferentes relaciones, conceptos y objetos construidos, se van integrando y diferenciando sucesivamente de tal forma que al ir pasando de un nivel a otro; por ejemplo del nivel *intra-* al *inter-* o del *inter-* al *trans-*, de diferentes maneras, aumenta la densidad del entramado de relaciones. El descubrimiento de invariantes generalizadores, que dan sentido a los conceptos y a las definiciones y que conducen a la tematización en teorías, favorece la estructuración del conocimiento.De esta forma éste cambia y se enriquece.

Los conceptos adquiridos de esta manera pasan a ser, en las estructuras mentales del sujeto, objetos sobre los que nuevamente se pueden llevar a cabo acciones que a su vez se convierten en fuentes de nuevas relaciones y transformaciones para recomenzar el ciclo de aprendizaje a un nivel mayor de abstracción. El paso entre los distintos niveles de manejo de los conceptos matemáticos requiere siempre del proceso de la abstracción reflexiva.

El Papel de la Abstracción reflexiva

Un concepto central en la epistemología de Piaget es el de abstracción reflexiva. En sus trabajos Dubinsky (Dubinsky, 1986, 1989, 1990, 1991, 1993, 1996; Asiala, et al, 1996) afirma que este concepto es una herramienta poderosa en el estudio del pensamiento matemático avanzado y que puede proporcionar una base teórica que contribuya a la comprensión de lo que es el pensamiento matemático y de la forma en la que se puede ayudar a los alumnos a desarrollar las capacidades que éste requiere. Recientemente (Asiala 1997; Clarck,1998), Dubinsky y sus colaboradores han trabajado en la aplicación de este importante concepto al análisis del aprendizaje de las matemáticas al nivel superior y al diseño de material y de métodos de enseñanza efectivos.

La idea de abstracción reflexiva no es simple y en la obra de Piaget no está explicada en un solo lugar; parece ser algo en lo que Piaget trabajó durante mucho tiempo después de completar su trabajo con niños. La mayoría de los trabajos de Piaget se refieren al desarrollo del conocimiento matemático en la edad temprana. Rara vez trató el tema para edades posteriores a la adolescencia por ello es interesante analizar la manera como Dubinsky y sus colaboradores han intentado extender este concepto a los temas más avanzados mostrando que es posible, no sólo conjeturar y discutir, sino obtener evidencia que sugiere que conceptos como los de inducción matemática, función, cálculo de predicados y espacios topológicos, derivadas o teoría de grupos se puedan analizar en términos de extensiones de las mismas nociones que Piaget usó para describir cómo los niños construyen el conocimiento de conceptos como proporción, medición y otros de la aritmética elemental.

Para desarrollar la noción de abstracción reflexiva para el pensamiento matemático avanzado, Dubinsky aísla lo que considera sus elementos fundamentales. Para ello usa la noción de esquema como una colección coherente de objetos, procesos y relaciones entre ellos. De esta manera la tendencia de un sujeto a invocar un esquema para entender, manejar y organizar un problema puede asociarse con su conocimiento de un concepto particular de las matemáticas.

Cada persona posee un número de esquemas de acuerdo a su conocimiento matemático. Estos esquemas están interrelacionados en una organización compleja. Utilizando la idea de abstracción reflexiva conjuntamente con los conceptos propios de las matemáticas como disciplina de estudio, se pueden aislar porciones de esa estructura compleja para encontrar descripciones explícitas de las posibles relaciones entre esquemas y generar de esta manera una descomposición genética de cada concepto o grupo de conceptos.

La descomposición genética de un concepto es un instrumento de análisis que permite interpretar los datos obtenidos del estudio de la forma en la que los estudiantes responden ante situaciones matemáticas específicas y que puede modificarse de acuerdo a los datos obtenidos de la investigación con alumnos, hasta llegar a lo que se considera una interpretación adecuada de los mismos. Al igual que otras herramientas de análisis, esta descomposición genética no es única, es decir, diversos investigadores pueden construir descomposiciones genéticas diferentes de un mismo concepto, ya que no pretende ser una descripción exacta de la forma en la que se aprende el concepto. Sin embargo, la descomposición genética representa una forma razonable de acercarse a la manera en la que los estudiantes pueden construir dicho concepto matemático.

Si bien no es posible observar directamente los esquemas, objetos o procesos que un individuo posee, éstos se pueden inferir a partir de observaciones de lo que el individuo hace al enfrentar un problema o al tratar de entender un fenómeno. Son estos actos de interpretación, operación, solución a problemas y de planteamiento de nuevas preguntas los medios mediante los cuales las personas construyen conocimientos matemáticos nuevos.

Se puede afirmar que el conocimiento matemático consiste de una colección de esquemas que no podemos especificar directamente; todo lo que podemos hacer es observar la propensión del sujeto a responder a cierto tipo de problemas de una manera relativamente consistente que podemos describir en términos de esquemas. Cuando el individuo tiene éxito decimos que el problema ha sido asimilado por los esquemas. Cuando no lo tiene, en condiciones favorables, sus esquemas pueden acomodarse para manejarlo y es, en este proceso, que entra en juego la abstracción reflexiva.

Para describir cómo se pueden construir nuevos esquemas a partir de los existentes se pueden usar los cinco tipos de construcción mencionados anteriormente: interiorización, coordinación, tematización, generalización completiva y reversibilidad, que se ejemplifican a continuación.

Una parte importante de la comprensión de un concepto matemático es la construcción de un proceso mediante el cual se pueden obtener resultados de la aplicación del concepto. Esto se logra mediante la interiorización. Un ejemplo de esto, en el caso del concepto de función (Dubinsky, 1989) se presenta cuando el sujeto responde a una situación en la que aparece una función construyendo un proceso mediante el cual se obtiene el valor de la función para un valor particular de su dominio.

La coordinación requiere la combinación de dos o más procesos. Esta puede ilustrarse, nuevamente para el concepto de función: en la composición de funciones se requiere de la combinación de dos procesos, cada uno de ellos correspondiente a una de las funciones, para dar lugar a un nuevo objeto matemático: la función composición.

Un ejemplo de tematización se puede ilustrar con el concepto de integral indefinida. En el cálculo de la integral definida se obtiene un número, en cambio, cuando se calcula la integral indefinida se obtiene como resultado una nueva función; el éxito en la realización de esta nueva operación requiere que el proceso del cálculo del área y de la integral definida se haya convertido, para el sujeto, en un objeto que puede variar cuando uno de sus parámetros cambia.

La generalización completiva es la forma más común generada por la abstracción reflexiva. Por ejemplo el concepto de función, que inicialmente se introduce para transformar números, se generaliza para incluir funciones que transforman otro tipo de objetos como matrices o conjuntos.

La reversibilidad de un proceso también aparece en un gran número de actividades en matemáticas: la resta, la división, la función inversa, son algunos conceptos, entre otros muchos, que en su construcción requieren de un proceso de reversión de otro proceso.

Es posible proponer como hipótesis, que la construcción de todos los conceptos matemáticos se puede describir en estos términos. Con base en esta hipótesis la teoría epistemológica de Piaget puede usarse en el diseño de instrumentos de evaluación y de enseñanza y, en particular, se puede emplear para diseñar un modelo diagnóstico para el aprendizaje de las matemáticas. Para lograrlo, es conveniente tomar en cuenta los elementos antes mencionados conjuntamente con las características propias no únicamente de la matemática sino del concepto específico con el que se pretende trabajar.

Presupuestos del modelo aplicado a matemáticas

Los presupuestos del modelo de diagnóstico que podemos plantear en base a la teoría de Piaget en torno al aprendizaje de los conceptos matemáticos son los siguientes:

- El conocimiento matemático se construye a través de la acción dialéctica del sujeto sobre los objetos matemáticos.

- En esta interacción el sujeto enfrenta situaciones problemáticas ante las cuales, y mediante los procesos de asimilación y acomodación, modifica sus estructuras de conocimiento matemático.

- En los procesos de desequilibraciones y equilibraciones sucesivas requeridos para la construcción de los conceptos matemáticos, la abstracción reflexiva juega un papel preponderante para lograr el paso de un nivel a otro en el conocimiento de cada concepto y teoría.

- Aunque la evolución de cada sujeto sea diferente, las respuestas de sujetos en el mismo nivel presentan elementos comunes y distinguibles y éstos se pueden emplear para construir instrumentos de diagnóstico. Cada estructura incluye a las anteriores y supone una mejora con relación a ellas.

Finalidad

La epistemología de Piaget, aunada a la aplicación que de ella hacen Dubinsky y sus colaboradores, permite desarrollar una descomposición genética para cada concepto matemático. Ésta es una herramienta que permite establecer una secuencia que describe una forma posible en la que se da el aprendizaje del concepto y que puede ser probada experimentalmente, refinándose hasta el punto en el que el diseño de los ítems relativos al concepto sea viable y efectivo.

La descomposición genética permite:

- Investigar en cuál nivel de la descomposición propuesta se encuentra cada alumno. Conviene aclarar aquí que, dada la no unicidad de la descomposición genética, los niveles obtenidos están dados en términos de la que se está utilizando en un caso particular.

- Determinar qué tipo de noción del concepto tiene el alumno.

- Evaluar las posibilidades del alumno en cuanto a su profundización en el conocimiento del concepto.

- Analizar las relaciones que el alumno ha establecido entre diferentes conceptos

- Analizar el tipo de estrategias de aprendizaje que el alumno requiere para pasar a un nivel más profundo de conocimiento del concepto.

Objetivos

A partir de la descomposición genética de un concepto se pueden elaborar exámenes, cuestionarios, instumentos de observación o guiones de entrevista clínica con la finalidad de:

- Aportar información objetiva para agrupar a los alumnos que poseen un mismo tipo de concepción relativa al concepto en estudio.

- Pronosticar las posibilidades de un alumno para continuar estudios que requieren de conceptos matemáticos más avanzados.

- Adecuar los aprendizajes de las matemáticas con el objeto de que la mayor cantidad de alumnos posible pueda lograr una verdadera y profunda comprensión de los conceptos matemáticos.

Valoración

Las ventajas que presenta este tipo de diagnóstico son las siguientes:

- Abre el camino para la investigación en la enseñanza de las matemáticas aportando datos sobre posibles formas en las que los estudiantes las aprenden.

- Permite la estructuración de secuencias didácticas y el diseño de situaciones problemáticas adecuadas para guiar a los estudiantes por el camino de un aprendizaje serio y crítico de las matemáticas.

- Posibilita la utilización de las matemáticas aprendidas en el análisis y solución de problemas en otras disciplinas.

Entre las limitaciones que este tipo de modelo diagnóstico presenta se encuentran las siguientes:

- La elaboración de descomposiciones genéticas para cada concepto matemático supone un trabajo teórico y experimental arduo y laborioso.

- Los instrumentos pueden ser difíciles de construir y de estandarizar.

- Requiere de el uso de instrumentos personalizados, al menos en la etapa de investigación, que son difíciles de analizar y requieren de mucho tiempo para hacerlo.

- Se requiere de un entrenamiento en los elementos del modelo, en matemáticas y en los métodos requeridos para el análisis de los datos.

Proceso del diagnóstico

En el proceso del diagnóstico interesa entender la forma en la que el estudiante maneja y aborda los problemas matemáticos para dilucidar qué tipo de concepción de los conceptos tiene y cuáles son sus posibilidades de avance hacia niveles superiores, más que contar con una puntuación determinada que caracteriza a cada estudiante. En el diagnóstico es fundamental conocer las estrategias cognitivas del sujeto y sus posibilidades de modificación para poder conducirlo a niveles superiores de conocimiento.

En este proceso, el investigador toma necesariamente un papel activo para poder encontrar las razones por las cuales el sujeto adopta ciertos métodos o ciertas definiciones. El investigador observa y toma nota de las acciones del sujeto para su análisis posterior. Para ello es necesario combinar los métodos experimental y clínico de investigación; ellos permiten encontrar, a través del uso de la descomposición genética en el caso de las matemáticas, el tipo de concepción que tiene el alumno de la manera más precisa posible.

La exploración en todo este proceso tiene el fin de aclarar el tipo de concepción que tiene el estudiante respecto a la comprensión de un concepto o tema de las matemáticas, lo cual permite tomar decisiones más confiables a fin de guiarlo en el camino hacia niveles superiores de comprensión.

Si el diagnóstico se hace con fines de evaluación o de clasificación, el modelo presentado anteriormente permite la obtención de un mayor número de elementos que guíen la calificación y clasificación de los estudiantes.

Sección III.3 Un modelo para el diagnóstico del aprendizaje de las matemáticas

La teoría psicogenética afirma que los conceptos matemáticos, al igual que los de otras disciplinas, se construyen mediante la participación activa del sujeto. Algo que diferencía a los conceptos matemáticos de los de otras disciplinas es su naturaleza abstracta, desligada en muchas ocasiones de la realidad empírica. Esta naturaleza de los conceptos matemáticos es la que hace que se les considere inaccesibles, o al menos difíciles de aprender.

La enseñanza tradicional de las matemáticas no ha abordado frontalmente los problemas que los alumnos tienen cuando intentan trabajar con problemas de esta disciplina. La investgación que se ha realizado recientemente en el campo de la didáctica de las matemáticas muestra claramente que esa forma de enseñar no ha sido efectiva, que los alumnos no comprenden cabalmente lo que se intenta enseñar y que en lugar de aprender el significado de los conceptos, los estudiantes se limitan a memorizar fórmulas o algoritmos que difícilmente pueden emplear en situaciones ligeramente diferentes a las que se ejemplifican en el salón de clase. La enseñanza tradicional se ha limitado a dejar el aprendizaje en una etapa superficial en la que los conceptos matemáticos se usan memorísticamente sin que haya comprensión real de su significado. La evaluación y el diagnóstico del conocimiento matemático se han mantenido asimismo en este nivel, contribuyendo de esa manera, probablemente, a que por mucho tiempo no se haya cuestionado la eficacia de los procesos de enseñanza de las matemáticas.

En la búsqueda de formas más efectivas de enseñar los conceptos de las matemáticas, en particular aquéllos correspondientes a las matemáticas avanzadas, que requieren un profundo nivel de abstracción, el diagnóstico juega un papel importante. El diagnóstico es un proceso que permite acercarse a la forma en que los estudiantes usan y construyen sus conocimientos matemáticos.

El concepto de abstracción reflexiva, así como la dialéctica sujeto–objeto del constructivismo proporcionan un marco conceptual que se puede usar, como se mencionó en la sección anterior, como punto de partida para describir, en principio, el aprendizaje de cualquier concepto matemático conjuntamente con su adquisición. Éste es el primer elemento necesario en el análisis diagnóstico.

No bastan, sin embargo, esos elementos para diseñar y analizar instrumentos de diagnóstico o de evaluación. Es necesario concretar los elementos de la interacción sujeto – objeto y de la abstracción reflexiva en términos de los requerimientos del aprendizaje de los conceptos matemáticos que se incluirán en los instrumentos y del nivel de los estudiantes que participarán en dicho proceso.

En el proceso de diagnóstico se requiere de una descripción particular del estado de conocimiento de los sujetos. Para encontrarla es necesario analizar y entender lo que implica la construcción del conocimiento de las matemáticas. Conjuntamente con los dos elementos antes mencionados es posible sentar las bases para obtener una descripción de cualquier concepto matemático que pueda ser útil en el diseño de instrumentos de medida y de diagnóstico, así como de procedimientos didácticos útiles en el aula.

El resultado del uso per se de la abstracción reflexiva y el modelo de construcción de conceptos matemáticos es quizás demasiado ex post facto para esperar que tenga alguna relación real con la forma en que los estudiantes llevan a cabo el proceso de construcción de dichos conceptos. Para ello es necesario agregar un tercer elemento esencial en el estudio de cualquier concepto: el esfuerzo de observación de los estudiantes durante su proceso de construcción de los conceptos matemáticos.

Es la síntesis de estos tres ingredientes lo que constituye una base para lograr diseños adecuados de instrumentos de medida y principios confiables para el análisis de los datos que se obtienen de la aplicación de tales intrumentos.

En los últimos años un grupo de investigadores (Asiala et al, 1996) ha desarrollado un marco teórico para la conducción de la investigación en enseñanza de las matemáticas. A partir de las bases del constructivismo de Piaget se intenta llevar a cabo investigaciones cualitativas y cuantitativas que permitan, por una parte, el desarrollo del propio marco teórico y, por otra, profundizar en la comprensión de la forma en la que los conceptos matemáticos se aprenden y así poder avanzar en el diagnóstico, la evaluación y la instrucción de dichos conceptos.

El análisis de cualquier concepto matemático a la luz de esta propuesta conceptual requiere de un primer análisis de los conceptos matemáticos en términos de los elementos relacionados con la interacción sujeto –objeto y con la abstracción reflexiva requeridos en su construcción. A este análisis se le denomina, como ya se ha mencionado, descomposición genética del concepto.

El estudio de la forma en que un individuo aprende un concepto matemático en particular se lleva a cabo mediante refinamientos sucesivos de la descomposición genética del concepto conforme el investigador interactúa, a través de la observación y el análisis de instrumentos, con el conocimiento de los alumnos. La investigación se inicia con un análisis teórico de la epistemología del concepto en cuestión, es decir, con lo que significa para los investigadores entender el concepto y la forma en la que esa comprensión puede ser construida por el sujeto. La descomposición preliminar se basa en el análisis epistemológico del concepto llevado a cabo por los investigadores expertos. Este análisis se aplica al diseño de instrumentos y de instrucción; de su aplicación se obtienen datos que permiten reconsiderar la descomposición genética original de manera que sea más compatible con los datos experimentales. Este procedimiento puede repetirse tantas veces cuantas se considere necesario para que la descomposición genética generada refleje de manera adecuada los datos empíricos.

Los objetivos de este marco teórico de investigación consisten, como ya se ha dicho con anterioridad, en incrementar la comprensión de la manera en la que la construcción de los distintos conceptos matemáticos se lleva a cabo y en desarrollar diseños de instrucción y técnicas de evaluación y diagnóstico que permitan un mejor aprendizaje de los conceptos matemáticos.

La descomposición genética de un concepto es una descripción en términos de la teoría y con base en datos empíricos acerca del aprendizaje de los conceptos matemáticos involucrados y de la forma en la que el sujeto puede hacer la construcción que lo conducirá a su comprensión. Un concepto no tiene necesariamente una sola descomposición genética, ni ésta es tampoco la forma exacta en que el sujeto lo aprende, pero la descomposición genética de un concepto incluye la observación preliminar del proceso de aprendizaje y sirve como una buena guía para el diagnóstico y para el diseño de actividades para usarse en su instrucción.

El análisis que lleva a una posible descomposición genética incluye todos los elementos de la teoría de Piaget que se describieron en la sección anterior, haciendo una adecuación pertinente al contexto de las matemáticas. Esta descomposición debe componerse de una síntesis de los datos obtenidos de la observación de estudiantes mientras intentan entender el concepto en cuestión; datos que se pueden obtener mediante una diversidad de instrumentos: tests, entrevistas, trabajos en clase, etc., conjuntamente con el marco conceptual antes mencionado y un análisis detallado de los conceptos matemáticos involucrados.

El modelo diagnóstico resultante puede pensarse en forma sistémica ya que no es un modelo lineal sino un sistema con retroalimentación y consiste de la organización de los cinco tipos de construcción analizados en la sección III.2.

El uso de un marco teórico de esta naturaleza plantea algunos problemas importantes a considerar. Por una parte cabe preguntarse cómo puede desarrollarse la descomposición genética del concepto en cuestión; por otra parte se encuentra el problema usual en pedagogía de cuál puede ser la relación entre el marco teórico y el grado en el que el análisis de los datos proporciona una visión adecuada de lo que ocurre en la mente de quien aprende el concepto.

Dentro de este marco teórico se reconoce que lo que una persona sabe y es capaz de hacer no es necesariamente asequible en un momento dado o en una situación determinada. En el problema del conocimiento entran en juego dos componentes: el aprendizaje del concepto y la posibilidad de emplearlo cuando es necesario.

La reflexión es una parte importante del aprendizaje y del conocimiento. Las matemáticas en particular contienen muchas técnicas y algoritmos que deben usarse frente a situaciones de naturaleza diferente. Una gran cantidad de personas logran aprender estos procedimientos, pero la comprensión de los conceptos matemáticos va más lejos, sobrepasa la posibilidad de hacer cálculos, independientemente de cuán sofisticados sean; es necesario ser consciente de la forma en la que estos procedimientos funcionan y tener una idea de cuál puede ser el resultado de su puesta en marcha, sin necesidad de llevar a cabo todos los pasos específicos; es necesario asimismo tener la posibilidad de trabajar con variaciones de cada algoritmo y entender las relaciones entre diversos procedimientos.

De lo considerado anteriormente puede pensarse que el conocimiento matemático consiste en una tendencia a llevar a cabo construcciones mentales frente a situaciones problemáticas. La evolución del conocimiento matemático se lleva a cabo mediante la reconstrucción de situciones con las que previamente se ha trabajado. La reconstrucción contiene al conocimiento anterior pero no en una forma lineal sino en una forma compleja que implica la construcción de nuevas relaciones y nuevas formas de ver el problema, asociadas con lo que Piaget llama asimilación y acomodación.

La comprensión de un concepto matemático se inicia con la posibilidad de manipular mental o físicamente un objeto de conocimiento previamente construido.

Un modelo diagnóstico basado en este marco teórico y en esta idea de lo que significa el conocimiento matemático debe contener los elementos necesarios para que la descomposición genética tenga sentido como instrumento de análisis y de diseño. En este marco teórico se parte de definiciones específicas de lo que se supone constituyen escaños indispensables en el proceso de construcción del conocimiento matemático: las acciones, los procesos y los objetos.

Acciones, procesos y objetos

Los objetos matemáticos pueden referirse a cualquier tipo de concepto matemático que un individuo construye en algún momento de la evolución de su conocimiento de esta disciplina.

Cuando se trabaja con matemáticas existen una gran cantidad de acciones que un sujeto puede usar para manipular esos objetos o para hacer cálculos con ellos. Estas acciones van más allá del cálculo cuyo resultado es una respuesta numérica. Es posible que un sujeto trabaje de otra manera utilizando acciones sobre los objetos; por ejemplo ejecutando paso a paso un algoritmo, aplicando directamente algún teorema o insertando números en una fórmula conocida.

Por nivel de conceptualización **acción** se entiende un conjunto de recetas o actividades en las que el sujeto se puede involucrar utilizando un concepto matemático. Estos procedimientos pueden ser complicados e incluir decisiones en el camino pero lo que distingue a este nivel es que éstas se toman en base a reacciones o estímulos que el sujeto percibe como externos. Esto significa que el individuo requiere de instrucciones claras, precisas y comprensibles que le den los detalles o los pasos a seguir conforme el procedimiento avanza. Esta estructura de conocimiento es completamente local, es decir, el sujeto solamente la entiende localmente, no percibe el panorama total que involucran sus acciones y no razona sobre el procedimiento en forma global. El sujeto sabe qué hacer en cada paso, pero no mucho más que eso. Por ejemplo, un estudiante que puede resolver paso a paso una ecuación algebraica de primer grado pero es incapaz de hacerlo saltando algunos pasos, o un estudiante que puede resolver ecuaciones cuadráticas únicamente usando una fórmula, está restringido a un nivel de tipo acción en su manejo de la incógnita. Para este estudiante la factorización de ecuaciones cuadráticas o la solución de ecuaciones lineales que involucran parámetros presentan dificultades imposibles de superar. Dentro de este marco teórico las dificultades del individuo se deben a que esos nuevos problemas requieren de un nivel de comprensión del concepto más profundo o más abstracto que el que maneja en ese momento.

El nivel de acción es un nivel en el que el manejo matemático del concepto es limitado; sin embargo, juega un papel sumamente importante en el inicio de la construcción de los nuevos conceptos.

Cuando el sujeto reflexiona sobre las acciones, éstas se interiorizan, pueden concebirse como algo interno, propio del individuo y sobre el cual puede tener algún control. Las acciones se convierten en procesos. La interiorización permite al sujeto ser consciente de una acción, reflexionar sobre ella y combinarla con otras acciones.

Una concepción a nivel **proceso** involucra la habilidad del sujeto de manipular con mayor libertad objetos cognitivos. En este caso el sujeto es consciente del procedimiento en forma más global y puede manipular los conceptos que conoce, compararlos y transformarlos en su mente, sin ayuda de apoyos que él concibe como externos. Por ejemplo, en el caso del concepto de función una concepción de tipo proceso permite al sujeto percibir a la función como algo que puede recibir una o más entradas o valores de la variable independiente, efectuar una o más operaciones sobre esas variables y obtener los resultados correctos como resultado de la aplicación de la función a esos valores. En este nivel el sujeto no necesita llevar a cabo explícitamente todas las operaciones, le es posible imaginarlas con mayor o menor detalle y saltarse algunos pasos sin correr el riesgo de cometer un error. De esta manera, un estudiante a este nivel puede manipular una expresión algebraica o combinar varias de ellas en una sola, hacer una composición de funciones e incluso revertir el proceso que ha aplicado a una función para obtener la función inversa.

Es posible trabajar con procesos para formar nuevos procesos. Esto se logra por ejemplo mediante la reversión. Un caso de ella ocurre cuando un estudiante de cálculo puede haber interiorizado la acción de obtener la derivada de una función y ser capaz de hacerlo para un gran número de ejemplos usando varias técnicas. Si el proceso se interioriza el estudiante debe ser capaz de revertirlo para resolver problemas en los que se da una función y se desea encontrar una función cuya derivada es dicha función.

También se pueden conseguir nuevos procesos mediante la coordinación de varios de ellos. Por ejemplo la coordinación de los procesos de diferenciabilidad y de linearidad se pueden coordinar para llegar a la idea de función lineal derivable. Esta coordinación permite extender y reorganizar los procesos utilizando nuevamente la abstracción reflexiva para, a través de la acomodación, construir nuevos esquemas.

Por otra parte, es posible reflexionar sobre un proceso e interiorizarlo de tal forma que se convierta en un objeto, para ello se requiere de la encapsulación. Cuando el individuo no sólo es consciente de los pasos individuales del proceso, sino es consciente del proceso como una totalidad y puede avanzar y retroceder revirtiendo las actividades mentales, realizando transformaciones o actuando mediante acciones o procesos sobre él, se dice que el sujeto tiene una comprensión de nivel **objeto** de dicho concepto.

En este caso puede decirse que el proceso ha sido encapsulado en un objeto. Cuando se efectúan acciones o procesos sobre un objeto a menudo es necesario desencapsularlo para recuperar el proceso que le dió origen, para usar sus propiedades y para tener mayores posibilidades de manipulación. En el caso del álgebra elemental, cuando la variable se concibe como un objeto es posible que el sujeto sea capaz de comprender los problemas asociados al manejo de polinomios, por ejemplo, pensando en el problema de calcular las raices de un polinomio convirtiendo la expresión algebraica que lo representa en una ecuación en la que la variable puede manipularse para obtener la solución; en este caso es necesario desencapsular el concepto de polinomio para hacer uso de las propiedades y de los procesos que le dieron origen para encontrar el valor de la variable que hace que la ecuación sea válida en una situación particular. En general puede decirse que la encapsulación de un proceso en un objeto es algo difícil de lograr (Sfard, 1989; Sfard, 1994).

Una vez construidos los objetos y los procesos se pueden relacionar entre ellos de diversas maneras. Esto conduce a la construcción de estructuras de conocimiento más organizadas que conforman esquemas conceptuales. En la construcción de los esquemas intervienen los mecanismos del conocimiento de los que se ha ya hablado con anterioridad.

Los esquemas pueden a su vez ser tratados como objetos y nuevamente a través del paso por los niveles de acción, proceso y objeto llegar a la construcción y organización de esquemas de un nivel superior. Cuando esto ocurre puede decirse que el esquema ha sido tematizado. Por ejemplo, el concepto de variable en el álgebra elemental puede extenderse a concebir conjuntos en los que las variables son funciones, polinomios o matrices sobre los cuales se pueden efectuar operaciones y definir propiedades, llegando así a la construcción del esquema de espacio vectorial que puede aplicarse a conceptos como el de transformaciones lineales o espacios duales.

En muchas ocasiones es importante que el sujeto pueda pensar alternativamente en una misma entidad matemática como proceso y como objeto.

Uso de estos elementos en el diagnóstico

Las acciones, procesos, objetos y esquemas que pueden definirse con base en el análisis teórico de los conceptos matemáticos y en los datos recabados mediante el uso de instrumentos destinados a validar la descomposición genética y a refinarla, permiten definir etapas por las que el sujeto debe pasar en la construcción de un concepto determinado y permiten también decidir en cuál de esas etapas se encuentra un sujeto en un momento específico.

La descomposición genética derivada del marco teórico aquí presentado permite por una parte el diseño de instrumentos para diagnosticar la posición del estudiante a lo largo de las etapas cognitivas que señala la descomposición y por otro el análisis de los resultados obtenidos de la aplicación de los mismos. Provee además información que ayuda a la toma de decisiones respecto a los pasos necesarios para guiar al estudiante hacia una mejor comprensión del concepto y para el desarrollo de aquellas actividades que se considere deben llevarse a cabo para inducir esta mejora.

Si se pretende conocer algo acerca de lo que una persona sabe de algún concepto matemático el procedimiento a seguir consiste en ponerla ante una situación problemática, observar lo que ocurre, es decir ver cómo trabaja, qué tipo de argumentos usa, etc. y tratar de inferir algo sobre los esquemas que parece estar usando. Sin embargo, ni los esquemas ni las situaciones problemáticas pueden especificarse con precisión. Inclusive, el hecho de que una persona use un esquema en una situación no es una garantía de que usará uno similar en lo que al observador le parece será una situación análoga, aun cuando las dos situaciones ocurran en un tiempo cercano.

Lo que es importante en el diagnóstico es comprender aquéllo que determina la naturaleza de la respuesta de un individuo en una situación dada. Esto es lo que pretende el modelo diagnóstico desarrollado en este trabajo. Modelo que se desarrolla alrededor de las construcciones básicas que un sujeto hace mediante el proceso de abstracción reflexiva, que conduce a través de los niveles de acción, proceso y objeto a la construcción de los conceptos matemáticos.

En la enseñanza tradicional muy pocos estudiantes llegan a desarrollar una concepción a nivel objeto de los conceptos matemáticos. La mayoría se queda incluso a nivel de acción en su comprensión de dichos conceptos. Este tipo de concepción no permite la aplicación de los conceptos a situaciones nuevas, aun en el caso en que sean muy similares a aquéllas mediante las cuales fue aprendido el concepto; además, los conceptos aprendidos a este nivel se olvidan con facilidad.

La investigación en enseñanza de las matemáticas muestra que actualmente el conocimiento profundo de las matemáticas es un caso atípico. En términos del marco teórico que aquí se utiliza esto podría expresarse diciendo que muy pocos estudiantes logran alcanzar una concepción a nivel objeto de los conceptos matemáticos que estudian en la escuela a través de didácticas de corte tradicional. Por supuesto, lo deseable es lograr que la mayoría de los estudiantes alcancen un nivel de conceptualización al menos a nivel proceso de los conceptos, que entiendan lo que significan y los puedan aplicar. El caso ideal consistiría en lograr que un gran número de estudiantes terminaran sus estudios con una concepción a nivel objeto de todos los conceptos matemáticos estudiados.

Al enfrentar un problema matemático los estudiantes no pueden, con cierta frecuencia, aplicar las ideas específicas de las que disponen. Su conocimiento matemático consiste en la tendencia a recordar procedimientos memorizados carentes de significación. De acuerdo al marco teórico que se presenta en esta sección esto puede explicarse en términos de la no linealidad de la construcción de los conceptos matemáticos. El conocimiento matemático crece a través de avances, retrocesos, altos en el camino, comprensiones parciales y en ocasiones no totalmente correctas, síntesis, construcción de relaciones, generalizaciones y deducciones en las que se vuelve una y otra vez al mismo concepto.

Puede ser difícil decidir si una actividad particular de un sujeto lo sitúa contundentemente en un nivel de acción, de proceso o de objeto. La distinción no es intrínseca al concepto matemático, tiene que ver con la relación entre el sujeto y una idea. Por ello la entrevista clínica y la aplicación de instrumentos diversos juega un papel importante en el diagnóstico y en la evaluación basadas en la descomposición genética. Todo este procedimiento debe hacerse con mucho cuidado. El análisis de los resultados obtenidos de la aplicación de los instrumentos resulta también laborioso y requiere de mucho tiempo y dedicación. La decisión que se tome con base a este análisis afecta el camino que seguirá el estudiante y por ello debe llevarse a cabo de manera muy responsable.

La idea global de un proceso de diagnóstico de esta naturaleza consiste en que la información obtenida arroje luz sobre las posibles construcciones mentales de los estudiantes. Esta finalidad no puede perderse de vista. Para lograrlo, los instrumentos escritos pueden complementarse con entrevistas y así llegar a un diagnostico más acertado de las capacidades de cada estudiante.

La construcción del conocimiento en este sentido puede representarse mediante una espiral ya que las acciones que se interiorizan para construir procesos y puesto que éstos a su vez se interiorizan para construir nuevos objetos, se repiten una y otra vez a niveles mayores de abstracción.

Este modelo diagnóstico puede concebirse en tres etapas: análisis del concepto, diseño de instrumentos y observación de los estudiantes.

La primera parte se logra a través de una primera descomposición genética del concepto. Esta debe hacerse por expertos en la enseñanza de las matemáticas en el marco de la teoría antes señalada. A partir de los datos obtenidos de primeros instrumentos y de la observación de los estudiantes cuando tratan de entender este concepto, la descomposición genética puede revisarse y refinarse hasta llevarla a una forma en la que sea útil en la medida del nivel en el que los estudiantes se encuentran respecto a ese concepto en particular.

Es importante enfatizar que el análisis matemático del concepto en términos del modelo aquí propuesto acerca de las construcciones mentales que un individuo puede llevar a cabo al aprenderlo, es una conjetura. El objetivo de esta conjetura consiste en intentar dar sentido a las reacciones de un sujeto ante situaciones problemáticas dentro del campo de las matemáticas, sustentándose en la epistemología de Piaget que ha demostrado su potencial de aplicación en situaciones muy variadas. Por otra parte, en este modelo no se toman las reacciones de los estudiantes ante los problemas como punto de partida de inferencias que serían difíciles de verificar; el uso de este modelo considera únicamente el hecho de que la descomposición del concepto sea compatible con las respuestas observadas, es decir, si el análisis teórico proporciona una explicación satisfactoria del trabajo de los estudiantes.

El modelo de diagnóstico discutido en esta sección ha sido aplicado con éxito, como ya se ha mencionado, a la comprensión de la construcción de varios conceptos matemáticos muy diversos y a la explicación de los obstáculos que los estudiantes enfrentan durante el proceso de construcción.

En las siguientes secciones se abordará la descripción de las dificultades de los estudiantes ante el concepto de variable en el álgebra elemental y se revisará lo que se conoce acerca de estas dificultades a través de la investigación para proceder posteriormente a la presentación de una descomposición genética de dicho concepto. La descomposición se aplicará al análisis de ítems destinados a probar la efectividad del modelo matemático de medida presentado en el capítulo II.

Sección III.4 El concepto de variable en el álgebra: algunos resultados de la investigación

Las matemáticas que se enseñan en el nivel universitario, como el Cálculo Diferencial e Integral, la Probabilidad y la Estadística demandan del estudiante una gran capacidad de abstracción. Los conceptos que las constituyen son complejos y muy estructurados, lo que hace que su comprensión sea imposible si no se sustenta sobre una base muy sólida constituida por las ideas más elementales del álgebra y de la geometría.

El caso del álgebra es particularmente importante. Esta materia es básica para un buen manejo de las nociones de las matemáticas avanzadas y es por ello que se le dedica mucho tiempo de instrucción al nivel de la enseñanza secundaria. Sin embargo, los estudiantes que ingresan a la universidad y que han cursado al menos cinco años de matemáticas, con un gran énfasis en el estudio del álgebra, muestran dificultades y una falta de comprensión de los conceptos elementales de esta disciplina.

Los resultados de los exámenes de admisión a las universidades, los resultados de las evaluaciones de los alumnos en los cursos propedéuticos de carácter remedial que se imparten en ellas y los resultados de la investigación muestran que sus conocimientos de álgebra son superficiales.

La identificación de los problemas específicos de los estudiantes en esta materia y de sus causas, así como la búsqueda de formas alternativas de enseñanza diseñadas para superar estos problemas son temas que requieren de investigación diseñada específicamente con este propósito.

Con el fin de profundizar en la comprensión de las dificultades reales que los estudiantes presentan al enfrentarse a la solución de problemas algebraicos, es conveniente poner atención a un concepto central a esta materia y que puede ser la fuente de dichas dificultades: el concepto de variable.

El manejo de expresiones, la solución de problemas y la modelación de situaciones reales mediante el uso del álgebra requieren de una comprensión global y flexible del concepto de variable. No sería exagerado decir que, a nivel elemental, el álgebra gira en torno a esta idea. Sin embargo, a pesar de su papel protagónico, este concepto es muy difícil de definir. La variable en el álgebra se presenta con caracterizaciones que varían según el problema en el que ésta está involucrada.

La investigación concerniente al aprendizaje del álgebra se ha llevado a cabo fundamentalmente con niños de 12 a 16 años, edad en la que usualmente se estudia álgebra en la escuela. Los resultados obtenidos (English y Sharry, 1996; Reggiani,1994; Usiskin, 1988; Ursini, 1996; Bernarz et al, 1991; Filloy y Rojano 1984; Filloy y Rojano 1989; Gascon, 1984; Ruiz y Rodriguez, 1994; Brust, 1997) indican que los estudiantes, independientemente de que inicien apenas sus estudios o hayan llegado a niveles más avanzados, tienen dificultades al manejar todas las caracterizaciones de la variable. En un estudio reciente (Lozano, 1998) se encontró que la comprensión de los estudiantes que se encuentran en distintos grados de la educación secundaria es muy limitada, que únicamente pueden trabajar con problemas muy simples, que un aumento ligero en la complejidad de los problemas induce en los estudiantes un comportamiento que parece responder más a la presencia de ciertos símbolos contenidos en la pregunta misma que a la reflexión sobre el problema, que los estudiantes tienen muchas dificultades para simbolizar problemas verbales utilizando variables y que la noción de relación entre variables les causa asimismo muchas dificultades. El comportamiento de los estudiantes no cambia a su paso por el sistema escolar, aun cuando estén cursando la materia de álgebra en ese momento, y el tipo de respuesta que los estudiantes de distintos niveles dan a los ítems presentados se mantiene prácticamente invariable. Estos resultados ponen de manifiesto que la enseñanza del álgebra en la escuela es ineficaz y no ayuda a los alumnos a lograr la abstracción que se requiere en el paso del aritmética al álgebra.

Dentro de los estudios acerca del concepto de variable se ha hecho un énfasis especial en la detección de los errores que más freuentemente cometen los estudiantes; en este sentido se han elaborado explicaciones acerca del desarrollo de la abstracción algebraica en los que se ha encontrado que la posibilidad de que los alumnos manejen el concepto de variable como un objeto matemático involucra un proceso largo y penoso que sobrepasa el tiempo que se dedica a este concepto en la escuela y por esa razón son pocos los estudiantes que logran alcanzar una comprensión profunda de los conceptos que constituyen esta disciplina (Sfard y Linchevsky, 1994).

La investigación respecto a la potencialidad que tienen los estudiantes para llegar a manejar las distintas caracterizaciones del concepto de variable es muy escasa, pero existen evidencias preliminares (Arzarello et al, 1993; Bernarz et al, 1991; Kaput, 1996; Reggiani, 1994) que apuntan hacia respuestas positivas y que indican que para lograrlo se requiere de una enseñanza que aborde cada caracterización de una manera más explícita.

En los niveles más avanzados de matemáticas se ha hecho también investigación que refuerza los resultados antes mencionados (Trigueros y Ursini, 1995; Ursini y Trigueros 1997; Gascón, 1993; Gascon, 1998b; Chevallard, 1994). A pesar de ello, en la instrucción universitaria, continúa partiéndose del supuesto de que los estudiantes dominan el concepto de variable a partir de un conjunto de ejercicios y prácticas que la involucran en diferentes formas. Este supuesto señala implícitamente que con práctica suficiente los estudiantes aprenden, por una parte, a manejar el concepto de variable en sus distintas manifestaciones de una forma apropiada y, por otra que, integran estas experiencias para formarse un concepto de variable sólido que las incluya.

En esta parte del trabajo se aborda el problema de las dificultades de los estudiantes frente al concepto de variable a la luz de la investigación que se ha hecho al respecto.

Acerca del concepto de variable

El concepto de variable es de fundamental importancia en el desarrollo y comprensión de cualquier rama de las matemáticas.

Este concepto se introduce por primera vez en la enseñanza dentro del marco de los cursos elementales de álgebra. En ellos los estudiantes se enfrentan a las ideas de generalización y de modelación que han guiado el desarrollo de las matemáticas, dejando atrás el manejo de operaciones específicas mediante el uso de números naturales.

Es también en los cursos de álgebra donde la mayoría de los estudiantes enfrenta grandes dificultades. La generalización de la aritmética al álgebra parece ser un obstáculo difícil de superar (Arzarello et al, 1994; Bell y Malone, 1993; Bernarz y Janvier, 1994; Bolea et al, 1998b; Radford y Grenier, 1996). El aprendizaje del concepto de variable, de su valor como herramienta matemática en la solución y en la modelación de problemas y de su manejo en situaciones diversas está atrás de estas dificultades.

El concepto de variable es complejo. Cuando se revisa su papel dentro del álgebra se encuentra que este concepto se usa con significados diversos en diferentes contextos y que dependiendo de ellos se maneja de distintas maneras. Esta variedad en las formas de empleo hace que el concepto de variable sea difícil de definir y que los estudiantes tengan dificultades para aprenderlo. (Wagner, 1983; Usiskin, 1988, Bolea et al. 1998c; Clement, 1982).

Dentro del marco del álgebra hay ocasiones en las que dada una tarea específica aparecerá una caracterización de la variable también específica, pero en muchas otras situaciones se plantean problemas, que se usan frecuentemente en la didáctica, que requieren que el estudiante trabaje con diferentes formas de usar la variable. Por ejemplo cuando se pide al estudiante que encuentre la ecuación de la recta que pasa por el punto (7,1) con pendiente 5, este problema se puede abordar a partir de la relación $y = mx + b$ en la que el estudiante considera a las variables como números generalizados. Cuando se toman los datos para ponerlos en la fórmula, b es la incógnita que debe determinarse para, por último, llegar a una relación funcional entre las variables x e y. En este ejemplo aparecen los diferentes significados de la variable y el estudiante debe no sólo manejar cada uno de ellos sino, además pasar de uno a otro sin dificultad, aun cuando las distintas caracterizaciones estén revestidas de la misma representación simbólica.

Un usuario competente del álgebra puede asignarle diferentes significados a la variable dependiendo del problema que se le presenta: distingue ecuaciones de tautologías, puede simplificar expresiones y manejar la idea de variación en relaciones funcionales. Para los novatos, estas manifestaciones diversas, que para un experto resultan insignificantes, pueden convertirse en un obstáculo difícil de salvar como lo muestran los estudios de Matz y Kaput con estudiantes que se inician en el álgebra (Matz, 1982; Kaput, 1983; Kaput, 1987).

Parte del problema con las distintas caracterizaciones de la variable proviene de la simbolización. Las variables matemáticas se representan mediante símbolos, sin embargo, se usan los mismos símbolos para diferentes caracterizaciones de la variable y al mismo tiempo, diferentes símbolos se usan para la misma cosa. Esto esconde las diferencias entre las caracterizaciones y hace difícil su comprensión. Existen varios estudios que señalan cómo el simbolismo oscurece los significados para los estudiantes que se inician en el álgebra; entre ellos se pueden mencionar los de Thorndike, Van Engen, Wagner y Matz (Thorndike et al; 1923, Van Engen, 1953; Wagner, 1981; Matz, 1982). En la enseñanza del álgebra es indispensable superar esta situación ya que siempre que se enfrenta un problema es necesario poder pasar flexiblemente de una interpretación a otra del mismo símbolo y para lograrlo es importante disponer de herramientas que permitan diagnosticar las dificultades reales de los estudiantes, para después buscar formas de resolverlas mediante apoyos didácticos.

En términos generales se puede hablar de tres formas distintas en las que la variable puede manifestarse en el álgebra elemental: La variable como número generalizado, es decir como un número no determinado que aparece en generalizaciones y en métodos generales; la variable como incógnita o como un número desconocido y la variable en una relación de variación con otras variables que denominaremos variable en relación funcional.

En 1980 Küchemann hizo un estudio clásico con alumnos de 13 a 15 años en el que, desde una perspectiva piagetiana, clasificó las respuestas de los alumnos con respecto al significado de la variable, organizándolas en términos de su dificultad. El orden introducido por este autor sugiere que para los alumnos lo más simple es trabajar con incógnitas, después con la variable como número generalizado, siendo la más difícil la caracterización de las letras como variable en una relación funcional. Además, él mismo señala que cada uno de estos usos puede enseñarse a distintos niveles de complejidad y que en cada uno de ellos los estudiantes que se inician encuentran dificultades.

Sin embargo, otros trabajos (Lopez, 1996; Kieran, 1984; Kieran y Filloy, 1989; Ursini, 1994) no parecen confirmar esta aseveración. En ellas se sugiere que el uso de la variable como incógnita es tal vez el más favorecido en la enseñanza del álgebra en el aula y que por esta razón puede aparecer como el más simple. Estos estudios sugieren que el uso de la variable como número general es el que realmente introduce a los alumnos al pensamiento algebraico y que la reificación de este concepto, es decir, su encapsulación en formade objeto, juega un papel definitivo en el paso del manejo aritmético al algebraico y facilita la diferenciación y la integración de todos los usos o caracterizaciones de la variable.

Dificultades de los Alumnos frente al Concepto de Variable

La variable como incógnita

En este caso la variable se presenta como un número desconocido. Desde la primaria los alumnos empiezan a trabajar con problemas de este tipo aunque la incógnita no se representa mediante una literal sino mediante otros signos: una raya, un cuadrito, un espacio vacío, etc. Este signo se sustituye por la letra al iniciarse la enseñanza del álgebra.

Si los problemas son muy simples los estudiantes no tiene grandes dificultades. Usando generalmente la estrategia de trabajar hacia atrás o por ensayo y error, la solución se encuentra fácilmente. Sin embargo, cuando la incógnita se encuentra dentro de una ecuación que tiene una estructura más compleja, las dificultades empiezan a surgir, haciéndose más evidentes cuando se trata de ecuaciones que requieren de más de un paso en la solución. La investigación concerniente a este aspecto de la variable ha encontrado que los estudiantes tratan de aplicar las estrategias útiles en ecuaciones más simples en lugar de operar con la incógnita específica (Kieran, 1984; Filloy y Rojano 1984, 1989; Hercovics y Lenchevski, 1991; ,Cortés, 1992; Drouhard, 1992; Rojano 1994).

La presencia del signo de igualdad en las expresiones algebraicas induce a los estudiantes a percibir la variable como incógnita; es decir, los estudiantes muestran grandes dificultades para distinguir el papel de la variable en una tautología, generalmente manejan a la variable como incógnita y se enfrentan a dificultades de interpretación en lo que obtienen después de una manipulación o suponen que la expresión se trata de una ecuación en la que el número de valores que la variable puede tomar depende del grado de la misma.

La presencia de dificultades para conceptualizar la variable como incógnita y simbolizarla se reconfirma en el caso de los ítems que presentan problemas verbales. Cuando el problema es muy simple, los estudiantes tienden a resolver el problema sin plantear una ecuación. Si bien identifican la incógnita del problema, esto es, conceptualizan la incógnita, no la simbolizan y determinan su valor por procedimientos aritméticos. Cuando la complejidad del problema no les permite usar un acercamiento aritmético dan una respuesta arbitraria poniendo en evidencia su incapacidad para plantear la ecuación y también para conceptualizar la incógnita.

Resumiendo puede decirse que los estudiantes manejan la incógnita a un nivel muy elemental, en el que apenas la reconocen como número desconocido y su capacidad para identificarla dentro de un problema más general y de simbolizarla adecuadamente son muy limitadas.

Aun en las ecuaciones se nota ente los estudiantes una tendencia a evitar la manipulación y a determinar el valor de la variable por simple inspección. Tal parece que cuando la instrucción en un ítem pide resolver la ecuación, los estudiantes reaccionan automáticamente buscando un valor por inspección, evitando la manipulación.

La enseñanza tradicional no parece ayudar a los estudiantes a superar las dificultades que se han señalado e incluso tiende a reforzar estrategias equivocadas o estrategias que no los conducen a darle significado a la variable. Las investigaciones acerca de experimentos de enseñanza diseñados para superar las dificultades específicas encontradas muestran distintas posibilidades de superarlos, aunque señalan que se requiere de un esfuerzo considerable de todos los involucrados en el proceso de enseñanza aprendizaje.

La variable como número general

La variable es una herramienta que se usa en matemáticas para expresar generalización . Cuando se quiere expresar un patrón, una regularidad o un método general en matemáticas se usan variables para indicar los números generales involucrados.

La investigación acerca de la forma en que los estudiantes trabajan con la variable como número general ha mostrado que la mayoría tiene dificultades para entender esta caracterización. Se encontró que al enfrentrarse a un número general representado por una literal los estudiantes producen una gran variedad de interpretaciones del símbolo y que éstas varían además con el tipo de problema en el que se está trabajando. La conducta más frecuente de los estudiantes consiste en ignorar la variable, o en asignarle un valor específico, aunque en algunas ocasiones la interpretan como un objeto determinado . En términos generales se puede decir que los estudiantes son capaces de manipular la variable como numero general cuando esta aparece en expresiones abiertas muy simples, pero que cuando se aumenta ligeramente la complejidad las dificultades en su manejo se ponen de manifiesto. La capacidad para construir una expresión que involucra un número general aumenta ligeramente cuando el símbolo tiene un referente claro y la expresión que se obtiene representa un resultado, tal como el perímetro de una figura. A pesar de que los estudiantes pueden reconocer patrones generales cuando se pregunta respecto de los pasos correspondientes a un número específico, son incapaces de expresar la regla general que gobierna al patrón y que si bien son capaces de reconocer al símbolo como la representación de algo indeterminado en casos simples, muestran grandes dificultades cuando se requiere de un nivel de abstracción un poco más elevado (Küchemann, 1980; Booth, 1984; Ávila, et al, 1990; Bolea et al, 1998a; Chevallard, 1989; Gascón, 1993).

Todos los estudios coinciden en que los estudiantes se desorientan cuando tienen que interpretar símbolos que representan números generales y en que su tendencia es a especificarlos de una manera arbitraria.

Es importante señalar que algunos estudios, particularmente aquellos realizados con ambientes computacionales, muestran la potencialidad de los estudiantes que se inician en el álgebra para manejar la variable como número general y su capacidad de transferir esta capacidad al contexto de la clase de álgebra (Noss, 1985; Sutherland, 1987; Ursini, 1994, Ursini y Trigueros, 1997, Elias Calles, 1996). Estos resultados sugieren que un acercamiento distinto al tradicional puede reducir las dificultades encontradas y ayudarles a trabajar con la idea de números generales y su representación simbólica en una forma apropiada. También se encuentra en ellos una sugerencia implícita acerca de que la instrucción tradicional no ofrece suficientes oportunidades para que el alumno construya la idea de número general y para que desarrolle un significado para el símbolo que se usa para representarlo.

La Variable en una Relación Funcional

Las variables se pueden usar para expresar una relación entre dos cantidades que cambian. Cuando se usa así, las características principales de la variable son su variación en un rango de valores y el hecho de que el cambio en una de las variables produce un cambio en el valor de la otra. Esta característica se enfatiza cuando la relación funcional se concibe como la expresión del cambio.

En el curriculum actual las funciones juegan un papel importante, pero se definen mediante dos conjuntos y una regla de correspondencia que asigna a un elemento de un conjunto exactamente un elemento del otro conjunto. En esta definición de función, la idea de cambio y la relación entre el movimiento de las variables está ausente.

Además, cuando se usa esta definición de función, los estudiantes encuentran obstáculos en las ideas de dominio y rango de variación, independientemente de que la función esté representada en forma algebraica o gráfica. Los alumnos tienen también dificultades para entender las funciones constantes y más aún las discontinuas. (Markovits et al., 1986; Sfard, 1989; Dubinsky, 1991; Dubinsky y Harel, 1992).

El desarrollo de acepciones complementarias de las relaciones funcionales requiere del desarrollo de formas versátiles de pensar, es decir, de la capacidad de moverse libremente entre modos de pensamiento secuenciales y globales (Thomas y Tall, 1991), o dicho de otra manera, de tener la posibilidad de considerar por una parte a las variables en una correspondencia estática y ser capaz de calcular el valor de una de ellas en correspondencia con el valor asignado a la otra, o, por otra parte, de pensar en las variables como entidades que se mueven una en relación a la otra dentro de un rango de valores y de manejar así la idea de cambio.

La idea de cambio es difícil para los alumnos. La dependencia de los cambios de una de las variables de los de la otra es algo que los alumnos no ven con facilidad. Ellos perciben los procesos dinámicos como estáticos (Bernarz y Dufour-Janvier, 1991; Trigueros y Cantoral, 1992) y cuando se describen procesos dinámicos los alumnos se fijan solamente en algunas características esenciales que se pueden representar puntualmente dejando de lado la concepción de la relación funcional en forma global. Cuando el cambio se representa en forma gráfica o mediante otros códigos usuales en las matemáticas, también se encuentran dificultades y nuevamente éstas están asociadas a la tendencia de los estudiantes a ver a las variables como algo estático (Kieran, 1992; English y Sharry, 1996).

Los resultados encontrados ponen de manifiesto que los estudiantes manejan de manera adecuada la noción de correspondencia entre cantidades, independientemente de la representación que se use para expresar una relación (tabla, gráfica, expresión analítica). La noción de variación conjunta, sin embargo, les resulta totalmente ajena. Se observa también que si bien los estudiantes no tienen dificultades para determinar el valor específico de una variable cuando se conoce el valor de la otra, son incapaces de determinar los intervalos de variación correspondientes.

El manejo que los estudiantes hacen de las tablas y de las gráficas sugiere que recurren a métodos intuitivos para interpretar la variación. Los estudiantes tienen dificultades con el manejo de la noción de continuo numérico y esto les impide establecer intervalos en los que el comportamiento de la función tiene cierta propiedad, por ejemplo, los intervalos en que la función crece o decrece.

La simbolización de la relación funcional entre variables presenta asimismo serias dificultades, aunque estas dificultades disminuyen cuando se explican las variables que deben ser relacionadas y se asigna en el ítem un símbolo específico para designarlas.

La manipulación de las relaciones funcionales, por ejemplo efectuar sustituciones, transposiciones, agrupación de términos semejantes, etc. no representa dificultad para los estudiantes cuando se presentan en forma aislada, pero cuando las tienen que llevar a cabo en forma sucesiva, las dificultades aumentan notablemente.

En términos generales puede decirse que en problemas que involucran las relaciones entre variables las respuestas comunes de los estudiantes muestran inconsistencias y generalizaciones inadecuadas aun en el caso de problemas muy simples. La mayor parte de los estudiantes no parece concebir a la relación como un proceso de transformación, ni como un proceso dinámico de variación conjunta. Su concepción parece limitarse a la idea estática de correspondencia uno a uno.

Todos los estudios parecen mostrar que las dificultades que los estudiantes tienen para conceptualizar las variables en relación funcional son aún mayores que las que tienen para manejar la variable como incógnita y como número general.

La Variable en los Libros de Texto

Con el objeto de tener una idea más clara del manejo del concepto de variable en el ámbito escolar se revisaron algunos de los libros más frecuentemente usados en los niveles primario y secundario en México.

En cuanto al concepto de variable se encontró lo siguiente:

Una revisión de los textos más comunes en México muestra que en primaria los estudiantes encuentran por primera vez a las variables en una relación funcional, representada mediante símbolos; sin embargo, el énfasis sobre ellas no es con respecto a la relación funcional en sí, ni respecto al papel de las variables sino simplemente como formas abreviadas de recordar algo.

En el primer año de secundaria se presentan a los estudiantes diversas expresiones aritméticas generalizadas que involucran el uso de los símbolos para representar variables como números generales. Sin embargo, no se va más lejos de una presentación informal y se nota una ausencia de actividades conducentes a la construcción de la idea de variable como número general.

También en el primer año de secundaria aparece brevemente de manera formal el concepto de función. En esta presentación se hace énfasis en la idea de correspondencia, pero no en la idea de cambio o de variación de dos variables en forma conjunta .

El uso de literales para representar incógnitas empieza de una manera formal en los libros del segundo año de secundaria. Antes, como ya se ha mencionado, aparecen este tipo de problemas pero en ellos sólo se usa ocasionalmente la literal como símbolo.

No se encontró en los libros revisados ninguna actividad enfocada a la construcción de los distintos usos de la variable. Las actividades se centran primordialmente en la representación simbólica de la variable y en su uso en la solución de ejercicios en los que se le concibe como incógnita y en algunos problemas de índole más general.

Una primera observación que es importante hacer respecto a esta revisión es que los libros revisados difieren muy poco en la forma como abordan los conceptos del álgebra. Los ejemplos utilizados y los ejercicios propuestos son también muy similares.

Algunos resultados obtenidos de la observación en el aula

En algunas investigaciones relativas al concepto de variable se incluye como parte fundamental del diseño experimental la observación de la manera en la que el álgebra se enseña en el salón de clase. Esta información complementa la presentada anteriormente en la que los restados provienen básicamente de la aplicación de distinto tipo de instrumentos escritos y de entrevistas con los estudiantes.

Los resultados que se obtienen de la observación en clase muestran que durante sus explicaciones los profesores de matemáticas hacen un uso amplio de los distintos aspectos de la variable dando por sentado que son comprendidos por los alumnos (López, 1996). Los profesores pasan de un aspecto a otro en un mismo problema sin dar ninguna explicación que apoye a los alumnos en su comprensión de los diferentes papeles que la variable va asumiendo a lo largo de la solución del problema o cuando se pasa de un problema a otro.

Se constata, por otra parte que el tiempo que los profesores de matemáticas destinan a la mecanización de los procedimientos en la resolución de ejercicios y problemas es siempre mucho mayor que el que destinan a ejercicios de aplicación o de conceptualización. Esto propicia que los estudiantes pierdan la relación con la parte conceptual y se concentren en la memorización de procedimientos y algoritmos.

Por otra parte se observa que cuando los alumnos cometen un error durante la clase, este error es muy frecuentemente ignorado por el profesor; en los casos en que se retoma, el mayor énfasis del profesor se concentra en la repetición del procedimiento o algoritmo pero rara vez se discuten los aspectos conceptuales involucrados. Este tipo de actitudes de los profesores tiende a reforzar la consideración, muy común entre los alumnos, de que el álgebra consiste en una serie de reglas de manipulación y algoritmos que deben memorizar y aplicar de manera mecánica para resolver los ejercicios que se les proponen, y se deja de lado la oportunidad de ayudar a los alumnos a desarrollar una comprensión significativa de los conceptos algebraicos, en particular del de variable que, en principio deben quedar claros al finalizar el curso.

Por otra parte, los profesores que trabajan con estudiantes que inician sus estudios universitarios encuentran que a pesar de haber cursado al menos cinco años de matemáticas en la escuela secundaria y preparatoria, en los cuales el énfasis mayoritario está puesto en el manejo del álgebra, los estudiantes continúan mostrando dificultades con el manejo de expresiones algebraicas. Estas dificultades repercuten negativamente en sus estudios del Cálculo

Diferencial e Integral, de otras materias de matemáticas avanzadas y de materias de los distintos campos profesionales que usan la matemática como herramienta.

Algunas consideraciones adicionales

Como hemos visto anteriormente, el concepto de variable es complejo y su aprendizaje presenta muchas dificultades a los estudiantes. Este problema se agudiza por el énfasis exagerado que se pone en el aspecto de variable como incógnita tanto en los libros de texto como en la instrucción.

Todos los resultados antes mencionados dan muestras de que el aprendizaje del concepto de variable logrado por los estudiantes a su paso por el sistema escolar es poco significativo. Si bien se encuentra que los alumnos son capaces de reconocer el papel que juega la variable en expresiones y problemas muy simples, un ligero aumento en la complejidad de los mismos provoca generalizaciones inadecuadas y la búsqueda de soluciones memorizadas o por inspección que no son acordes a los objetivos planteados por la instrucción.

Las estrategias de los estudiantes están dominadas por procedimientos que no han sido interiorizados, esto los deja anclados a un nivel de acción en los términos del marco teórico presentado en la sección anterior, que se manifiesta por ejemplo en la necesidad de hacer explícitos los pasos que siguen en el proceso mental de solución y usarlos como soporte para continuar, sin ser capaces de analizarlos y detectar posibles errores. Muchas de sus acciones parecen estar provocadas por estímulos que perciben como externos y que los inducen a responder de cierta manera específica, sugiriendo nuevamente un anclaje a un nivel de acción.

La atención de los estudiantes se centra en las características superficiales de las expresiones y en los procedimientos aritméticos involucrados en su solución. La exposición a los cursos y la lectura de los textos no desarrolla en forma significativa la posibilidad de generalizar procedimientos y patrones.

Los resultados provenientes del análisis de lo que ocurre en el aula sugieren que el tipo de problemas que se detectan en estas investigaciones están fuertemente ligados a la forma en la que suele enseñarse el álgebra. Esta situación, aunada a la posibilidad de que el concepto de variable se desarrolle lentamente, como sugieren los estudios, podría apuntar hacia la necesidad de replantear la forma en la que este concepto se aborda.

Dada la complejidad del concepto de variable es tal vez demasiado optimista esperar que la práctica y el acercamiento a los textos permita a los alumnos alcanzar un concepto en el que estén integradas las diversas caracterizaciones de la variable.

Es posible que muchas de las dificultades antes mencionadas provengan de la falta de construcción, por parte de los estudiantes, del concepto de variable. Esta construcción debe incluir todos los aspectos vistos anteriormente y la posibilidad de pasar de una concepción a otra del concepto de variable con flexibilidad y acoplándose a las exigencias de los problemas que se intentan resolver.

Hemos visto hasta aquí, cómo la idea de variable puede considerarse no como un concepto monolítico, sino como uno con múltiples facetas dentro de las que destacan las concepciones como incógnita, como número general y como parte de una relación funcional.

La comprensión integral del concepto de variable requiere del entendimiento y el manejo adecuados de esas tres facetas, además de la posibilidad de pasar de una a otra de manera dinámica y reversible.

Una primera forma de acercarse al estudio de la comprensión del concepto de variable por parte de los estudiantes requiere de un análisis del grado en que los alumnos han comprendido los distintos aspectos que lo componen.

Tanto los textos de matemáticas como la instrucción en clase no hacen suficiente énfasis en los diferentes aspectos del concepto de variable. Lo más usual es que se enfatice la idea de variable como incógnita y se preste poca o ninguna atención a las otras manifestaciones con la idea subyacente de que el alumno por sí solo puede construir los otros significados cuando se le presentan.

Cuando los estudiantes llegan a los cursos avanzados de matemáticas, sus problemas con el uso del álgebra subsisten. Es posible que esto se deba a que la generalización e integración de las distintas facetas del concepto de variable no se da en la forma esperada.

Para ahondar en esta hipótesis es conveniente diseñar y experimentar con un mayor número de instrumentos de diagnóstico y evaluación del conocimiento del álgebra que aborden directamente las distintas caracterizaciones en las que este concepto se puede presentar y que permitan posteriormente incidir en la forma en la que éstos se enseñan.

El diseño de instrumentos basados en marcos teóricos relevantes, como el que se presentó en la sección anterior, permiten un análisis a mayor profundidad de la comprensión de los alumnos de los conceptos matemáticos. En la siguiente parte del trabajo se pretende analizar, a través del uso del marco teórico constructivista antes mencionado al concepto de variable. Es decir se presenta una descomposición genética que permitirá el diseño de instrumentos y el análisis de los resultados obtenidos de su aplicación. En particular se utilizará el análisis de los resultados de la aplicación de un instrumento en particular, que permite distinguir los aspectos concernientes a la diferenciación entre las distintas caracterizaciones del concepto de variable en el álgebra elemental.

III.5 Descomposición genética del concepto de variable y su uso en el diseño de ítems

El aprendizaje significativo de los conceptos del álgebra es, como se ha mencionado anteriormente un proceso largo y difícil de completar para muchos de los estudiantes que pasan por las escuelas secundarias. Los errores más comunes que los alumnos cometen se han tipificado, pero aún hace falta un marco teórico que de sentido al análisis de las respuestas de los estudiantes y que sirva como sustento para el diseño de ítems útiles en el diagnostico y en la evaluación de los conocimientos de los sujetos acerca del concepto de variable en el álgebra elemental. El marco teórico desarrollado en la sección III.3 puede utilizarse en particular con este fin.

El concepto de variable como se ha visto anteriormente, es multifacético e involucra, como un todo, distintos aspectos. La mayor parte de las investigaciones tienden a centrar su atención sólo en alguna de las formas en las que este concepto se manifiesta en los problemas algebraicos. En el presente trabajo consideramos tres aspectos del concepto de variable, destacados en los trabajos realizados destacados por otros investigadores (Usiskin, 1988; Ursini, 1994; Ursini, 1996), que pensamos son los más relevantes para un manejo eficaz y una comprensión significativa de los problemas algebraicos elementales: el uso de la *variable como incógnita*, el uso de la *variable como número general* y el uso de la variable en una relación funcional.

La comprensión del concepto de variable implica, desde nuestra perspectiva, la posibilidad de superar la simple realización de cálculos y operaciones con letras o con símbolos, para alcanzar una comprensión de las razones por las que funcionan estos procedimientos, la capacidad de prever hacia dónde conducen y la posibilidad de establecer relaciones entre las distintas formas en las que la variable se manifiesta en el contexto del álgebra elemental.

Construir el concepto de variable como un objeto matemático en una forma integral implica la capacidad de emplear alguna de las formas en las que este concepto se manifiesta en una situación específica, así como pasar de una de estas manifestaciones a otra de manera flexible e integrar los distintos aspectos como componentes de un mismo ente matemático.

El trabajo adecuado con cada uno de los aspectos en los que la variable puede presentarse implica la posibilidad de: interpretar, en un problema dado, el significado de la variable, es decir, reconocer el papel que la variable juega en esa situación; manipularla y operar sobre ella; así como utilizarla con el fin de representar situaciones problemáticas mediante símbolos a los que se atribuya el significado correspondiente.

Es de esperarse que después de varios años de contacto con el álgebra, los estudiantes universitarios sean capaces de manejar los tres aspectos del concepto de variable simultáneamente y que la comprensión de cada uno de ellos esté igualmente desarrollada. Las matemáticas avanzadas que el estudiante aborda a este nivel requieren como un prerrequisito la capacidad de manejar el concepto de variable algebraica como un objeto, ya que este concepto será generalizado para incluir dentro de él nuevas facetas e interpretaciones.

La comprensión de la forma en la que los estudiantes entienden y manejan el concepto de variable puede lograrse tomando como punto de partida un análisis cuidadoso de lo que significa entender el concepto, aislando sus componentes y describiendo explícitamente las relaciones entre ellas. La descomposición genética de un concepto (Asiala et al 1996) puede ser una herramienta útil para investigar la forma en la que los estudiantes lo entienden. En el caso del estudio que aquí nos concierne, se presenta una descomposición del concepto de variable que se obtiene, por una parte, de un análisis cuidadoso de este concepto y, por otra, de la revisión de las dificultades que tienen los estudiantes al enfrentarse a él y que han sido reportadas en la literatura. Este análisis se enfoca hacia los aspectos que parecen ser más relevantes desde el punto de vista de un experto y , puesto que el trabajo se realiza con estudiantes universitarios que han tenido contacto con el álgebra durante varios años, este interés se centra en destacar los aspectos que se consideran importantes para que los estudiantes manejen con soltura y comprendan el concepto de variable como un objeto matemático integrado.

En un primer acercamiento, la descomposición del concepto de variable se basó en la forma en la que los investigadores entendíamos el concepto y en su experiencia como maestros y como aprendices en cuanto a las construcciones mentales que consideraban necesarias para lograr su apropiación (Quintero et al, 1994; 1995; Trigueros et al, 1996). Esta descomposición se ha refinado a partir de la información que se obtuvo de un primer análisis de las respuestas de los estudiantes a los instrumentos de investigación.

A continuación se presenta una nueva descomposición del concepto de variable que contiene los aspectos que pueden considerarse básicos para la comprensión de este concepto:

Variable como incógnita

La conceptualización de la *variable como incógnita* implica:

- reconocer e identificar en una situación problemática la existencia de algo desconocido que se puede determinar
- reconocer el símbolo que aparece en una ecuación como un ente que puede tomar valores específicos
- sustituir el o los valores de la variable que hacen que la ecuación sea verdadera
- determinar la incógnita que aparece en ecuaciones o problemas llevando a cabo las manipulaciones y operaciones algebraicas o aritméticas necesarias
- identificar la incógnita en una situación específica y simbolizarla adecuadamente a través de una ecuación.

Habrá quien considere inadecuada la terminología "*variable como incógnita*" por el hecho de que una incógnita no es variable dado que representa un valor fijo. Sin embargo, puede considerarse que la primera percepción de las literales al trabajar con un problema algebraico es, o tendría que ser, la de símbolos que representan cualquier valor y que es en un segundo momento cuando se define su papel específico dentro de la expresión o problema en el que aparecen. Así, por ejemplo, frente a una ecuación, se toma consciencia de que la variable representa valores específicos sólo después de llevar a cabo, de hecho o mentalmente, las manipulaciones necesarias que permitan darse cuenta de que se trata efectivamente de una ecuación y no, por ejemplo, de una tautología. Por esta razón nos parece que el uso de la terminología "*variable como incógnita*" es adecuada.

Variable como número general

En el caso de la *variable como número general* su conceptualización requiere la capacidad de:

- reconocer patrones y reglas en secuencias numéricas y en familias de problemas
- reconocer el símbolo como una representación de un objeto indeterminado
- desarrollar la idea de método general, distinguiendo los elementos variables de los invariantes en situaciones problemáticas similares, hasta llegar a la simbolización de un método general y del objeto general sobre el cual actúa
- simplificar o desarrollar expresiones algebraicas

Variables en relación funcional

La conceptualización de las *variables en relación funcional* requiere:

- reconocer la correspondencia entre cantidades en sus diferentes representaciones: tabla, gráfica, problema verbal o expresión analítica.
- determinar los valores de la variable dependiente cuando se conocen los de la variable independiente
- determinar los valores de la variable independiente cuando se conocen los de la variable dependiente
- reconocer la variación conjunta de las variables que intervienen en una relación en cualquiera de sus formas de representación
- determinar los intervalos de variación de una de las variables cuando se conocen los de la otra

- expresar una relación funcional de manera tabular, gráfica y/o analítica, a partir de los datos de un problema

Para cada uno de estos aspectos de la variable, la descomposición pone de manifiesto distintos niveles de abstracción a los que este concepto se puede manejar. El manejo de la variable como un objeto matemático a este nivel, implica, como ya se ha mencionado, la posibilidad de integrar las diferentes manifestaciones de la variable en un solo concepto y de pasar flexiblemente de una a otra en cualquier situación.

Con el fin de poder aplicar este análisis teórico del concepto de variable al diseño de ítems y de tests que pretendan evaluar conocimientos respecto a este concepto, es necesario reconocer dentro de esta descripción cuál o cuáles comportamientos o respuestas de los estudiantes representan una concepción a nivel de acción del concepto de variable, cuáles representan una concepción a nivel de proceso y cuáles representan una concepción a nivel de objeto.

Conceptualización de la variable a nivel - acción

En el caso de la *variable como incógnita* un estudiante que se encuentra a un *nivel - acción* en su comprensión del concepto de variable puede:

- Reconocer en una situación problemática la existencia de algo desconocido que se puede determinar.
- Identificar la incógnita en una situación determinada.
- Reconocer la literal que aparece en una ecuación como un número específico que se puede calcular

Los indicadores de que un estudiante se encuentra en este *nivel - acción* en su comprensión de la *variable como incógnita* son los siguientes:

- Dificultades para discriminar cuando una variable en un problema representa un número general o una incógnita.
- Respuestas inducidas directamente por la presencia de signos externos o por la aparición de palabras clave que actúan como estímulos externos que conducen a dar respuestas memorizadas sin que haya un análisis de la situación específica en la que la ecuación aparece.
- Solución de los problemas por simple inspección o mediante el uso de procedimientos aritméticos.
- Tendencia a evitar la manipulación de las ecuaciones cuando se pide encontrar el valor de la incógnita.
- Interpretación de las diferentes apariciones del símbolo como diferentes números.
- Imposibilidad de interpretar el símbolo en una ecuación cuando aparece más de una vez.
- Dificultades para desligar al símbolo del referente específico en una situación problemática dada.
- Utilización del signo de igualdad como conexión entre los distintos pasos que se siguen en la solución de un problema.
- Necesidad de explicitar paso a paso todo el proceso de solución de una ecuación dada.

Los indicadores de que un estudiante se encuentra en este *nivel - acción* en su comprensión de la *variable como número general* son los siguientes:

- Dificultades para discriminar cuando una variable en un problema representa un número general o una incógnita.

- Respuestas inducidas directamente por la presencia de signos externos o por la aparición de palabras clave que actúan como estímulos externos que conducen a dar respuestas memorizadas sin que haya un análisis de la situación específica en la que la expresión aparece.
- Dificultad para reconocer el símbolo en una expresión como un objeto indeterminado.
- Confusión entre una ecuación y una tautología.
- Necesidad de igualar una expresión algebraica a un número para manejarla como una ecuación.
- Dificultades para simplificar o desarrollar expresiones abiertas, cuando éstas no son my simples.
- Necesidad de efectuar paso a paso la simplificación o el desarrollo de expresiones.
- Posibilidad de construir una expresión que involucra un número general únicamente cuando el símbolo tiene un referente claro y la expresión que se obtiene representa algún resultado específico como un perímetro o un área.
- Dificultad para generalizar patrones en términos de variables, aunque se tenga la capacidad de hacerlo únicamente en términos numéricos.

Los indicadores de que un estudiante se encuentra en este *nivel - acción* en su comprensión de la *variable en la relación funcional* son los siguientes:

- Reconocimiento de la correspondencia entre cantidades en sus diferentes representaciones: tabla, gráfica, problema verbal o expresión analítica.
- Dificultad para reconocer la correspondencia entre variables.
- Posibilidad de determinar los valores de la variable dependiente cuando se tiene la expresión analítica y los valores de la variable independiente son conocidos.

- Respuestas inducidas directamente por la presencia de signos externos o por la aparición de palabras clave que actúan como estímulos externos que conducen a dar respuestas memorizadas sin que haya un análisis de la situación específica en la que la relación aparece.
- Imposibilidad de determinar una relación de variación conjunta.
- Dificultades para determinar intervalos de variación.
- Dificultades en el manejo de variables continuas.
- Conceptualización de cualquier tipo de relación como una relación discreta.
- Recurrencia a métodos intuitivos para interpretar la variación.
- Dificultad para reconocer intervalos de crecimiento y decrecimiento de la función.
- Uso de métodos aritméticos para encontrar la relación funcional.
- Generalización de la propiedad de linearidad para cualquier tipo de relación.
- Interpretación de problemas verbales en términos de relaciones entre cantidades específicas.
- Necesidad de una fórmula explícita para interpretar el comportamiento de una relación funcional representada de otra manera.
- Uso del símbolo de igualdad para sintetizar la información proporcionada en un problema y para encadenar los pasos seguidos en un razonamiento.
- Interpretación de la relación funcional limitada a aquélla de correspondencia uno a uno.
- Posibilidad de encontrar intervalos de variación analizando únicamente el comportamiento de la función en los valores extremos del intervalo y generalizando el comportamiento de la función al interior del intervalo.

Además de estos indicadores a nivel de cada una de las caracterizaciones de la variable, en un problema más complejo, los estudiantes que se encuentran a este nivel tienen dificultades para resolver el problema a menos que cuenten con una guía específica que pueda ser memorizada y que les permita aplicar con seguridad un algoritmo.

Los estudiantes se muestran dificultades en la solución de una situación específica por la dificultad de reconocer en un problema complejo las distintas caracterizaciones que la variable asume en el transcurso de la solución.

Conceptualización de la variable a nivel de proceso

En el caso de la *variable como incógnita* un estudiante que se encuentra a un *nivel - proceso* en su comprensión del concepto de variable puede:

- Reconocer las diferentes apariciones de la variable en una ecuación como representaciones de un mismo número.
- Sustituir los valores de la variable que hacen que la ecuación sea verdadera.
- Determinar el valor de la incógnita que aparece en ecuaciones o problemas llevando a cabo las manipulaciones y las operaciones algebraicas o aritméticas necesarias.
- Posibilidad de reconocer entre ecuaciones y expresiones algebraicas abiertas y tautologías.
- Posibilidad de reflexionar acerca de la pertinencia de las manipulaciones en una ecuación.
- Posibilidad de llevar a cabo el procedimiento de solución de la ecuación sin necesidad de explicitar todos los pasos.
- Interpretar correctamente el signo de igualdad.
- Posibilidad de identificar la incógnita en un problema dado, aun cuando se manifiesten todavía dificultades para simbolizar la ecuación que resuelve el problema.

En el caso de la *variable como número general* un estudiante que se encuentra a un *nivel - proceso* en su comprensión del concepto de variable puede:

- Reconocer patrones y reglas en secuencias numéricas simples.
- Reconocer patrones en familias de problemas simples.
- Simplificar expresiones algebraicas sencillas.
- Desarrollar expresiones algebraicas sencillas.
- Manipular expresiones para llevar a cabo un análisis que permita distinguir entre el papel que juega el símbolo en una expresión que puede ser, en apariencia, una ecuación o una tautología.
- Posibilidad de interpretación de las expresiones sin dejarse llevar por la presencia de palabras claves o signos externos.
- Posibilidad de manipular expresiones abiertas más complejas para simplificarlas o para desarrollarlas.
- Posibilidad de simbolizar expresiones simples sin necesidad de explicitar paso a paso el proceso de pensamiento seguido.
- Capacidad de construir expresiones que involucran objetos generales sin necesidad de tener un referente explícito.
- Posibilidad de generalizar patrones simples en términos algebraicos, utilizando símbolos.

En el caso de la *variable en una relación funcional* un estudiante que se encuentra en un *nivel - proceso* en su comprensión del concepto de variable puede:

- Determinar los valores de la variable independiente cuando se conocen los de la dependiente.
- Determinar intervalos de variación de una de las variables cuando se conocen los de la otra.
- [illegible]

- Expresar relaciones funcionales simples en forma de tabla, gráfica o de expresión analítica a partir de los datos de un problema.
- Pasar de una expresión analítica a una tabla o a una gráfica que le correspondan.
- Analizar el comportamiento de funciones simples expresadas en términos de tablas o gráficas sin necesidad de contar con una fórmula analítica que relacione las variables de manera explícita.
- Interpretar la noción de variación conjunta.
- Posibilidad de conceptualizar el continuo numérico en problemas de variación.
- Posibilidad de inferir el comportamiento de funciones simples sin necesidad de calcular puntos específicos.
- Posibilidad de interpolar valores dada una función en representación tabular o gráfica.
- Posibilidad de encontrar intervalos de variación analizando el comportamiento de la función al interior del intervalo.
- Posibilidad de encontrar simbólicamente relaciones funcionales simples a partir de problemas verbales, datos en tablas o gráficas.
- Posibilidad de aplicación de razonamiento proporcional para simbolizar y trabajar con funciones lineales.
- Posibilidad de reconocer que no todas las funciones representadas en forma de tabla tienen un comportamiento lineal.
- Posibilidad de ejecución de varias acciones tales como sustitución, transposición y agrupación de términos semejantes para encontrar los valores de la variable independiente o dependiente en el caso de una relación funcional expresada en forma analítica.
- Posibilidad de concebir la relación funcional como un proceso dinámico de asociación entre variables continuas o discretas según sea el caso, sin limitarse a la idea estática de correspondencia uno a uno.

Aun cuando los estudiantes que se encuentran en un *nivel - proceso* en su conceptualización del concepto de variable son capaces de manejar este concepto con mayor eficacia en la solución de problemas y ejercicios, así como de diferenciar las diferentes caracterizaciones de este concepto, sus posibilidades de integración de las diferentes caracterizaciones de la variable son limitadas y ello hace que la posibilidad de generalizar este concepto a dominios más abstractos de las matemáticas esté restringida a casos muy particulares o se haga de una manera incorrecta, adjudicando propiedades a la variable que puede poseer en un contexto, pero no en otro.

Conceptualización de la variable a nivel - objeto

En el caso de la *variable como incógnita* un estudiante que se encuentra en un *nivel - objeto* en su comprensión del concepto de variable es capaz de realizar todas las operaciones antes mencionadas y además es capaz de:

- Identificar la incógnita en una situación específica, independientemente de la complejidad del problema.
- Manipular las ecuaciones a fin de obtener el valor de la incógnita comprendiendo el proceso a seguir y con la posibilidad de reflexionar críticamente para encontrar errores en algún momento del proceso de solución de ecuaciones.
- Simbolizar la incógnita adecuadamente en problemas o situaciones problemáticas utilizando un símbolo pertinente y reconociendo el papel que ese símbolo juega en la ecuación.
- Posibilidad de plantear ecuaciones en situaciones variadas, en las que a la incógnita se le asigna un significado específico.
- Posibilidad de elección del método más adecuado para resolver una ecuación dada, sin importar el nivel de complejidad de la misma.

En el caso de la *variable como número general* un estudiante que se encuentra en un *nivel – objeto* en su comprensión del concepto de variable es capaz de:

- Desarrollar la idea de método general en una situación dada distinguiendo aquéllos elementos que son variables de los que no lo son en situaciones problemáticas similares.
- Analizar el comportamiento de expresiones algebraicas en las que los parámetros pueden cambiar, prediciendo adecuadamente el comportamiento de la situación a la que la expresión se refiere al variar los parámetros.
- Simbolizar métodos generales utilizando símbolos que representan la generalización y reconociendo el objeto general sobre el cual actúa.
- Posibilidad de interpretación de la variable más adecuada para representar un objeto general en cualquier situación dada.

En el caso de la *variable en la relación funcional* un estudiante que se encuentra en un *nivel - objeto* en su comprensión del concepto de variable es capaz de:

- Reconocer la variación conjunta de las variables que intervienen en una relación en cualquiera de sus formas de representación.
- Expresar una relación funcional a partir de los datos de un problema en cualquiera de sus representaciones: tabular, gráfica o analítica.
- Pasar de una representación de la relación funcional a otra.
- Reconocer el comportamiento de la relación en diferentes intervalos de variación de la variable independiente.
- Reconocer el comportamiento de la relación para diferentes intervalos de la variable dependiente.
- Reconocer la relación entre variables en diferentes problemas que involucren cambio.
- Identificar los problemas en los que la variación es continua de aquellos en los que no lo es.

- Generalizar el concepto de relación funcional a contextos nuevos.
- Representar analíticamente relaciones entre variables independientemente de su complejidad.
- Posibilidad de elección de la representación más adecuada para representar una relación entre variables en una situación particular.
- Posibilidad de interpretar el significado de una relación funcional en términos de la situación específica que representa.
- Capacidad de distinguir las relaciones lineales de las que no lo son ananlizando e interpretando el comportamiento de las variables relacionadas.
- Concebir la relación entre variables como un proceso de transformación en el que se reconocen los invariantes.
- Concebir la relación entre variables como un proceso dinámico de variación conjunta.
- Posibilidad de interpretar la relación funcional tanto de manera global, en todo el dominio de variación, cuanto de manera local, es decir en intervalos específicos de variación.

Para cada uno de estos aspectos de la variable, la descomposición pone de manifiesto distintos niveles de abstracción a los que este concepto se puede manejar. El manejo de la variable como objeto matemático a este nivel implica la posibilidad de integrar las diferentes manifestaciones de la variable en un solo concepto y de pasar flexiblemente de una a otra en cualquier situación.

Diseño de ítems para medir el concepto de variable

El objetivo de la descomposición genética consiste en desarrollar herramientas que permita, por una parte, diseñar situaciones que el alumnos enfrentará durante la investigación o la docencia y, por otra, analizar sus respuestas. A través de las respuestas de los estudiantes a los instrumentos de investigación se intenta lograr una comprensión de los significados que los estudiantes han construido para los conceptos, en este caso del concepto de variable a su paso por el sistema educativo, lo que permitirá posteriormente, formular propuestas que puedan incidir en las formas de instrucción.

Si bien es difícil diseñar situaciones en las que la variable juegue un papel específico único, sí es posible diseñar situaciones en las que se destaque uno de los aspectos de la variable de manera preponderante. Es en este tipo de situaciones en las que las respuestas de los alumnos pueden proporcionar la información necesaria para interpretar la forma en la que conciben el papel de la variable y su forma de trabajar con ella. Es claro que en la mayoría de las situaciones de solución de problemas, los tres aspectos de la variable que se han utilizado como base para su descomposición aparecen interrelacionados y que la posibilidad de interpretar la relevancia de cada uno de ellos en los distintos momentos de la solución de problemas constituye un elemento esencial para un manejo competente de la variable en el álgebra. El diseño de los instrumentos de destinados al análisis de la forma en la que los estudiantes manejan el concepto de variable debe reflejar esta situación.

Tomando como punto de partida las consideraciones anteriores y la descomposición del concepto de variable presentada al inicio de esta sección, se diseñó un cuestionario, que ha sido validado y utilizado en otros estudios relativos a la comprensión del concepto de variable (Ursini y Trigueros, 1997, Trigueros y Ursini, 1998; Lozano, 1998). En esta parte del trabajo se presenta el cuestionario y se hace un análisis de los ítems que lo componen en términos de la descomposición genética. Es este cuestionario el que se utilizará para probar el modelo de medida desarrollado en el capítulo anterior.

Definición del constructo

El constructo a medir mediante el instrumento es el de variable en el álgebra. Este es, como se ha visto, el concepto fundamental de esta materia y, por su complejidad, un concepto difícil de manejar.

Dentro de las matemáticas, el paso del aritmética al álgebra depende esencialmente de la comprensión de este concepto que a su vez es la clave para la comprensión de los conceptos más avanzados de las matemáticas. La posibilidad de usar flexiblemente el concepto de variable depende en gran medida de la capacidad de diferenciar e integrar las distintas formas en que el concepto se manifiesta, es decir, reconocerlas, combinarlas de manera efectiva y tener la posibilidad de pasar de una de ellas a otra sin dificultad.

Definición de los Factores

Con el instrumento se pretende diagnosticar hasta qué punto los estudiantes manejan el concepto de variable en el álgebra. Este concepto puede manifestarse, como ya se ha mencionado repetidamente, bajo caracterizaciones diversas que pueden aparecer claramente diferenciadas o entremezcladas en la solución de problemas algebraicos; siendo este último el caso más común.

Puesto que la comprensión del álgebra requiere de la posibilidad de diferenciar entre estas caracterizaciones y de integrarlas en el manejo de las situaciones problemáticas que se presentan, es conveniente en primer lugar diagnosticar si los alumnos las pueden diferenciar.

Por esta razón se tomaron como factores del costructo de variable precisamente estas caracerizaciones:

- La *variable como una incógnita*, es decir la variable como un número específico desconocido que se puede calcular en circunstancias dadas. La diferenciación de este factor implica la posibilidad de reconocer las expresiones algebraicas como ecuaciones en las que es necesario determinar un número específico, así como la posibilidad de manipular la variable independientemente de la operación o de las operaciones involucradas y por último la posibilidad de simbolizar problemas específicos en términos de ecuaciones. Estas posibilidades se pueden usar como indicadores en el diseño del instrumento, ya que permiten la posibilidad de traducción en ítems precisos que pueden medirse posteriormente usando métodos cuantitativos o cualitativos de análisis.

- La *variable como un número general.* En este caso la variable representa un número indeterminado involucrado en un método general. Lo que se espera del estudiante es la posibilidad de reconocer, manejar y simbolizar expresiones algebraicas, de considerar los posibles valores que la variable puede tomar en distintas situaciones o problemas y sobre todo, la capacidad de distinguirlas claramente de las ecuaciones. Nuevamente estas posibilidades se pueden usar como indicadores en el diseño del instrumento, ya que permiten la posibilidad de traducción en ítems precisos que pueden medirse posteriormente usando métodos cuantitativos o cualitativos de análisis.

- La *variable en una relación funcional*, es decir, variables que representan números cuyos valores cambian dentro de un cierto rango de valores ligados uno con otro mediante una relación. La idea de relación funcional puede considerarse desde dos perspectivas: Una estática, en la que la relación se concibe como la correspondencia entre dos conjuntos de valores que las variables involucradas pueden tomar en forma relacionada; otra perspectiva dinámica, en la que la variación de una de las variables determina la de la otra. El manejo competente del álgebra implica la capacidad de reconocer las relaciones funcionales y de interpretarlas en forma estática y en forma dinámica, dependiendo de la naturaleza del problema, la identificación de los valores que cada una de las variables puede tomar en términos de los que la otra puede tomar y la posibilidad de utilizar tablas, gráficas o fórmulas en la interpretación y en la solución de problemas. Estas posibilidades se pueden usar como indicadores en el diseño del instrumento, ya que permiten la posibilidad de traducción en ítems precisos que pueden medirse posteriormente usando métodos cuantitativos o cualitativos de análisis.

Estas tres caracterizaciones del concepto de variable están fuertemente interrelacionadas entre sí. En la mayoría de los problemas aparecen conjuntamente, aunque con diferentes grados de importancia. Por ello en el diseño del instrumento se buscarán ítems que permitan diferenciarlas lo más claramente posible, es decir, preguntas en las que el énfasis en uno de los aspectos se perciba con claridad.

Definición de las dimensiones

Para cada una de las caracterizaciones de la variable que se tomarán como factores en el diseño del instrumento, las diferentes formas en que cada una de ellas se puede presentar en la práctica se considerarán como las dimensiones del constructo, es decir:

- Manipulación. La manipulación de la variable implica la capacidad de manejarla algebraicamente, de combinarla con otras variables y de obtener su valor en caso de que sea necesario.

- Interpretación. La capacidad de interpretar la variable se refiere a la posibilidad de identificarla en la manifestación correspondiente y de asignarle valores sin necesidad de hacer manejos algebraicos.

- Simbolización. Cuando los problemas se presentan en forma verbal o en forma de patrón, tabla o gráfica, además de la capacidad de interpretar a la variable y de manipularla, es necesaria la posibilidad de escribirla como una fórmula utilizando símbolos.

Características Globales del Instrumento

Tomando en consideración los resultados de las investigaciones acerca del concepto de variable que se mencionaron anteriormente, el primer elemento que se consideró para el diseño del cuestionario fue la inclusión de ítems que abordaran cada una de las caracterizaciones del concepto.

Dado que se pretende explorar el grado de comprensión de los estudiantes del concepto de variable a través de sus distintas caractérizaciones, es importante evitar que los ítems utilizados en él contengan dificultades de manejo de expresiones algebraicas. En muchas ocasiones estas dificultades pueden esconder el grado de comprensión real del concepto al no permitir adentrarse en la capacidad que se quiere medir.

Por otra parte, las preguntas cuya respuesta requieren de la posibilidad de pasar de una caracterización del concepto a otra dentro de un mismo problema, deben excluirse de un primer instrumento exploratorio. Es importante deslindar, de una manera lo más transparente posible, si los estudiantes manejan las distintas manifestaciones del concepto de variable o en cuáles enfrentan dificultades.

Los textos y la docencia hacen, como hemos visto, un énfasis mayor en *la variable como incógnita*, por lo que es de esperar que sea ésta la caracterización de la variable que mejor manejen los estudiantes. Por esta razón, el diseño del instrumento aborda preferencialmente los otros dos aspectos del concepto: *la variable como número general* y *la variable en la relación funcional*, sin descuidar completamente el aspecto de *la variable como incógnita*.

Los ítems del cuestionario se diseñaron de dos maneras diferentes. En una primera versión se utilizaron en la forma de respuesta abierta. Esto permitió hacer un análisis cuantitativo, utilizando la teoría de respuesta a los ítems y la teoría clásica de los tests. Mediante investigación cualitativa basada en la descomposición genética del concepto y que aunada a los resultados de las investigaciones previamente mencionadas acerca del concepto de variable, se diseñaron opciones de respuesta con un fundamento cognitivo para cada uno de los ítems.

Objetivos de Medición

El cuestionario pretende adentrarse, como ya se ha mencionado, en los factores que caracterizan las distintas manifestaciones del concepto de variable

- la *variable como incógnita*
- la *variable como número general*
- la *variable en la relación funcional.*

Dentro de cada uno de esos factores se requiere considerar las dimensiones siguientes:

- manipulación
- interpretación
- simbolización

Cada uno de estos registros permite estudiar la comprensión del concepto desde perspectivas asociadas a la forma como se aborda el problema y pueden relacionarse con la teoría cognitiva que se ha desarollado aquí, con la cual se puede caracterizar la comprensión del concepto de variable de un sujeto a través de tres niveles:

- Acción
- Proceso
- Objeto

Estos niveles de comprensión pueden considerarse como indicadores de una comprensión del concepto de variable en un orden ascendente, es decir, que siguen un orden de menor a mayor profundización en la forma de comprender el concepto.

Presentación del cuestionario

En esta sección se presenta el cuestionario original en la versión de ítems de respuesta abierta, con el fin de hacer un análisis de las tareas que requiere la respuesta a cada uno de ellos y su relación con los niveles de la descomposición genética del concepto de variable.

CUESTIONARIO

En este ejercicio, solamente escribe una fórmula. NO CALCULES el número.

1. Escribe una fórmula que exprese: Un número desconocido multiplicado por 13 es igual a 127.___________________________
2. Escribe una fórmula que exprese: Un número desconocido multiplicado por la suma del mismo número desconocido con 2 es igual a 6.___________________________
3. Escribe una fórmula que exprese: Un número desconocido es igual a 6 más otro número desconocido.___________________________
4. Escribe una fórmula que exprese: Un número desconocido dividido por 5 y el resultado sumado a 7.___________________________

Para cada una de las siguientes expresiones ¿cuántos valores puede tomar la letra?

5. $x + 2 = 2 + x$ ___________________________
6. $3 + a + a = a + 10$ ___________________________
7. $x = x$ ___________________________
8. $4 + s$ ___________________________
9. $x + 5 = x + x$ ___________________________
10. $3 + a + a + a + 10$___________________________
11. $7x^2 = 2x - 5$ ___________________________
12. $\frac{x}{x^2 - 4} = 3$___________________________
13. $(x + 1)^2 = x^2 + 2x + 1$___________________________

14. $4 + x^2 = x(x + 1)$ ______________________

Para cada una de las siguientes expresiones escribe los valores que puede tomar la letra:

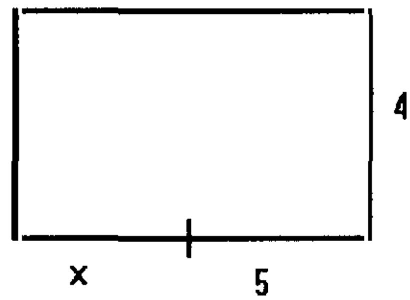
15. $13x + 27 - 2x = 30 + 5x$ ______________________
16. $(x + 3)^2 = 36$ ______________________
17. $4 + x = 2$ ______________________
18. $\frac{10}{1+x^2} = 2$ ______________________

Reduce las siguientes expresiones a una equivalente:

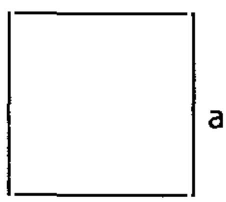
19. $(x^2 + 1)(x^2 - 2) =$ ______________________
20. $a + 5a - 3a =$ ______________________
21. $y^2 + 2y + 4y2 - 5y \quad -8 =$ ______________________

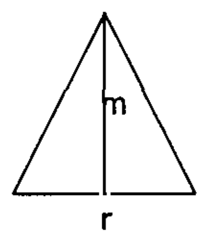
El perímetro de una figura se calcula sumando la longitud de sus lados. Escribe la fórmula que expresa el perímetro de la siguiente figura.

22.

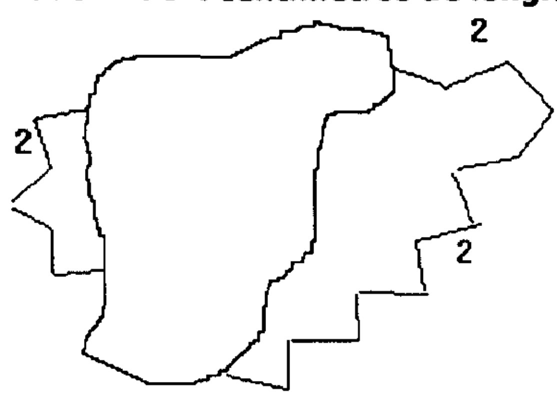


Escribe una fórmula para calcular el área de las siguientes figuras:

23. 

24. 

En la siguiente figura, el polígono no es completamente visible. Debido a que no sabemos cuántos lados tiene el polígono en total diremos que tiene N lados. Cada lado mide 2 centímetros de longitud.



25. Escribe una fórmula para calcular el perímetro del polígono.

Observa las siguientes figuras:

Número de puntos

Figura #1 ● 1

Figura #2 ● ● 4
● ●

Figura #3 ● ● ● 9
● ● ●
● ● ●

Figura #4

26. ¿Cuántos puntos hay en la figura # 4? ____________________
27. Dibuja la figura # 5 y da el número total de puntos. ____________________
28. Dibuja la figura # 6 y da el número total de puntos. ____________________
29. Imagínate que puedes seguir dibujando figuras hasta la figura # m. ¿Cuántos puntos en total tendrá la figura # m? ____________________

Si para hacer las figuras del ejercicio anterior vas agregando puntos

30. ¿Cuántos puntos agregas para pasar de la figura # 1 a la # 2? ____________________
31. ¿Cuántos puntos agregas para pasar de la figura # 2 a la # 3? ____________________
32. ¿Cuántos puntos agregas para pasar de la figura #m a la siguiente? ____________________
33. Escribe una fórmula que muestre como vas agregando puntos hasta llegar a la figura #m ____________________
34. Observa las siguientes igualdades y completa:

____________________ $1 + 2 + 3 = \frac{(3 \times 4)}{2}$

____________________ $1 + 2 + 3 + 4 = \frac{(4 \times 5)}{2}$

...

$1 + 2 + 3 + \dots + n =$

Si x + 3 = y

35. ¿Qué valores puede tomar **x**? ____________________
36. ¿Qué valores puede tomar **y**? ____________________

37. Si **y = 7 + x,** ¿qué les pasa a los valores de **y** cuando los valores de **x** aumentan? ______

Te encuentras en una papelería en donde se hacen fotocopias. Para evitarse estar haciendo multiplicaciones el empleado elabora una tabla.

38. Complétala.

número de copias	Precio
5	6.25
10	12.50
15	
	25.00
25	31.25
35	
	62.50
100	

39. Escribe una regla general donde **n** es el número de copias. ____________________

Observa la siguiente tabla.

40. Complétala y contesta las preguntas.

segundos	velocidad
0	0 m/s
10	30 m/s
15	
20	60 m/s
35	
50	
60	

41. Si aumenta el tiempo ¿qué le pasa a la velocidad, aumenta o disminuye? ___________

42. En una hoja aparte, sobre un sistema de coordenadas marca los puntos de la tabla y únelos trazando aproximadamente una curva.
43. Escribe la regla general que asocia a los números de la lista de la izquierda los números de la lista de la derecha. ______________________

Considera la siguiente expresión y=3+x.

44. Si queremos que los valores de **y** sean mayores que 3 pero más pequeños que 10, ¿qué valores puede tomar **x**? ____________________
45. Si **x** toma valores entre 8 y 15, ¿entre qué valores caerán los valores de **y**?________

De las siguientes dos expresiones n + 2 y 2 × n

46. ¿Cuál es más grande? ____________________
47. Justifica tu respuesta.____________________

El peso de la mercancía que se compra en el mercado se mide con una báscula. En el puesto de Don Panchito, por cada kilogramo de peso la charola de la báscula se desplaza 4 cm.

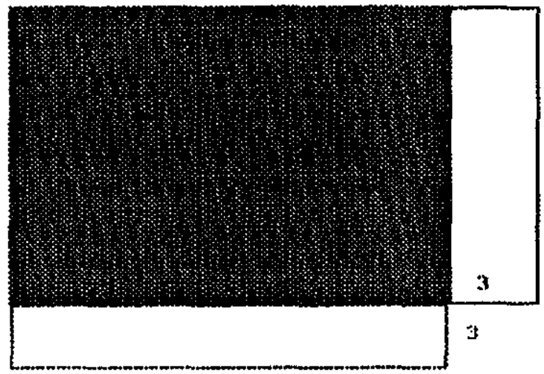
48. Encuentra la relación entre el peso de la compra y el desplazamiento de la charola.

____

49. Si la charola se desplaza 10.5 cm, ¿cuántos kilos pesa la mercancía? ____________

Escribe uno una fórmula para resolver los siguientes problemas:

50. El área total de la figura es 27. Calcula el lado del cuadrado sombreado.



51. Juan es 15 años mayor que Santiago. La suma de las dos edades es 41.¿Cuáles son las edades de Juan y Santiago? ____________________

52. Rentar un automóvil cuesta $25 por día, más $0.12 por kilómetro. ¿Cuántos kilómetros puede manejar Diego en un día, si sólo dispone de $40? ____________________

Observa los datos siguientes

x	y
0	0
10	100
-15	225
25	625
20	400
-10	100
15	225
-20	400

53. Determina qué pasa con el valor de **y** cuando el valor de **x** va creciendo ____________

54. ¿Para qué valor de **x,** alcanza **y** su valor máximo? ____________________

55. ¿Para qué valor de **x**, alcanza **y** su valor mínimo? ____________________

56. Escribe la regla general que relaciona a la variable **x** con la variable **y**. ____________

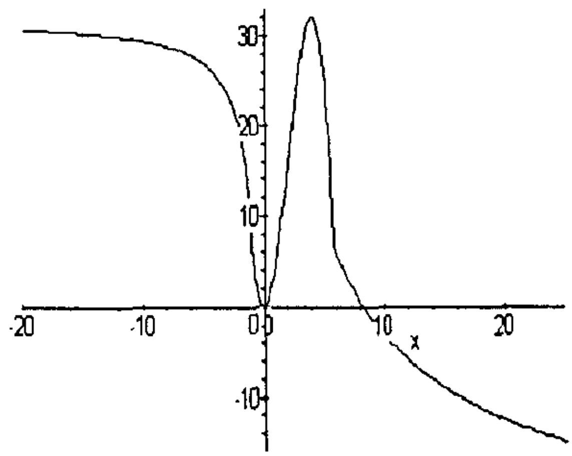
57. Si queremos que el valor de **y** esté entre 256 y 10000 ¿entre qué valores tiene que estar **x**? ____________________

58. Si **x** toma valores entre -2 y 26 ¿entre qué valores estará **y**? ____________________

Dada la expresión 40 - 15x - 3y = 17y - 5x

59. ¿Cuál es el valor de **y** para **x** = 16? ______________________
60. Para que el valor de **y** esté entre 1 y 5 ¿entre qué valores debe estar x?____________
61. Supón que **x** toma valores entre -5 y 5, ¿para qué valor de **x** alcanza **y** su valor máximo? ______________________

Dada la siguiente gráfica:



62. ¿Entre qué valores de **x**, los valores de **y** crecen? ___________________________
63. ¿Entre qué valores de **x**, los valores de **y** decrecen? _________________________
64. ¿Para qué valor de **x**, se obtiene el valor máximo de **y**? ______________________
65. ¿Para qué valor de **x** se obtiene el valor mínimo de **y**? _______________________

Análisis de las tareas para el cuestionario

A continuación se pone de manifiesto para cada ítem el aspecto de variable que éste pretende poner de manifiesto, la tarea matemática específica que se contempla en él y su relación con el marco teórico utilizado en el diseño:

Item Número **Tarea y relación con marco teórico**

1. Simbolización de la *variable como incógnita* con operación de producto. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable como incógnita.*

2. Simbolización de la *variable como incógnita* con distributividad y con múltiple aparición de la variable. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable como incógnita.*

3. Simbolización de la *variable en la relación funcional* con operación de suma. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable en una relación funcional.*

4. Simbolización de la *variable como número general* con operación de división. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de *la variable como número general.*

5. Interpretación de la *variable como número general* en tautología. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

6. Interpretación de la *variable como incógnita* con operación de suma y múltiple aparición de la variable. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como inc*ógnita.

7. Interpretación de la *variable como número general* en una tautología. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

8. Interpretación de la *variable como número general* con operación de suma. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

9. Interpretación de la *variable como incógnita,* con aparición múltiple de la incógnita y en la que la variable se encuentra de los dos lados de la ecuación. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable como incógnita.*

10. Interpretación de la *variable como número general* con múltiple aparición de la variable. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

11. Interpretación de la *variable como incógnita* con múltiple aparición la variable en forma cuadrática. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como incógnita.*

12. Interpretación de la *variable como in*cógnita, con cociente y multiple aparición de la variable en forma cuadrática. Transición del *nivel- proceso* al *nivel - objeto* en el manejo de la *variable como incógnita.*

13. Interpretación de la *variable como número general,* reconocimiento de una tautología con múltiple aparición de la variable en forma cuadrática. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

14. Interpretación de la *variable como incógnita* con distributividad y múltiple aparición de la variable en forma cuadrática. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como incógnita.*

15. Manipulación de la *variable como incógnita* con operación de suma y múltiple aparición de la variable. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable como incógnita.*

16. Manipulación de la *variable como incógnita* con aparición de la variable en forma cuadrática. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como incógnita.*

17. Interpretación de la *variable como incógnita* con operación de suma y solución en los negativos. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la incógnita.

18. Manipulación de la *variable como incógnita,* con operación de división y con la variable en forma cuadrática. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como incógnita.*

19. Manipulación de la *variable como número general,* con operación producto y con aparición múltiplede la variable en forma cuadrática. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la variable como número general.

20. Manipulación de la *variable como número general,* con operación de suma y con aparición múltiple de la variable. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable como número general.*

21. Manipulación de la *variable como número general,* con operación de suma y con aparición múltiple de la variable en forma cuadrática. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable como número general.*

22. Simbolización de la *variable como número general* a partir de la interpretación geométrica, con distributividad y en la que el símbolo está dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

23. Simbolización de la *variable como número general* a partir de la interpretación geométrica con operación de producto y con símbolo dado. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable como número general.*

24. Simbolización de la *variable como número general* a partir de la interpretación geométrica, con operación de producto y con símbolo dado. Transición de *nivel - objeto* a *nivel - proceso* en el manejo de la *variable como número general.*

25. Simbolización de la *variable como número general* con interpretación geométrica, símbolo dado y la idea de generalización incluida. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable como número general.*

26. Hacer evidente el patrón a generalizar. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

27. Hacer evidente el patrón a generalizar. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

28. Hacer evidente el patrón a generalizar. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

29. Simbolización de la *variable como número general* a partir de la interpretación de patrones, con símbolo dado y con la idea de generalización incluida. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

30. Hacer evidente el patrón a generalizar. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

31. Hacer evidente el patrón a generalizar. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

32. Simbolización de la *variable como número general* a partir de la interpretación de patrones, con símbolo dado y con la idea de variación incluida. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

33. Simbolización de la *variable como número general* a partir de la generalización de patrones, con símbolo dado y con la idea de variación incluida. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

34. Simbolización de la *variable como número general* a partir de la generalización de patrones con símbolo dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como número general.*

35. Interpretación de la *variable en la relación funcional,* concepto dinámico de dominio. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

36. Interpretación de la *variable en la relación funcional,* concepto dinámico de rango. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable en la relación funcional.*

37. Interpretación de la *variable en la relación funcional,* concepción dinámica de la variación de las variables con énfasis en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

38. Manipulación de la *variable en la relación funcional* en la forma de tabla, énfasis en la idea de correspondencia. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

39. Simbolización de la *variable en la relación funcional* a partir de la generalización de los datos de una tabla y con símbolo dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

40. Manipulación de la *variable en la relación funcional* expresada en forma de tabla. . Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

41. Interpretación de la *variable en la relación funcional,* dinámica de la variación conjunta de las variables y con énfasis en el dominio. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

42. Graficación de la *variable en la relación funcional* a partir de una tabla. . Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

43. Simbolización de la *variable en la relación funcional* a partir de la generalización de la correspondencia entre los datos de una tabla y en la que el símbolo no está dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

44. Interpretación de la *variable en la relación funcional,* concepción dinámica de la variación en un intervalo, relación dominio rango dado este último. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

45. Interpretación de la *variable en la relación funcional,* concepción dinámica de la variación en un intervalo, relación dominio rango dado el dominio. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la variable en al relación funcional.

46. Interpretación de la *variable como número general,* comparación de la variación en dos expresiones. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable como número general.*

47. Interpretación de la *variable como número general,* comparación y justificación de la variación en dos expresiones. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable como número general.*

48. Simbolización de la *variable en la relación funcional* a partir un problema formulado verbalmente y en el que no está dado el símbolo, generalización de la idea de variación dinámica. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

49. Manipulación de la *variable en la relación funcional* a partir de una tabla y una posible fórmula que la expresa, con énfasis en la relación entre el dominio

y el rango. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable en la relación funcional.*

50. Simbolización de la *variable como incógnita* a partir de la interpretación de la variable en un problema geométrico, con aparición de la variable en forma cuadrática y en la que el símbolo no está dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como incógnita.*

51. Simbolización de la *variable como incógnita* a partir de la interpretación de la variable en un problema expresado verbalmente, con múltiple aparición de la variable, con asociatividad y en la que el símbolo no está dado. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable como incógnita.*

52. Simbolización de la *variable como incógnita* a partir de la interpretación de la variable en un problema expresado verbalmente, con operación de suma y en la que el símbolo no está dado. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable como incógnita.*

53. Interpretación de la *variable en la relación funcional* a partir de la lectura de los datos desordenados de una tabla que representa una relación cuadrática, concepción dinámica de la variación conjunta de las dos variables. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

54. Interpretación de la *variable en la relación funcional,* a partir de la lectura de los datos desordenados de una tabla que representa una relación cuadrática, concepción estática de la relación, búsqueda de valor máximo en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

55. Interpretación de la *variable en la relación funcional*, a partir de la lectura de los datos desordenados de una tabla que representa una relación cuadrática, concepción estática de la relación, búsqueda de valor mínimo en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

56. Simbolización de la *variable en la relación funcional*, a partir la lectura de los datos desordenados de una tabla que representa una relación cuadrática y en la que el símbolo no está dado. Transición de *nivel - proceso* a *nivel - objeto* en el manejo de la *variable en la relación funcional.*

57. Interpretación de la *variable en la relación funcional* a partir de la lectura de los datos desordenados de una tabla que representa una relación cuadrática, concepción dinámica de la variación, noción de intervalo en el dominio dado el intervalo en el rango. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

58. Interpretación de la *variable en la relación funcional* a partir de la lectura de los datos desordenados de una tabla, concepción dinámica de la variación, noción de intervalo en el rango dado el intervalo en el dominio. Transición del *nivel - proceso* al *nivel - objeto* en el manejo de la *variable en la relación funcional.*

59. Manipulación de la *variable en la relación funcional* a partir de una fórmula, concepción estática de la función. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable en la relación funcional.*

60. Interpretación de la *variable en la relación funcional* a partir de una fórmula, concepción dinámica de la variación, noción de intervalo en el

dominio dado el intervalo en el rango. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable en la relación funcional.*

61. Interpretación de la *variable en la relación funcional* a partir de una fórmula, concepción dinámica y estática de la relación, búsqueda de valor en el dominio dado el valor en el rango. Transición del *nivel - acción* al *nivel - proceso* en el manejo de la *variable en la relación funcional.*

62. Interpretación de la *variable en la relación funcional* a partir de una gráfica. Interpretación dinámica de la variación conjunta de las variables a partir de una gráfica, noción de intervalo en el dominio dada la propiedad que debe cumplirse en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

63. Interpretación de la *variable en la relación funcional* a partir de una gráfica. Interpretación dinámica de la variación conjunta de las variables a partir de una gráfica, noción de intervalo en el dominio dada la propiedad que debe cumplirse en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

64. Interpretación de la *variable en la relación funcional* a partir de una gráfica. Interpretación estática de la variación conjunta de las variables a partir de una gráfica, noción de valor en el dominio dada la propiedad que debe cumplirse en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional.*

65. Interpretación de la *variable en la relación funcional* a partir de una gráfica. Interpretación estática de la variación conjunta de las variables a partir de una gráfica, noción de valor en el dominio dada la propiedad que

debe cumplirse en el rango. Transición de *nivel - acción* a *nivel - proceso* en el manejo de la *variable en la relación funcional*.

De entre todas estas preguntas se hizo una selección de treinta para probar el modelo de medida que se presentó en el capítulo II del presente trabajo. Estas treinta preguntas se escribieron en la modalidad de opción múltiple y se presentan en la siguiente sección.

Sección III.6.1 Aplicación del modelo a los ítems diseñados y análisis de resultados

El objetivo de esta sección consiste en analizar la plausibilidad de la aplicación del modelo de medida descrito en el capítulo II a una situación particular.

La aplicación del modelo de medida descrito en el capítulo II de este trabajo requiere del diseño de un cuestionario y de su uso con alumnos para después estimar los parámetros y analizar la información obtenida. Para ello se tomó como punto de partida el cuestionario relativo al concepto de variable en el álgebra que se describió en la sección anterior y que ya ha sido probado y validado. Se eligieron, de entre todas las preguntas que componen dicho cuestionario, treinta ítems representativos de las distintas caracterizaciones de la variable y que, en términos del marco teórico descrito anteriormente, pudieran dar cuenta de posibles transiciones entre el *nivel - acción* y el *nivel - proceso* o el *nivel - proceso* y el *nivel - objeto* descritos en el modelo diagnóstico.

Todos los ítems del cuestionario son preguntas que intentan medir el nivel de conocimiento del concepto de variable en el álgebra elemental y, por ello es de esperarse que en cuanto al modelo de medida desarrollado en este trabajo los casos que se presentarán corresponden a los casos 6, 7 y 8 del segundo submodelo descrito en la sección II.3 del presente trabajo.

En esta parte del trabajo se presentan en primer lugar los ítems seleccionados con las opciones de respuesta incluidas. Es importante recalcar que dichas opciones se diseñaron tomando en consideración los resultados reportados en la literatura disponible acerca del aprendizaje del concepto de variable. A continuación se analizan los ítems del cuestionario desde la perspectiva del marco teórico de análisis; se describe brevemente la metodología empleada para la estimación de los parámetros del cuestionario, que se basa en la metodología descrita en la sección II.7 del capítulo anterior y se muestran y discuten los resultados obtenidos.

Una vez conocidos los parámetros de cada una de las opciones de respuesta a los ítems del cuestionario, éstos se analizan a la luz del modelo de medida y, al mismo tiempo, en términos del modelo diagnóstico. Puesto que en la presentación del modelo de medida se hace énfasis en la utilidad de las gráficas de las probabilidades de elección de las distintas opciones de respuesta a los ítems, tanto en el espacio fase cuanto en el plano probabilidad de elección contra capacidad de comprensión de la variable, se incluyen dichas gráficas en el análisis de cada ítem.

Finalmente se concluye esta sección comentando los resultados de la aplicación del modelo: sus ventajas y sus limitaciones.

Elección de preguntas del cuestionario original para la prueba del modelo

De los 65 ítems que componen el cuestionario original que se mostró en la sección anterior, se hizo una selección de 30 preguntas para hacer más fácil la prueba del modelo matemático de medida que se presentó en el capítulo II. Estas preguntas se reescribieron utilizando un formato de respuesta de opción múltiple, como lo requiere el modelo.

A continuación se muestra el nuevo cuestionario. Conviene aclarar, sin embargo, que si bien en la forma en la que aquí se presenta, la opción correcta de respuesta a cada pregunta aparece como la opción 1, en el cuestionario respondido por los estudiantes los distractores se colocaron de manera aleatoria en cada una de las preguntas. La razón por la cual el orden de los distractores se cambia en esta presentación responde a la facilidad de identificación de los parámetros con el modelo propuesto y a una simplificación en la interpretación de los mismos.

CUESTIONARIO

1. ¿Cuál es la expresión que representa al siguiente enunciado: Un número desconocido es igual a 6 más otro número desconocido.?
   1) $x = 6 + y$
   2) $x = 6 + x$
   3) $x + 6 = y$

2. ¿Cuál es la expresión que representa al siguiente enunciado: Un número desconocido dividido por 5 y el resultado sumado a 7?
   1) $x/5 + 7$
   2) $x/5 = y + 7$
   3) $(x +7)/5$

3. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $x+2 = 2+x$
   1) Un numero infinito de valores
   2) Un valor

3) Uno $x = 1$

4. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $x = x$
   1) Un número infinito de valores
   2) Uno, $x = 0$
   3) Ningún valor

5. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $4 + s$
   1) Un número infinito
   2) Uno, $s = -4$
   3) Uno, $s = 4$

6. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $x+5 = x+x$
   1) Uno
   2) Tres
   3) Ninguno

7. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $\frac{x}{x^2 - 4} = 3$
   1) Dos
   2) Uno
   3) Ninguno

8. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $(x + 1)^2 = x^2 + 2x + 1$
   1) un número infinito de valores
   2) uno
   3) dos

9. ¿Cuántos valores puede tomar la letra en la siguiente expresión?

$$4+x^2=x(x + 1)$$

1) uno
2) dos
3) un número infinito

10. En la siguiente expresión encuentra los valores que puede tomar la letra: $(x + 3)^2 = 36$

1) $x = 3$ y $x = 9$
2) $x = 3$
3) $x = 9$

11. En la siguiente expresión escribe los valores que puede tomar la letra: $4+x=2$

1) $x = -2$
2) $x = 2$
3) $x = 6$

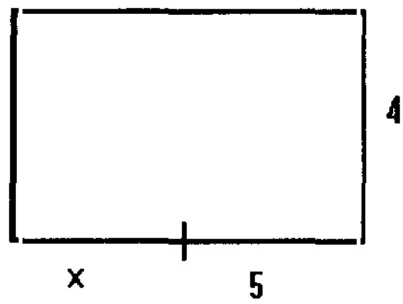
12. En la siguiente expresión escribe los valores que puede tomar la letra: $\frac{10}{1+x^2} = 2$

1) $x = \sqrt{2}$ o $x = - \sqrt{2}$
2) $x = \sqrt{2}$
3) ninguno

13 ¿Cuál de las expresiones siguientes es equivalente a $(x^2 + 1)(x^2 - 2)$?

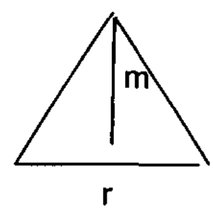
1) $x^4 - x^2 - 2$
2) $x^4 - 2$
3) $x^4 - 3x^2 - 2$

14. La fórmula para obtener el perímetro de la siguiente figura es:



1) P = (x + 5) + 4
2) 2(b + h)
3) (5 + 4)2

15. El área de la siguiente figura está dada por



1) (r m) / 2
2) (b h) / 2
3) 2(b h)

16. Observa las siguientes figuras:

Número de puntos

Figura #1  1

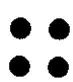
Figura #2  4

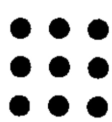
Figura #3  9

Figura #4

¿Cuántos puntos hay en la figura # 4?

¿Cuántos puntos hay en la figura # 5?

¿Cuántos puntos hay en la figura # 6?

Imagina que puedes seguir dibujando figuras hasta la figura # m. ¿Cuántos puntos en total tendrá la figura # m?

1) m^2
2) 2m
3) m+ m

17. Si para hacer las figuras del ejercicio anterior vas agregando puntos

¿Cuántos puntos agregas para pasar de la figura # 1 a la # 2?

¿Cuántos puntos agregas para pasar de la figura # 2 a la # 3?

La fórmula que muestra cuántos puntos se deben agregar para pasar de la figura #m a la #m + 1 es

1) 2m + 1
2) m x m
3) m + m

18. La siguiente tabla indica el precio que corresponde a un número dado de fotocopias en la papelería. La tabla esta incompleta.

Número de copias	Precio
5	6.25
10	12.50
15	
	25.00
25	31.25
35	
	62.50
100	

El precio que corresponde a 35 fotocopias es

1) p = 43.75
2) p = 42
3) x = 19.75

19. Usando los datos del problema anterior. La regla general que describe el precio de las fotocopias usando **n** para representar el número de fotocopias es

1) p = 1.25n
2) n = 1.25p
3) 5n = 6.25

20. Observa la siguiente tabla, que está incompleta:

Segundos	Velocidad
0	0 m/s
10	30 m/s
15	
20	60 m/s
35	
50	
60	

La regla general que asocia la velocidad al tiempo es:

1) v = 15t
2) 10s = 30v
3) d = v / t

Considera la siguiente expresión y=3+x.

21. Si queremos que los valores de **y** sean mayores que 3 pero más pequeños que 10, los valores que puede tomar **x** son:

1) de 0 a 7
2) 0, 1, 2, 3, 4, 5, 6, 7
3) 0 y 7

22. Si **x** toma valores entre 8 y 15, los valores de **y** estarán entre:

1) entre 11 y 18
2) 11, 12, 13, 14, 15, 16, 17, 18
3) 11 y 18

23. De las siguientes dos expresiones $n + 2$ y $2 \times n$, la más grande es:

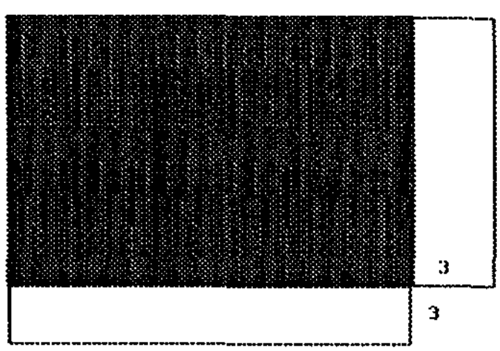
1) depende del valor de n
2) 2xn
3) 2 + n

24. El peso de la mercancía que se compra en el mercado se mide con una báscula. En el puesto de Don Panchito, por cada kilogramo de peso la charola de la báscula se desplaza 4 cm. La relación entre el peso de la compra y el desplazamiento de la charola está dada por:

1) $x = 4p$
2) 1kg = 4 cm
3) $4 = x$

25. El área total de la siguiente figura es 27. La fórmula para calcular el lado del cuadrado sombreado es:

1) $x^2 + 6x = 27$
2) $a^2 + b^2 = 27$
3) $L = \sqrt{27}$



26. Juan es 15 años mayor que Santiago. La suma de las dos edades es 41. La ecuación que permite encontrar las edades de Juan y Santiago es:

1) j + j + 15 = 41
2) s + s + 15 = 41
3) j + s = 41

27. Rentar un automóvil cuesta $25por día, más $0.12 por kilómetro. ¿Cuántos kilómetros puede manejar Diego en un día, si sólo dispone de $40?
    1) 125
    2) 15
    3) 12
28. Observa los datos siguientes:

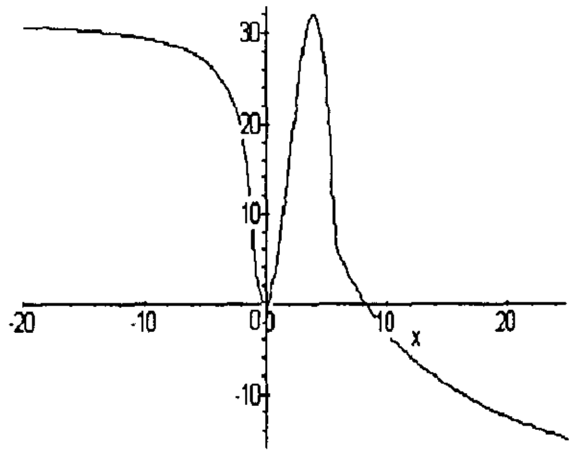
x	y
0	0
10	100
-15	225
25	625
20	400
-10	100
15	225
-20	400

Si **x** toma valores entre -2 y 26 ¿entre qué valores estará **y**?
1) entre 0 y 676
2) entre 4 y 676
3) entre –4 y 676

29. Dada la expresión 40 - 15x - 3y = 17y - 5x, el valor de **y** para **x** = 16 es:
    1) y = -6
    2) y = 10

3) $y = 60$

30. Dada la siguiente gráfica:



¿Entre qué valores de **x**, los valores de **y** decrecen?

1) entre $-\infty$ y 0 y entre 5 y ∞
2) entre ∞ y 5 y entre 5 y ∞
3) entre 10 y 20

Sección III.6.2 Análisis de los ítems de acuerdo al marco teórico

Cada uno de los ítems del cuestionario puede analizarse en términos del modelo de diagnóstico descrito anteriormente en este capítulo. Este marco teórico es el que se utilizará conjuntamente con el modelo de medida para analizar los resultados de la aplicación del cuestionario. A continuación se muestra el análisis de cada uno de los treinta ítems seleccionados para el cuestionario de prueba.

ÍTEM 1

7. ¿Cuál es la expresión que representa al siguiente enunciado: Un número desconocido es igual a 6 más otro número desconocido.?
   4) $x = 6 + y$
   5) $x = 6 + x$
   6) $x + 6 = y$

La simbolización de la variable en el enunciado que se presenta en este ítem requiere que el estudiante distinga que se encuentra frente a una situación en la que hay dos *variables en relación funcional* y que sea capaz de reconocer la forma de la dependencia a partir de los datos que se presentan en el enunciado del problema. Un estudiante en *nivel - proceso* puede reconocer la existencia de dos variables en correspondencia, pero tener dificultades al interpretar la forma en la que la relación se simboliza en una expresión. Un estudiante que se encuentra en *nivel - objeto* es capaz de interpretar la relación presentada en el enunciado y además simbolizarla en la forma correcta.

ÍTEM 2

8. ¿Cuál es la expresión que representa al siguiente enunciado: Un número desconocido dividido por 5 y el resultado sumado a 7.?

1) $x/5 + 7$
2) $x/5 = y + 7$
3) $(x + 7)/5$

La simbolización de esta expresión algebraica implica el manejo de la *variable como número general* como un objeto matemático sobre el cual se pueden hacer distintos tipos de operaciones. Los estudiantes que se encuentran en *nivel - proceso* en su forma de entender la variable, efectúan la primera operación, es decir, la división entre cinco y no pueden concebir a la expresión $x/5$ como un nuevo objeto sobre el que hay que operar nuevamente, por ello, igualan la expresión $x/5$ con una nueva variable, y, sin tomar en consideración el papel que juega el signo de igualdad, efectúan a continuación la segunda operación. Estos estudiantes se apoyan en el símbolo de igualdad para poder llevar a cabo los pasos consecutivos de la operación. Esto constituye un indicador de que se encuentran en *nivel - proceso* en su comprensión de la *variable como número general*.

ÍTEM 3

9. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $x+2 = 2+x$
   1) Un numero infinito de valores
   2) Un valor
   3) Uno, $x = 1$

La distinción entre una expresión abierta y una ecuación implica la posibilidad de concebir a la expresión como un objeto que es válido independientemente del valor que la literal toma. Si se piensa en ella como una expresión abierta en la que la variable se concibe en *nivel - proceso*, es posible verificar la validez de la expresión para valores particulares de la variable, pero, cuando no se distingue de una ecuación, se tiene una conceptualización de la variable en *nivel - acción*. En este problema los estudiantes que se encuentran en *nivel - acción* en su conceptualización de la *variable como número general* interpretan el signo de igualdad como un indicador de que la expresión que se presenta es una ecuación y, puesto que ésta es lineal, asumen que la variable puede tomar solamente un valor, que asignan generalmente por inspección.

ÍTEM 4

10. ¿Cuántos valores puede tomar la letra en la siguiente expresión? x = x

1) Un número infinito de valores
2) Uno, x = 0
3) Ningún valor

La distinción entre una expresión abierta y una ecuación implica la posibilidad de concebir a la expresión como un objeto que es válido independientemente del valor que la literal toma. Si se concibe el papel de la variable en la expresión como un *número general* en *nivel - proceso*, es posible, al igual que en el ítem anterior, verificar la validez de la expresión para valores particulares de la variable. Sin embargo, cuando el papel de la variable no está claro y se confunde la expresión con una ecuación, es imposible ver la igualdad que se presenta en este ítem como un todo. En este problema los estudiantes que se encuentran en *nivel - acción* en su conceptualización de la *variable como número general* interpretan el signo de igualdad como un indicador de que la expresión que se presenta es una ecuación y, puesto que ésta es lineal, utilizando un hecho memorizado, concluyen que la variable es una incógnita que puede tomar solamente un valor, que asignan generalmente por inspección. A diferencia del ítem 3, en este ítem los estudiantes intentan con mayor frecuencia manipular la expresión encontrándose con la expresión 0=0, que les causa una enorme confusión, de ahí que algunos de ellos concluyan que la variable no puede tomar ningún valor.

ÍTEM 5

11. ¿Cuántos valores puede tomar la letra en la siguiente expresión? 4 + s

1) Un número infinito
2) Uno, s = -4
3) Uno, s = 4

Para resolver correctamente este ítem se requiere que el estudiante tenga una concepción de la *variable como número general* en *nivel - proceso* y así sea capaz de considerar que en la expresión la variable puede tomar cualquier valor y no únicamente valores enteros. La consideración de la *variable como número general* en *nivel - acción* induce a los estudiantes a la idea de que para determinar los valores que la variable puede tomar es necesario considerar a la expresión como una ecuación; al no contar con un signo de igualdad, ni con un valor determinado después de éste, ellos lo introducen haciendo la expresión igual a cero. De esta manera son capaces de resolver la pregunta planteada.

ÍTEM 6

12. ¿Cuántos valores puede tomar la letra en la siguiente expresión? x+5 = x+x

1) Uno
2) Tres
3) Ninguno

Un estudiante que se encuentra en *nivel - proceso* en el manejo de la *variable como incógnita* es capaz de manipular la literal en la ecuación para transformar esta última en tal forma que le sea posible encontrar el valor que la hace verdadera. Puede, además reconocer que la variable que aparece en los dos lados de la expresión representa la misma cantidad. Cuando un estudiante se encuentra en *nivel - acción* en su conceptualización de la *variable como incógnita* puede tener problemas para reconocer el papel que la variable juega en la expresión e incluso puede presentar el fenómeno conocido como polisemia, que se refiere a la asignación de diferentes valores para cada aparición de la misma literal en una expresión.

ÍTEM 7

7. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $\frac{x}{x^2 - 4} = 3$

   1) Dos
   2) Uno
   3) Ninguno

La expresión que se presenta en este ítem es compleja. Su solución correcta requiere de un manejo eficaz de las propiedades de la *variable como incógnita* y la posibilidad de concebir el denominador como un objeto que se puede manipular como una entidad global, dando como resultado una expresión cuadrática que tiene dos soluciones. Este tipo de manipulación puede asociarse a un *nivel - objeto* en el manejo de la variable. Si el estudiante tiene una concepción en *nivel - proceso* de la variable no es capaz de pensar en el denominador como una totalidad y considera que la variable cuadrática que aparece en el denominador puede cancelarse con la variable que aparece en el numerador. Haciendo esta operación la ecuación se convierte en una ecuación lineal que tiene únicamente una solución.

ÍTEM 8

12. ¿Cuántos valores puede tomar la letra en la siguiente expresión? $(x + 1)^2 = x^2 + 2x + 1$

   1) un número infinito de valores
   2) uno
   3) dos

La dificultad de este ítem, en comparación con los ítems 3 y 4 radica en que en él se muestra una tautología en forma cuadrática. Es por ello que la distinción entre una expresión abierta y una ecuación se torna más difícil e implica la posibilidad de concebir a la expresión como un objeto que es válido independientemente del valor que la literal toma. Si se concibe a la expresión como un proceso, es posible verificar la validez de la expresión para valores particulares de la variable, pero, al no distinguirla de una ecuación, es imposible verla como un todo. La presencia del exponente induce fuertemente a los estudiantes que se encuentran en *nivel - proceso* a interpretar la expresión como una ecuación, a utilizar un hecho memorizado que les indica que una ecuación de segundo grado tiene dos soluciones y a encontrarlas mediante un método conocido.

ÍTEM 9

13. ¿Cuántos valores puede tomar la letra en la siguiente expresión?

$4 + x^2 = x(x + 1)$

1) uno
2) dos
3) un número infinito

La respuesta correcta a este ítem implica el reconocimiento de la *variable como incógnita* y la posibilidad de manipular la expresión correctamente a fin de encontrar el valor de la incógnita. Dicha manipulación requiere que el estudiante se encuentre en *nivel - proceso* en su comprensión de la incógnita, de tal suerte que sea capaz de darse cuenta de que los términos cuadráticos se cancelan y lo que queda es una ecuación lineal simple. La elección de la segunda opción indica que si bien los estudiantes interpretan la *variable como incógnita,* muestran una propensión a buscar en la expresión dada la presencia de signos que les permitan dar respuesta a la pregunta; en este caso, el hecho de que en el lado izquierdo de la expresión aparezca el exponente cuadrático induce a los estudiantes a responder que la variable puede tomar dos valores, recordando un hecho memorizado.

ÍTEM 10

14. En la siguiente expresión encuentra los valores que puede tomar la letra:

$(x + 3)^2 = 36$

1) x = 3 y x = 9
2) x = 3
3) x = 9

La solución correcta de este ítem implica la posibilidad de manipular la *variable como incógnita* en una ecuación, tomando en consideración todas las posibilidades. Los estudiantes que pueden manejar de esta manera la *variable como incógnita* se encuentran en *nivel - proceso* en su conceptualizacion; mientras que aquéllos que responden en forma incorrecta utilizan procedimientos automáticos o memorizados para encontrar la solución o cometen errores en la manipulación, por lo que puede considerarse que se encuentran en *nivel - acción* en la forma en la que comprenden el concepto de *variable como incógnita.*

ÍTEM 11

15. En la siguiente expresión escribe los valores que puede tomar la letra: 4+x=2

1) x = -2
2) x = 2
3) x = 6

La solución de esta ecuación, si bien no involucra la manipulación de la *variable como incógnita,* implica la necesidad de aceptar como resultado un número negativo. Los alumnos que se encuentran en *nivel - acción* en el manejo de la incógnita, cuando enfrentan ecuaciones de este grado de dificultad, tienen dificultades para aceptar valores negativos como solución. Los estudiantes que se encuentran en *nivel - proceso* en el manejo de la *variable como incógnita* no tienen problema en aceptar a un número negativo como solución de una ecuación.

ÍTEM 12

12. En la siguiente expresión escribe los valores que puede tomar la letra:

$$\frac{10}{1+x^2} = 2$$

1) $x = \sqrt{2}$ o $x = -\sqrt{2}$
2) $x = \sqrt{2}$
3) ninguno

La respuesta correcta a esta pregunta implica que el estudiante se encuentra en *nivel - objeto* en el manejo de la *variable como incógnita*, dado que es capaz de manipular la literal en una ecuación cuadrática cuya presentación sigue un formato poco común en la instrucción, y transformar la ecuación para encontrar los valores de la incógnita que la satisfacen. Cuando un estudiante se encuentra en *nivel - proceso* en su conceptualización de la *variable como incógnita* suele tener problemas con el manejo de las ecuaciones cuadráticas, el hecho de que la variable se encuentre en el denominador de una expresión suele causar mayores dificultades, frente a esta situación, el alumno es capaz de identificar solamente una de las soluciones o concluir que no es posible resolver una ecuación cuando tiene esa forma.

ÍTEM 13

14 ¿Cuál de las expresiones siguientes es equivalente a $(x^2 + 1)(x^2 - 2)$?

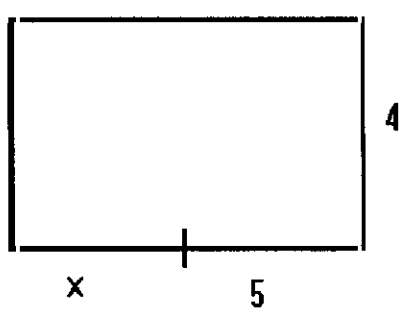
1) $x^4 - x^2 - 2$
2) $x^4 - 2$

3) $x^4 - 3x^2 - 2$

La solución correcta de este ítem implica la concepción de la *variable como número general* en *nivel - proceso*, es decir que el estudiante haya interiorizado la posibilidad de operar con las variables de tal manera que pueda efectuar sin problemas las operaciones necesarias para desarrollar la expresión que se le presenta y transformarla en otra equivalente. Si estas acciones no se han interiorizado, los estudiantes no son capaces de aplicar correctamente la propiedad distributiva del producto y aplican únicamente la acción de elevar al cuadrado cada una de las variables, sin tener en cuenta el proceso involucrado en la manipulación implicada por la presencia de los dos paréntesis.

ÍTEM 14

14. La fórmula para obtener el perímetro de la siguiente figura es:

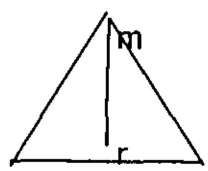


1) P = (x + 5) + 4
2) 2(b + h)
3) (5 + 4)2

La respuesta a esta pregunta en forma correcta implica la capacidad de asignar a las dimensiones de la figura cualquier valor y representarlas mediante cualquier número general. Es decir, implica ver a la figura dada como representante de una clase general de figuras y calcular perímetro utilizando cualquier símbolo que no esta asociado a un valor predeterminado. En este caso, además, la variable no se refiere a la totalidad de una de las dimensiones de la figura por lo que es necesario agruparla con la cantidad apropiada y considerar a esa suma como un todo. Los estudiantes que son capaces de hacer esto se encuentran en su conceptualización de la *variable como número general* en *nivel - objeto,* mientras que quienes tienen dificultades al encontrar la respuesta correcta, aunque conocen la fórmula geométrica que se utiliza en la solución no son capaces de utilizarla más que en *nivel – proceso,* en el que requieren aún de utilizar los nombres de las variables que aparecen comúnmente en los apuntes o en los textos o son incapaces de agrupar los términos necesarios para identificar una dimensión de la figura.

ÍTEM 15

15. El área de la siguiente figura está dada por



1) (r m) / 2
2) (b h) / 2
3) 2(b h)

Este ítem es análogo al ítem 14; sin embargo en él la variable se indica directamente en las dimensiones del problema. La respuesta a esta pregunta en forma correcta implica la capacidad de asignar a las dimensiones de la figura cualquier valor y representarlas mediante cualquier número general. Es decir, implica ver a la figura dada como representante de una clase general de figuras y calcular su área utilizando cualquier símbolo que no esta asociado a un valor predeterminado. Los estudiantes que son capaces de hacer esto se encuentran en *nivel - proceso* de conceptualización de la *variable como número general,* mientras que quienes tienen dificultades al encontrar la respuesta correcta, aunque conocen la fórmula geométrica que se utiliza en la solución no son capaces de utilizarla más que en *nivel – acción,* en el que utilizan en el cálculo los nombres de las variables que aparecen comúnmente en los apuntes o en los textos y no toman en cuenta el símbolo utilizado en este triángulo específico.

ÍTEM 16

16. Observa las siguientes figuras:

		Número de puntos
Figura #1	●	1
Figura #2	● ● ● ●	4
Figura #3	● ● ● ● ● ● ● ● ●	9
Figura #4		

¿Cuántos puntos hay en la figura # 4?
¿Cuántos puntos hay en la figura # 5?
¿Cuántos puntos hay en la figura # 6?
Imagina que puedes seguir dibujando figuras hasta la figura # m. ¿Cuántos puntos en total tendrá la figura # m?

1) m^2
2) 2m
3) m+ m

La respuesta correcta a este ítem requiere de la capacidad de los estudiantes de generalizar un patrón y escribir la relación que les permite encontrar el número de puntos correspondiente a cada una de las figuras que se presentan en el proceso de generalización, independientemente del número de puntos particulares. Un estudiante que se encuentra en *nivel - proceso* de conceptualización de la *variable como número general* puede interpretar el patrón involucrado, pero, al no haber interiorizado el proceso como un todo, le es imposible simbolizarlo utilizando un símbolo que represente a un número general. Un estudiante que se encuentra en *nivel - objeto* en la conceptualización de la *variable como número general,* ha interiorizado el proceso de generalización y es capaz de simbolizarlo utilizando la variable como el número general.

ÍTEM 17

17. Si para hacer las figuras del ejercicio anterior vas agregando puntos

¿Cuántos puntos agregas para pasar de la figura # 1 a la # 2?

¿Cuántos puntos agregas para pasar de la figura # 2 a la # 3?

La fórmula que muestra cuántos puntos se deben agregar para pasar de la figura #m a la #m + 1 es

1) 2m + 1
2) m x m
3) m + m

El ítem mide la capacidad de simbolización de la *variable como un número general* a través del reconocimiento, por una parte, del patrón y por otra parte de la variación de ese patrón. Los estudiantes que responden bien esta pregunta se encuentran muy probablemente en *nivel - objeto* en el manejo de la variable, ya que el reconocimiento de la variación de un patrón y de la forma de simbolizarlo implica un manejo muy flexible de la noción de variable y la interiorización de los procesos de generalización y simbolización. Los alumnos que eligen opciones incorrectas pueden encontrarse en *nivel - proceso* o en *nivel - acción*. En el *nivel - proceso* son capaces de reconocer parte de la variación del patrón. En el *nivel - acción* no reconocen el patrón asociado a la variación. El reconocimiento del patrón puede explicarse como la generalización de la acción de contar los puntos de la figura y la capacidad de predecir el número de puntos para cualquiera de las figuras. El reconocimiento de la variación del patrón involucra por una parte lo anterior y, por otra, la interiorización del proceso de comparar una figura con otra para obtener el número de puntos añadidos y la generalización del nuevo patrón mediante el cual se van añadiendo los puntos al pasar de una figura a la otra.

ÍTEM 18

18. La siguiente tabla indica el precio que corresponde a un número dado de fotocopias en la papelería. La tabla esta incompleta.

Número de copias	Precio
5	6.25
10	12.50
15	
	25.00
25	31.25
35	
	62.50
100	

El precio que corresponde a 35 fotocopias es

1) p = 43.75
2) p = 42
3) x = 19.75

La interpretación de la variable en la relación funcional cuando la información se presenta en forma de tabla, es un ejercicio que se hace con poca frecuencia en la instrucción. En este ítem se requiere que el estudiante sea capaz de conceptualizar la *variable en una relación funcional* en *nivel - proceso* ya que ha interiorizado la acción de relacionar cada pareja de la tabla con la variable dependiente que corresponde a la independiente; la interiorización de esta acción le permite interpolar y calcular los valores de la variable dependiente que no están contenidos en la tabla a partir de la información de la relación entre las variables dada por los valores contenidos en ella. Un estudiante que se encuentra en *nivel - acción* tiene dificultades para abstraer esta información, por lo que recurre a la inspección para determinar los valores, e incluso puede tener dificultades para representar el valor de la variable que se le pide de manera congruente con la información dada.

ÍTEM 19

19 Usando los datos del problema anterior. La regla general que describe el precio de las fotocopias usando **n** para representar el número de fotocopias es

1) $p = 1.25n$
2) $n = 1.25p$
3) $5n = 6.25$

La respuesta correcta a este ítem no sólo requiere, como en el caso del ítem 18, de la interpretación de la *variable en la relación funcional* a partir de los datos de una tabla, sino además, que el estudiante sea capaz de interiorizar el proceso mismo de relación para generalizarlo en una fórmula que la simbolice. Los estudiantes que han interiorizado el proceso requerido por la relación y que son capaces de simbolizarla correctamente se encuentran en *nivel - proceso* en su conceptualización del concepto *variable en la relación funcional*. La imposibilidad de lograrlo implica la falta de encapsulación del proceso correspondiente, por lo que puede decirse que los estudiantes se encuentran en *nivel - proceso* con respecto a su comprensión de la relación funcional entre variables.

ÍTEM 20

20. Observa la siguiente tabla, que está incompleta:

Segundos	Velocidad
0	0 m/s
10	30 m/s
15	
20	60 m/s
35	
50	
60	

La regla general que asocia la velocidad al tiempo es:

1) v = 15t
2) 10s = 30v
3) d = v / t

La respuesta correcta a este ítem requiere de la capacidad de los estudiantes de interiorizar el proceso de la relación que se presenta en la tabla, para generalizarlo en una fórmula que la simbolice. Los estudiantes que han interiorizado el proceso requerido por la relación y que son capaces de simbolizarla correctamente se encuentran en *nivel - proceso* en su conceptualización del concepto de *variable en la relación funcional.* La imposibilidad de lograrlo implica la falta de encapsulación del proceso correspondiente, por lo que puede decirse que los estudiantes se encuentran en *nivel - proceso* con respecto a su comprensión de la relación funcional entre variables. A diferencia del ítem 19, este ítem incluye una relación entre variables con las cuales los estudiantes están menos familiarizados.

ÍTEM 21

Considera la siguiente expresión y=3+x.

30. Si queremos que los valores de **y** sean mayores que 3 pero más pequeños que 10, los valores que puede tomar **x** son:
    1) de 0 a 7
    2) 0, 1, 2, 3, 4, 5, 6, 7
    3) 0 y 7

Este ítem se refiere nuevamente a la relación funcional entre dos variables, dada a través de su representación analítica. La respuesta correcta a este ítem indica una conceptualización de la *variable en la relación funcional* como objeto ya que implica la desencapsulación del objeto función en el proceso que le dio origen para poder identificar el intervalo de variación de la variable independiente cuando se conoce el de la dependiente. Este proceso de reversión es posible únicamente cuando se ha logrado una conceptualización de la *variable en la relación funcional* en *nivel – objeto* lo que implica la noción del continuo en los números reales. Los estudiantes que se encuentran en *nivel - proceso* son capaces de llevar a cabo el proceso de reversión únicamente para números específicos y muestran incapacidad de generalizar al continuo numérico.

ÍTEM 22

31. Si **x** toma valores entre 8 y 15, los valores de **y** estarán entre:
    1) entre 11 y 18
    2) 11, 12, 13, 14, 15, 16, 17, 18
    3) 11 y 18

La respuesta correcta a esta pregunta involucra la interiorización de la noción de variación, es decir, del hecho de que en la relación funcional la variable independiente puede tomar un número infinito de valores y que, correspondiente a cada uno de ellos, la variable dependiente tomará otros valores. Los estudiantes que manejan la idea de variable en *nivel - proceso* son capaces de pensar en dicha relación como una regla que asigna a un valor dado de la variable dependiente un valor de la independiente, pero al no haber interiorizado la noción de variación continua de la variable asignan únicamente valores enteros a la variable independiente y, como consecuencia, obtienen únicamente ciertos valores para la variable dependiente. La interiorización de la noción de variación continua permite al alumno pensar en la variable de una manera generalizada de tal suerte que pueden conceptualizarla de una manera dinámica. Esto puede explicarse en términos de la interiorización del proceso de mover la variable sobre la recta de los números reales o sobre un intervalo de éstos y obtener como resultado de la relación funcional la variación continua correspondiente a la variable dependiente.

ÍTEM 23

32. De las siguientes dos expresiones $n + 2$ y $2 \times n$, la más grande es:

    1) depende del valor de n
    2) 2xn
    3) 2 + n

La comparación de expresiones en las que aparece la *variable como un número general* no es problema para un estudiante que se encuentra en *nivel - proceso* en el manejo de la *variable como número general* pues ha interiorizado la acción de evaluar la variable para distintas cantidades. Este estudiante es capaz de tomar decisiones con base en diferentes comparaciones. Aquellos estudiantes que no están en este nivel de conocimiento tienen dificultades para llevar a cabo esta comparación. Pueden, por ejemplo, considerar únicamente algunos valores que la variable puede tomar y generalizar a partir de ellos, es decir, no son capaces de pensar en todo el rango de valores que la variable puede tomar.

ÍTEM 24

33. El peso de la mercancía que se compra en el mercado se mide con una báscula. En el puesto de Don Panchito, por cada kilogramo de peso la charola de la báscula se desplaza 4 cm. La relación entre el peso de la compra y el desplazamiento de la charola está dada por:

1) x = 4p

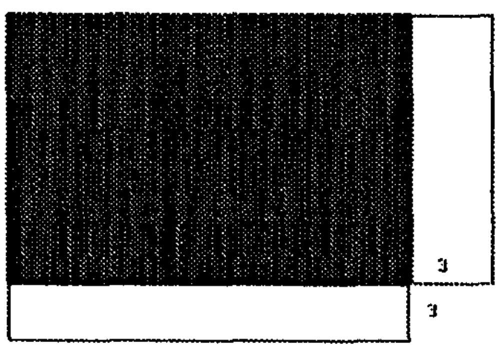
2) 1kg = 4 cm

3) 4 = x

La respuesta incorrecta de esta pregunta implica el uso del signo de igualdad como un recurso mental en el proceso de encadenar acciones. Los estudiantes que utilizan este recurso son capaces de interpretar la relación entre las variables pero no la pueden formalizar. Para responder correctamente a esta pregunta es necesario conceptualizar a *la variable en la relación funcional* en *nivel - objeto*; es decir, ser capaz de interpretar la relación entre las variable y simbolizarla de manera adecuada en cualquier tipo de problema.

ÍTEM 25

34. El área total de la siguiente figura es 27. La fórmula para calcular el lado del cuadrado sombreado es:



1) $x^2 + 6x = 27$

2) $a^2 + b^2 = 27$

3) $L = \sqrt{27}$

La respuesta correcta a esta pregunta implica la comprensión de la *variable como incógnita* como un objeto matemático pues requiere de la posibilidad de identificar en la figura cuál es la variable que se desea encontrar y simbolizarla correctamente. Los estudiantes que responden incorrectamente no han interiorizado la interpretación de la *variable como incógnita*, de tal manera que aún en aquellos casos en los que se encuentran en *nivel - proceso* en su conceptualización e interpretan correctamente el problema, identificando la variable deseada, tienen dificultades para relacionarla con los datos del problema, incorporados en la figura en cuestión

ÍTEM 26

35. Juan es 15 años mayor que Santiago. La suma de las dos edades es 41. La ecuación que permite encontrar las edades de Juan y Santiago es:

1) $j + j + 15 = 41$

2) $s + s + 15 = 41$

3) $j + s = 41$

Los estudiantes que se encuentran en *nivel - proceso* de conceptualización de la *variable como incógnita* pueden interpretar la situación que se presenta en el problema en términos de una ecuación; sin embargo, al no haber interiorizado el proceso de representación simbólica de las ecuaciones suelen confundir la información proporcionada y utilizar la variable de manera incorrecta en la simbolización de la ecuación. Cuando este proceso se ha interiorizado este problema se supera y el estudiante puede representar la situación en términos de una relación correcta entre las variables involucradas.

ÍTEM 27

36. Rentar un automóvil cuesta $25por día, más $0.12 por kilómetro. ¿Cuántos kilómetros puede manejar Diego en un día, si sólo dispone de $40?

1) 125

2)15

3) 12

Este ítem requiere de la interpretación y manipulación de la *variable como incógnita*. Los estudiantes que se encuentran en *nivel – proceso* en la conceptualización de la *variable como incógnita* pueden interpretar y simbolizar la situación que se presenta en el problema en términos de una ecuación y; a través de un proceso de manipulación o de asignación de valores a la variable que en él aparecen, responder mediante un número que parezca razonable. Pero como su capacidad de simbolización es limitada pueden con facilidad confundir la información proporcionada y utilizar la variable de manera incorrecta en la manipulación de la ecuación. Cuando este proceso se ha interiorizado este problema se supera y el estudiante puede representar la situación en términos de una relación correcta entre las variables involucradas, manipularla mediante las transformaciones pertinentes y encontrar el valor de la incógnita; lo que significa que se ha llegado a una conceptualización de la *variable como incógnita* en *nivel - objeto.*

ÍTEM 28

37. Observa los datos siguientes:

x	y
0	0
10	100
-15	225
25	625

20	400
-10	100
15	225
-20	400

Si **x** toma valores entre -2 y 26 ¿entre qué valores estará **y**?

1) entre 0 y 676
2) entre 4 y 676
3) entre –4 y 676

La respuesta correcta a este ítem requiere de la interpretación de los datos de una tabla que representa una relación funcional cuadrática. La interpretación correcta de esta información y su generalización para asignar a la variable dependiente todos los valores que puede tomar dados los límites del intervalo correspondiente a la variable independiente requiere de la conceptualización de la *variable en la relación funcional* en *nivel - objeto* en el que se ha encapsulado el proceso de relación entre las variables correspondientes y se ha superado el proceso de generalización de la relación lineal a cualquier tipo de relación representada mediante una tabla. Los estudiantes que se encuentran en *nivel - proceso* tienden a acudir a una relación lineal siempre que enfrentan datos en una tabla.

ÍTEM 29

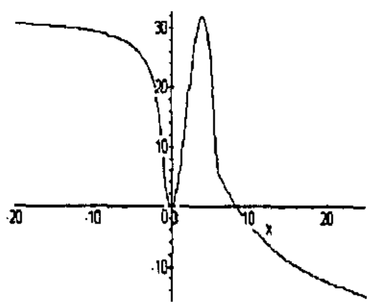
38. Dada la expresión 40 - 15x - 3y = 17y - 5x, el valor de **y** para **x** = 16 es:

1) 1y = -6
2) y = 10
3) y = 60

La solución de este ítem implica la ejecución de varias acciones de manipulación: sustitución, agrupación de términos semejantes y transposición de lás variables en una relación funcional. Esto representa una dificultad grande para los estudiantes que se encuentran en *nivel - acción* en su concepción de la *variable en la relación funcional,* ya que no han interiorizado la ejecución de dichas acciones en un proceso y le ejecución de la secuencia de varias de ellas les resulta complicado y los conduce a errores. Cuando un estudiante ha interiorizado estas acciones en un proceso, la ejecución de la secuencia de acciones no representa más un problema.

ÍTEM 30

30. Dada la siguiente gráfica:



¿Entre qué valores de **x**, los valores de **y** decrecen?

1) entre - ∞ y 0 y entre 5 y ∞
2) entre ∞ y 5 y entre 5 y ∞
3) entre 10 y 20

La lectura de datos a partir de una gráfica representa dificultades para muchos estudiantes. En el caso de este ítem se requiere la interpretación de la gráfica de forma dinámica en términos de intervalos de variación. Los estudiantes que se encuentran en *nivel - acción* en la conceptualización de la *variable en la* *relación funcional* pueden asignar valores a puntos específicos de la gráfica, sin embargo no han interiorizado esta acción en el proceso de asignación de valores de manera dinámica en un intervalo. Los estudiantes que se encuentran en *nivel - proceso* han interiorizado la acción de lectura punto a punto de la gráfica y son capaces de interpretar la variación de las variables expresadas mediante esta representación

Aplicación del modelo de medida a los ítems

A partir de una aplicación del cuestionario abierto de 65 preguntas a una muestra de 100 alumnos del primer año de escuela secundaria, que no han cursado estudios de álgebra se calcularon los valores que se usaron para el parámetro k_2 del modelo, que representa la probabilidad máxima de elección de la opción dos de cada una de las preguntas seleccionadas para el cuestionario de prueba, considerando a este parámetro como la proporción máxima de elección de dicha opción para cada ítem. El cálculo de este valor permite reducir el número de parámetros a estimar cuando se ajusta el modelo logístico para la segunda opción de respuesta. El valor del parámetro k_1 se consideró siempre igual a uno, ya que es de esperarse que un sujeto que domina el concepto seleccione con probabilidad igual a uno la opción correcta para cada una de las preguntas del cuestionario.

El cuestionario de prueba, que se presenta en esta sección, fue aplicado a una muestra de 284 alumnos que inician los estudios de licenciatura en los programas de Ciencias Sociales, Administración de Empresas, Relaciones Internacionales y Contabilidad de una universidad mexicana (Instituto Tecnológico Autónomo de México). El cuestionario en formato abierto se aplicó a una muestra comparable de 164 estudiantes con el fin de analizar las respuestas más frecuentes, compararlas con los resultados reportados en la literatura y utilizarlas en el diseño de las opciones del cuestionario de opción múltiple.

La estimación de los parámetros se llevó a cabo utlizando el mátodo descrito en la sección II.7 desl presente trabajo. La metodología seguida se describe a continuación.

En una primera fase se tomó como punto de partida el programa SYSTAT que permite estimar los parámetros de dificultad y de discriminación de los ítems de del cuestionario, así como los parámetros de hablidad de los sujetos que lo responden. El número de iteraciones necesarias para calcular los parámetros fue de 73. Estos parámetros utilizaron en la primera iteración para calcular el parámetro de accesibilidad de la opción de respuesta correcta de cada uno de los ítems. Manteniendo fijos los parámetros de la habilidad se hizo una modificación del programa SYSTAT que permitió ajustar una curva logística a la opción incorrecta de cada ítem y obtener de ahí los valores para los parámteros de accesibilidad de dichas opciones. La modificación introducida consiste en lograr que el programa admita valores negativos para los parámetros de discriminación y permita que el valor máximo de la probabilidad de elección sea menor que uno. El número de iteraciones necesarias para lograr la convergencia en este caso fue de 71.

Los valores obtenidos de este procedimiento se utilizaron como datos de entrada para un programa que ajusta los parámetros de accesibilidad, los de interacción, las condiciones iniciales y los valores correspondentes a la variable independiente. En este programa se utilizaron versiones modifcadas de las rutinas escritas en FORTRAN desarrolladas por Byrd, Lu, Nocedal y Zhu (1995), en las que se se utilizan un método cuasi-Newton y los métodos de gradiente conjugado y de Newton truncado desarrollados por Gómez y Morales (1997), además de incluirse una subrutina para resolver numéricamente el sistema de ecuaciones diferenciales. Estos paquetes se corrieron utilizando el acceso vía red electrónica al sitio NEOS en el que se utiliza de forma remota una computadora de alto poder. En este caso fueron necesarias 243 iteraciones de las dos fases del programa para obtener la convergencia deseada de 10^{-5} . Los valores finales de los parámetros se reportan a continuación.

El tamaño de la muestra resultó conveniente para utilizar el método descrito de tal forma que la mayor parte de los parámetros convergieran. Únicamente en el caso de los ítems 17, 25 y 28 no se logró la convergencia deseada.

Es importante mencionar que el método de ajuste utilizado en esta ocasión no es el único posible y que sería interesante hacer investigación para encontrar métodos de ajuste de los parámetros más efectivos y eficientes.

Los valores obtenidos para los parámetros de cada una de las opciones de elección de respuesta de los ítems del cuestionario de prueba se muestran en la siguiente tabla. Es conveniente recordar que los parámetros a_1 y a_2 representan la accesibilidad de cada una de las opciones; los parámetros $\boxed{X}$ y b_2 representan la interacción que ejerce una opción de respuesta sobre la probabilidad de elección de la otra; los parámetros k_1 y k_2 representan a máxima probabilidad de elección de las opciones de respuesta 1 y 2 respectivamente y *c.i.* representa el valor obtenido para las condiciones iniciales en el ajuste de los parámetros con la muestra que respondió los ítems.

TABLA III.6.1

Parámetros de los Items

Ítem	a_1	b_1	k_1	a_2	b_2	k_2	*c. i.*
1	1.21	1.52	1	-0.5	0.85	0.72	(.2.6)
2	0.69	0.98	1	-0.231	0.985	0.70	(.103.677)
3	0.21	1.27	1	-0.45	1.5	0.74	(.29.66)
4	0.33	1.17	1	-0.54	1.15	0.70	(.351.55)
5	0.66	0.931	1	-0.97	0.535	0.77	(.265.477)
6	0.48	-0.25	1	-0.41	0.45	0.83	(.13.78)
7	0.678	0.508	1	-0.745	0.927	0.87	(.144.849)
8	0.874	1.23	1	-0.521	0.789	0.71	(.29.64)
9	0.734	0.891	1	-0.57	1.01	0.84	(.124.789)
10	1.006	1.237	1	-0.574	0.267	0.91	(.257.684)
11	1.03	1.151	1	-0.223	0.33	0.64	(.34.641)
12	0.442	0.571	1	-0.487	0.871	0.84	(.189.76)
13	0.248	0.304	1	-0.247	0.785	0.84	(.089.68)
14	0.74	1.002	1	-0.22	0.998	0.74	(.18.68)
15	0.421	0.827	1	-0.643	0.372	0.52	(.363.561)
16	0.342	1.27	1	-0.41	0.665	0.92	(.13.56)
17	X	X	X	X	X	X	X
18	0.539	0.799	1	-0.745	0.32	0.68	(.41.526)
19	0.546	0.563	1	-0.778	1.278	0.92	(.03.886)
20	0.932	0.998	1	-0.436	0.987	0.94	(.062.91)

21	0 .431	1.22	1	-0.401	1.141	0.75	(.312.52)
22	0.884	1.003	1	-0.374	0.999	0.88	(.145.803)
23	1.29	1.455	1	-0.761	0.777	0.73	(.14.68)
24	0.651	1.627	1	-0.73	1.12	0.87	(.02.813)
25	X	X	X	X	X	X	X
26	0.679	1.01	1	-0.792	0.976	0.69	(.066.67)
27	0.678	-1.125	1	-0.941	0.845	0.858	(.133.544)
28	X	X	X	X	X	X	X
29	1.349	1.41	1	-0.363	0.454	0.944	(.08.89)
30	1.022	1.27	1	-0.651	0.51	0.82	(.17.75)

Haciendo un análisis de los parámetros de los ítems que se muestran en la tabla en términos de las relaciones con las cuales se definieron los diferentes casos que se podrían presentar en el modelo de medida, encontramos los siguientes resultados:

Items correspondientes al CASO 6 del Modelo de Medida

El caso 6 del modelo de medida es aquél que corresponde a aquellos ítems para los cuales el parámetro a_2 es negativo y que son tales que a_1/b_1 es mayor que k_2.

TABLA III.6.2

Ítems correspondientes al caso 6 del modelo de medida

Ítems	1	2	7	11	13	19	20	22	23	29

Items correspondientes al CASO 7 del Modelo de Medida

Los parámetros correspondientes al caso 7 del modelo de medida propuesto en este trabajo son aquéllos para los cuales el parámetro a_2 es negativo y que son tales que a_1/b_1 es menor que k_2.

TABLA III.6.3

Ítems correspondientes al caso 7 del modelo de medida

Ítems	3	4	5	9	10	12	14	15	16	18	21	24	30

Items correspondientes al CASO 8 del Modelo de Medida

Los parámetros correspondientes al caso 8 del modelo de medida propuesto en este trabajo son aquéllos para los cuales el parámetro a_2 es negativo y que son tales que a_1/b_1 es igual a k_2.

TABLA III.6.4

Ítems correspondientes al caso 8 del modelo de medida

Ítems	8	14

En la tabla III.6.1 puede notarse, además, que se encontraron dos ítems, los ítems 6 y 27 para los cuales el parámetro b_1 es negativo. Este es un indicador de que los distractores de los ítems no están bien diseñados y de que deben ser rediseñados. Por último, los parámetros de los ítems 17, 25 y 29 no convergieron. Esto puede deberse a que la gran mayoría de los estudiantes eligió la opción tres de respuesta o no respondió a los ítems por lo que los datos obtenidos no proveen suficiente información.

A continuación se hace el análisis del los ítems en términos del modelo de medida, conjuntamente con el modelo diagnóstico.

Análisis de los ítems en términos del modelo de medida

Los resultados obtenidos de la estimación de los parámetros y del análisis de las gráficas que se obtienen para cada uno de los ítems del cuestionario pueden ser analizados utilizando conjuntamente el modelo de medida y el modelo diagnóstico discutidos con anterioridad. En esta sección se presenta el análisis y las gráficas para cada uno de los ítems del cuestionario de prueba. El orden seguido en el análisis de los ítems corresponde a su clasificación de acuerdo a los diferentes casos del modelo, para facilitar su comparación. Conviene aclarar aquí que el análisis de los ítems se hace únicamente con respecto a los dos primeros distractores para hacer más simple la estimación de los parámetros; la tercera opción se introdujo en el cuesionario a partir de respuestas frecuentes de los estudiantes, con el fin de no dejar muy abierta la posibilidad de no responder al ítem.

Items correspondientes al caso 6

Análisis del ítem 1

Los parámetros para este ítem lo sitúan entre las preguntas correspondientes al caso 6 que se refiere a aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es mayor que k_2. La razón a_1/b_1 puede considerarse como la relación que hay entre la accesibilidad de la opción 1 a los ojos del sujeto y la interacción que representa la presencia de la otra opción, es decir, puede considerarse como una medida de qué tan atractiva resulta la opción correcta para el sujeto que responde al ítem. Puesto que b_1 en este caso es mayor que 1, la interacción que la otra opción presenta es grande y dado que a_1 es mayor que 1, puede considerarse que la opción correcta es accesible para el sujeto. Esto significa que en esta pregunta la opción correcta es fácil de reconocer a pesar de la interacción. Por otra parte, las condiciones iniciales se encuentran en la región del plano fase que se indica con la letra A en la figura III.6.1a, en la que no es posible reconocer el intervalo de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto.* Las gráficas que muestran la solución correspondiente a este ítem, tanto en el espacio de configuración, se muestran en la figura III.6.1b. En ella puede observarse que para el valor θ = 1.5 las curvas de probabilidad de elección de las dos opciones de respuesta en estudio se cruzan, y la probabilidad de elegir la opción correcta domina sobre la incorrecta; puede considerarse, entonces, que aquéllos sujetos cuyo valor de θ es mayor que 1.5 son quienes responderán correctamente esta pregunta.

Figura III.6.1a: Regiones en plano fase

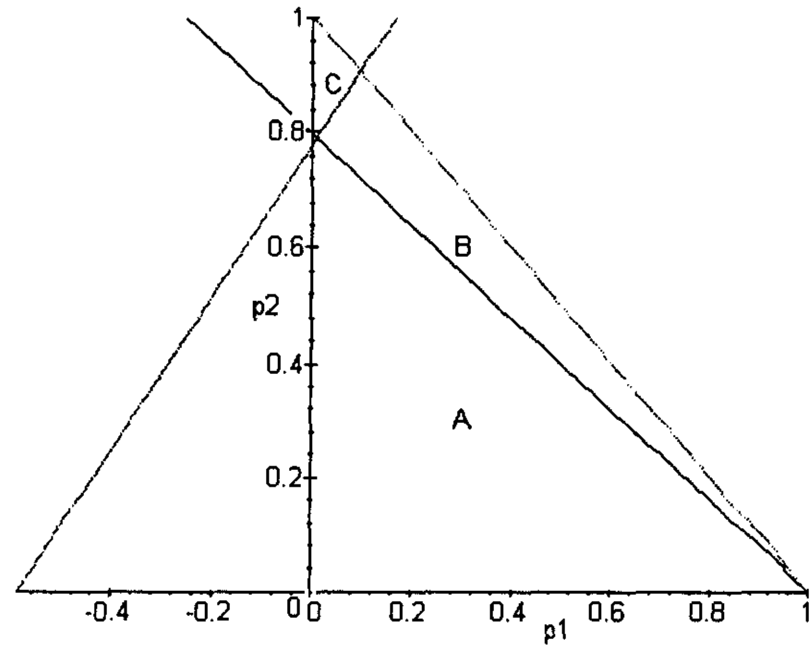
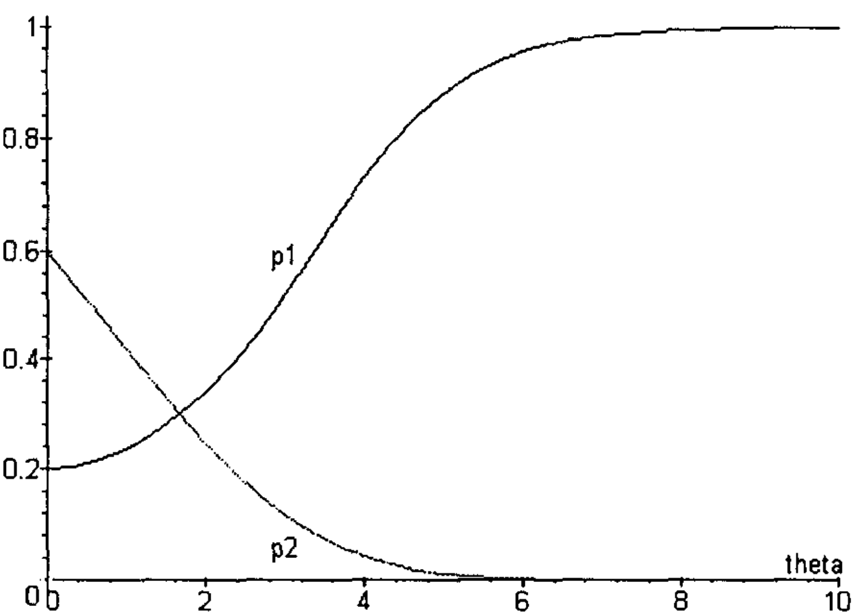


Figura III.6.1b: Curvas solucion



Análisis del ítem 2

Los parámetros de este ítem lo sitúan entre las preguntas correspondientes al caso 6, para los que a_2 es negativa y que son tales que a_1/b_1 es mayor que k_2. Como ya se mencionó, la razón a_1/b_1 puede considerarse como la relación que hay entre la accesibilidad de la opción 1 a los ojos del sujeto y la competencia que representa la presencia de la otra opción, es decir, puede considerarse como una medida de qué tan atractiva resulta la opción correcta para el sujeto que responde al ítem. Puesto que b_1 en este caso es casi igual a 1, las opciones de respuesta resultan similares para el sujeto que responde y dado que a_1 es igual a 0.68, puede considerarse que la opción correcta no es muy accesible para el sujeto. Esto implica que en esta pregunta la opción correcta no es fácil de reconocer. Además la interacción juega un papel importante. Por otra parte, las condiciones iniciales se localizan en la región del espacio fase que se indica con la letra B en la figura III.6.2a, lo cual indica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.2b, permite reconocer el intervalo de la variable independiente (0, 1.43) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto*, en la comprensión del concepto de *variable como número general*. El valor de la variable independiente θ para el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2 es de 5.17, por lo que puede considerarse que los sujetos que obtienen un valor de θ mayor a éste responderán correctamente este ítem.

Figura III.6.2a: Regiones en plano fase

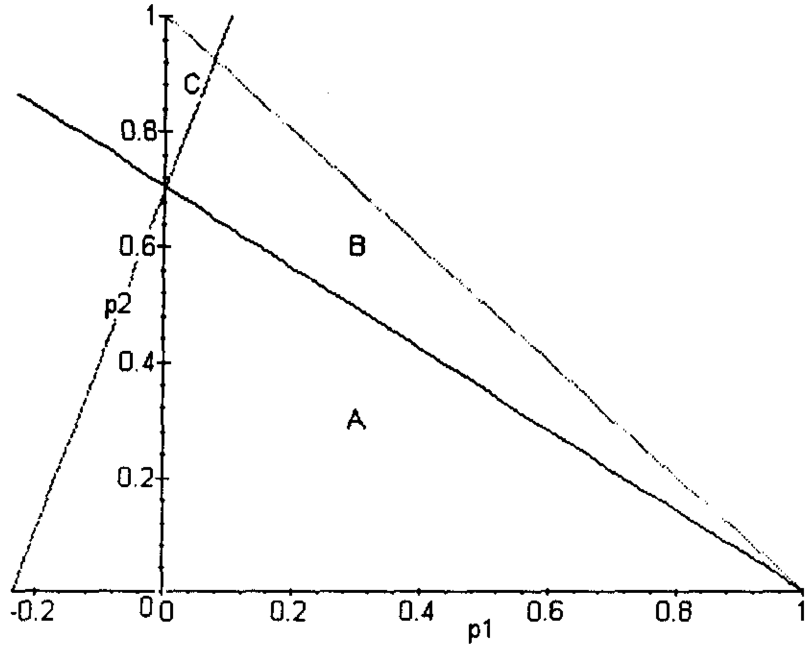
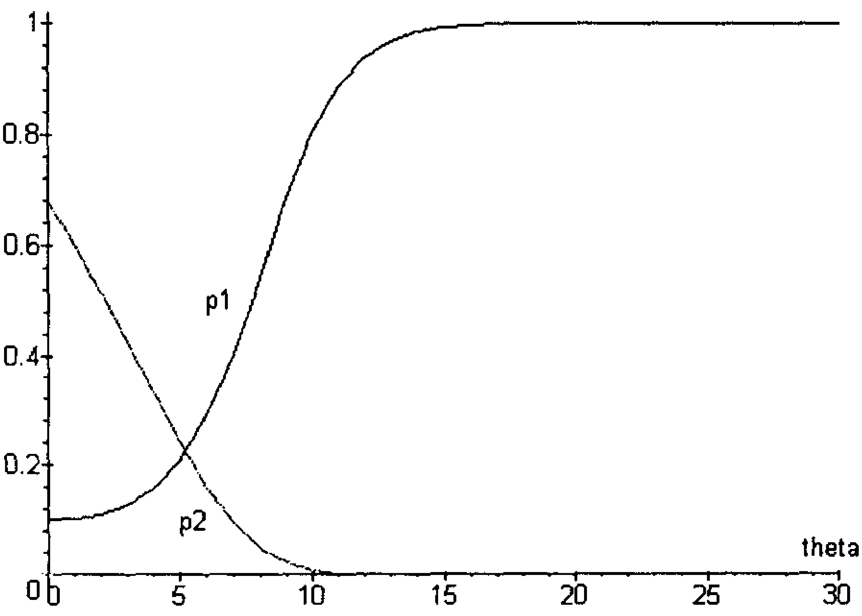


Figura III.6.2b: Curvas solucion



Análisis del ítem 7

Los parámetros para este ítem lo sitúan entre los reactivos correspondientes al caso 6. En este caso a_1 es 0.345, por lo que la respuesta correcta puede considerarse poco accesible. El valor del parámetro b_1 es en este caso menor que uno, por lo que la interacción que la opción 2 ejerce no es muy importante; esto implica que en esta pregunta la opción correcta no es fácil de reconocer pero que la interacción no es fuerte. Además, las condiciones iniciales se encuentran en la región del espacio fase denotada por B en la figura III.6.7a, lo cual indica que este ítem permite reconocer el intervalo de la variable independiente: (0,0.904) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto* en su comprensión de la *variable como incógnita*, esto se muestra también claramente en la Figura III.6.7b. El valor de la variable independiente θ para el que la probabilidad de responder la opción 1 domina sobre la probabilidad de responder la opción 2 es de 3.39, por lo que puede considerarse que los sujetos que obtienen un valor de θ mayor a éste responderán correctamente este ítem.

Figura III.6.7a: Regiones en plano fase

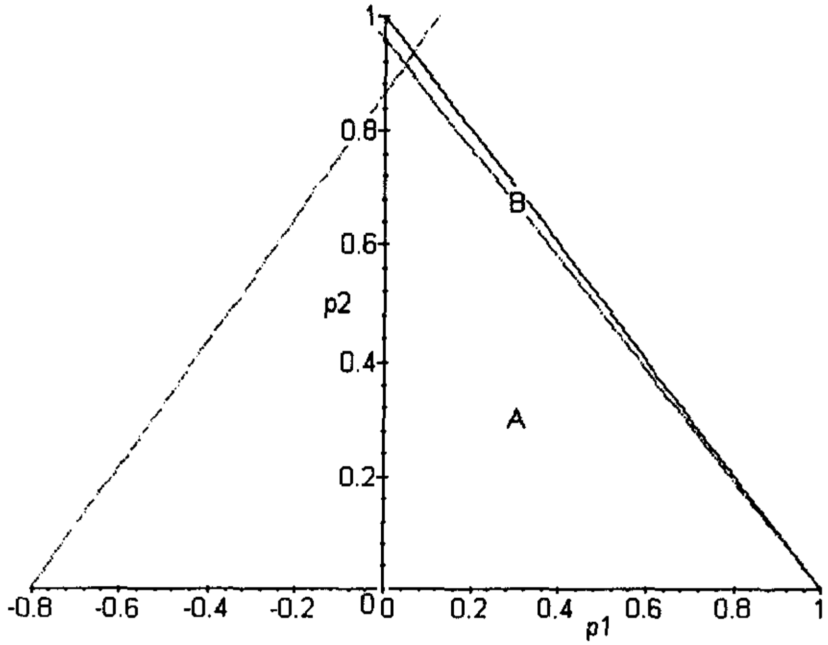
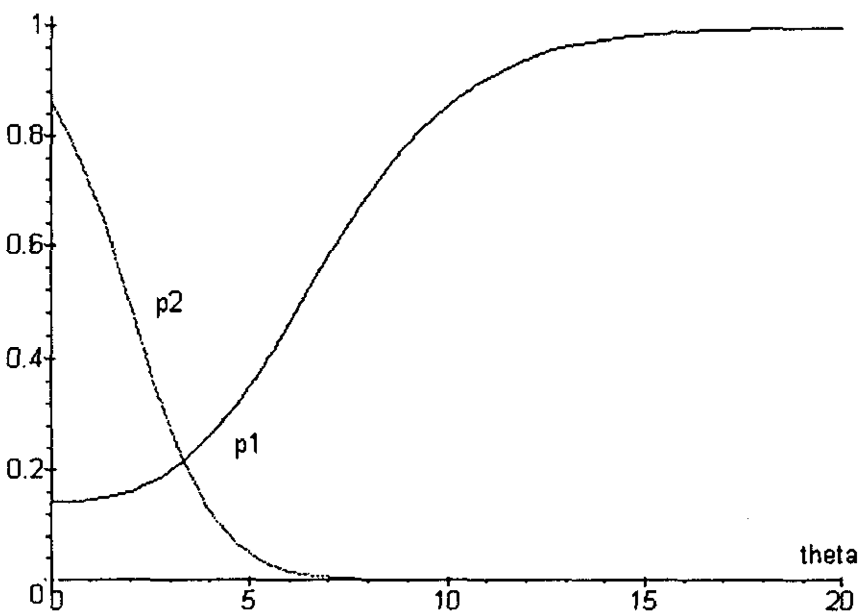


Figura III.6.7b: Curvas solucion



Análisis del ítem 11

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos incluidos en el caso 6. Puesto que b_1 en este caso es mayor que uno, la interacción que la otra opción presenta es grande y puesto que a_1 es mayor a uno, puede considerarse que la opción correcta es accesible para el sujeto; en este caso se tiene que la opción de respuesta correcta es fácil de reconocer a pesar de la interacción. Además, las condiciones iniciales se sitúan en la región del espacio fase indicada por la letra B en la figura III.6.11a, lo cual indica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.11b, permite reconocer el intervalo de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso* en su conceptualización de la *variable como incógnita*. Este corresponde al intervalo (0,1.06). El valor de la variable independiente θ para el que puede considerarse que los sujetos responderán correctamente a este ítem, en el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2 es de 2.92.

Figura III.6.11a: Regiones en plano fase

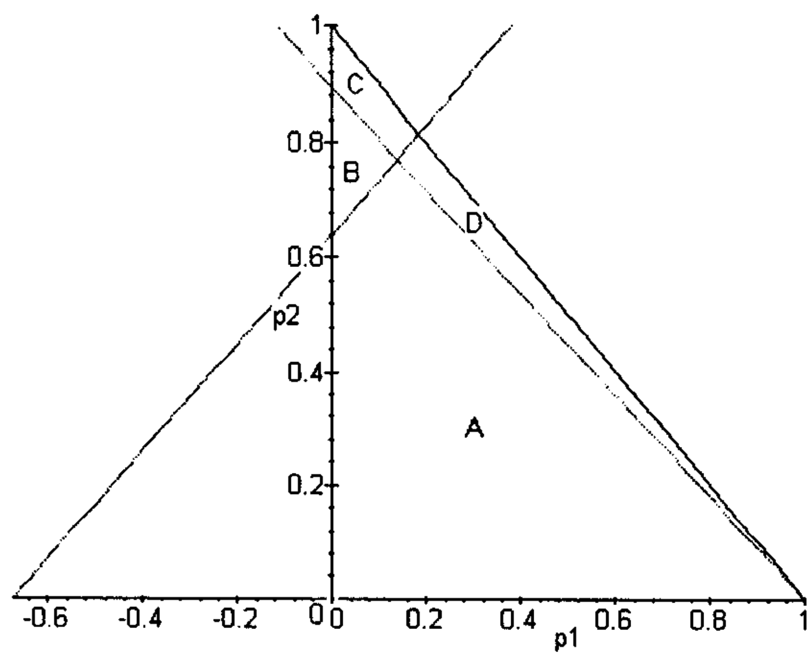
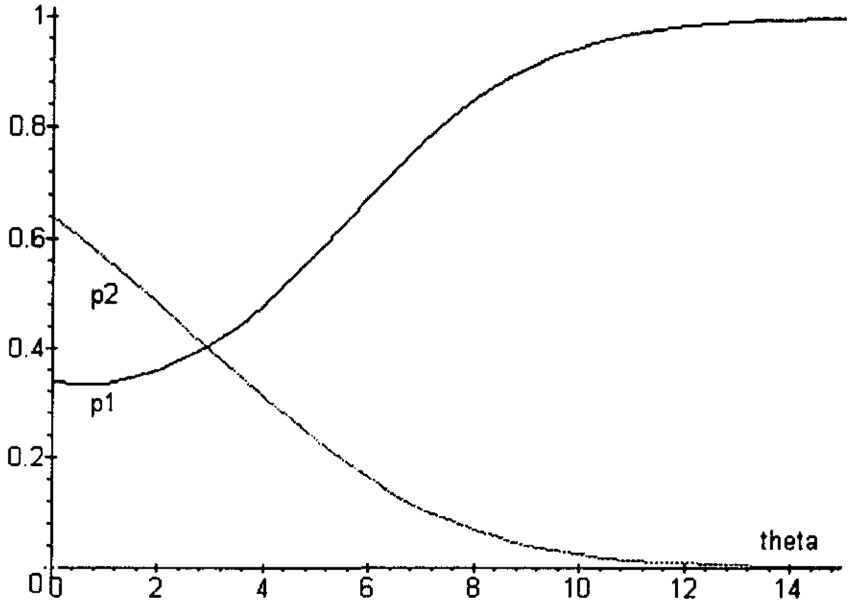


Figura III.6.11b: Curvas solucion



Análisis del ítem 13

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos correspondientes al caso 6. Puesto que a_1 es mucho menor que uno, puede considerarse que la opción correcta es poco accesible para el sujeto, aun cuando b_1 es asimismo menor que uno. La interacción que la opción 2 presenta no es muy fuerte, lo que implica que en esta pregunta la opción correcta es difícil de reconocer a pesar de que la interacción no es grande. Las condiciones iniciales indican que la región del espacio fase para este ítem pertenece a la región B de la figura III.6.13a, por lo que este ítem permite reconocer el intervalo de la variable independiente: (0, 1.2) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso,* lo que se muestra también claramente en la Figura III.6.13b. El valor de la variable independiente θ para el que la probabilidad de responder la opción 1 domina sobre la probabilidad de responder la opción 2 es de 6.69, por lo que puede considerarse que los sujetos que obtienen un valor de θ mayor a éste responderán correctamente este ítem.

Figura III.6.13a: Regiones en plano fase

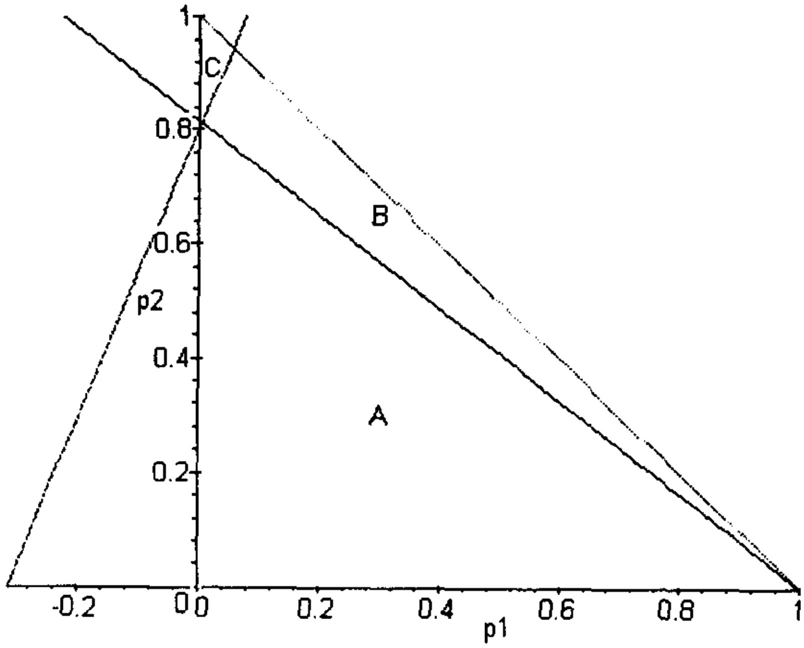
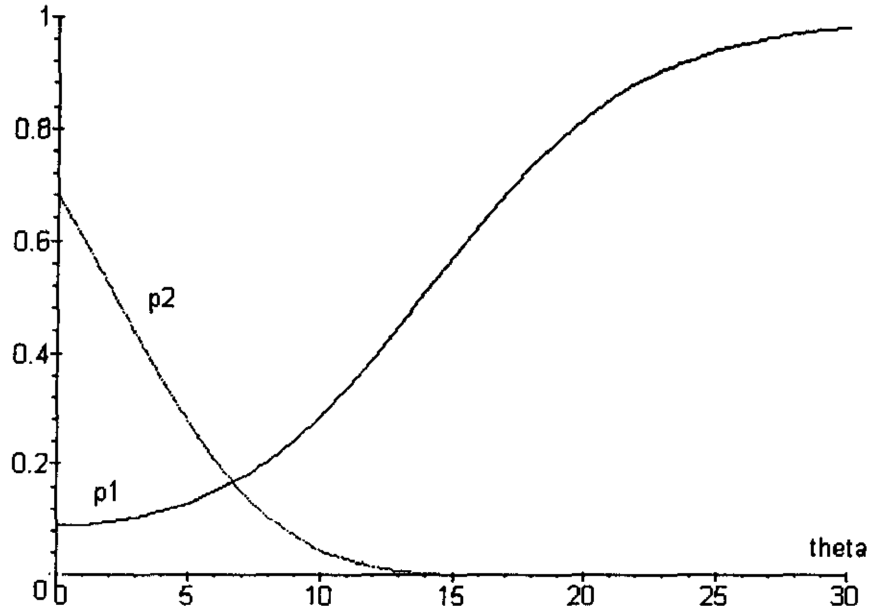


Figura III.6.13b: Curvas solucion



Análisis del ítem 19

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos pertenecientes al caso 6. En este caso a_1 es 0.546, por lo que la respuesta correcta puede considerarse poco accesible. El valor del parámetro b_1 es en este caso menor que uno, por lo que la interacción que la opción 2 ejerce no es muy importante; esto implica que en esta pregunta la opción correcta no es fácil de reconocer a pesar de que la interacción no es fuerte. Además, las condiciones iniciales se encuentran en la región del espacio fase denotada por A en la figura III.6.19a, lo cual indica que este ítem no permite reconocer el intervalo de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto* en su comprensión de la *variable como incógnita*; esto se muestra también claramente en la Figura III.6.19b. El valor de la variable independiente θ para el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2 es de 5.27, por lo que puede considerarse que los sujetos que obtienen un valor de θ mayor a éste responderán correctamente este ítem.

Figura III.6.19a: Regiones en plano fase

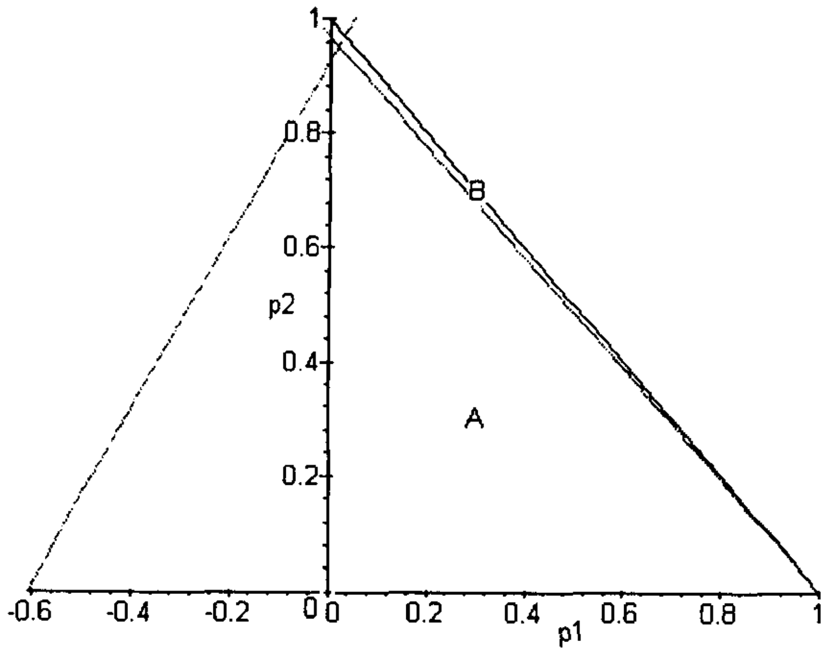
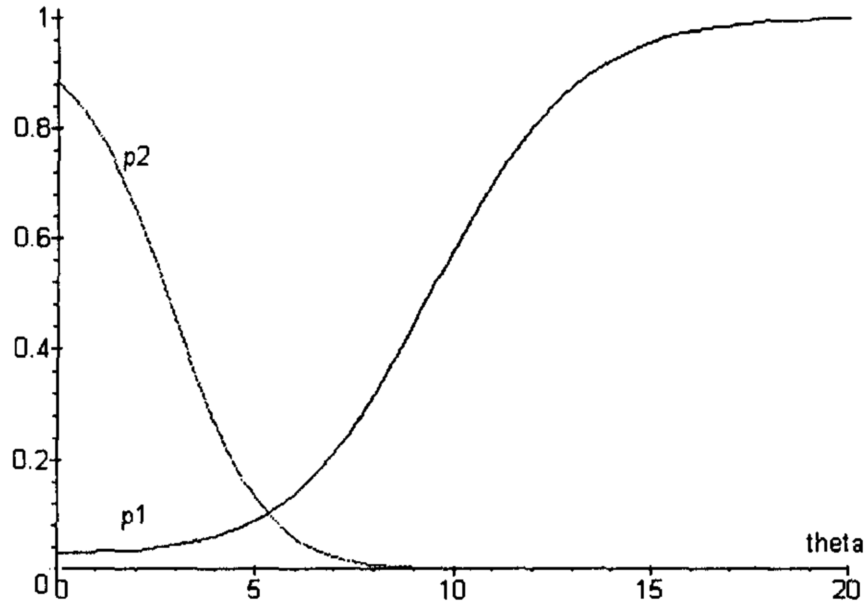


Figura III.6.19b: Curvas solucion



Análisis del ítem 20

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos correspondientes al caso 6. El parámetro a_1 es ligeramente menor a uno, lo que indica que la opción de respuesta correcta no es muy accesible a los sujetos. Puesto que b_1 en este caso es casi igual a uno, las opciones de respuesta resultan similares para el sujeto que responde, por lo cual puede considerarse que en esta pregunta la opción correcta no es demasiado fácil de reconocer y que además la interacción juega un papel importante. Las condiciones iniciales nuevamente se encuentran en la región del espacio fase correspondiente a la letra B de la figura III.6.20a, lo cual indica que la gráfica que muestra las curvas de probabilidad para este ítem (figura III.6.20b), permite reconocer el intervalo de la variable independiente (0, 1.17) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto* en su conceptualización de la *variable en la relación funcional*. El valor de la variable independiente θ para el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2 es de 5.37; es a partir de este valor de la variable independiente que puede considerarse que los estudiantes reponderán correctamente este ítem.

Figura III.6.20a: Regiones en plano fase

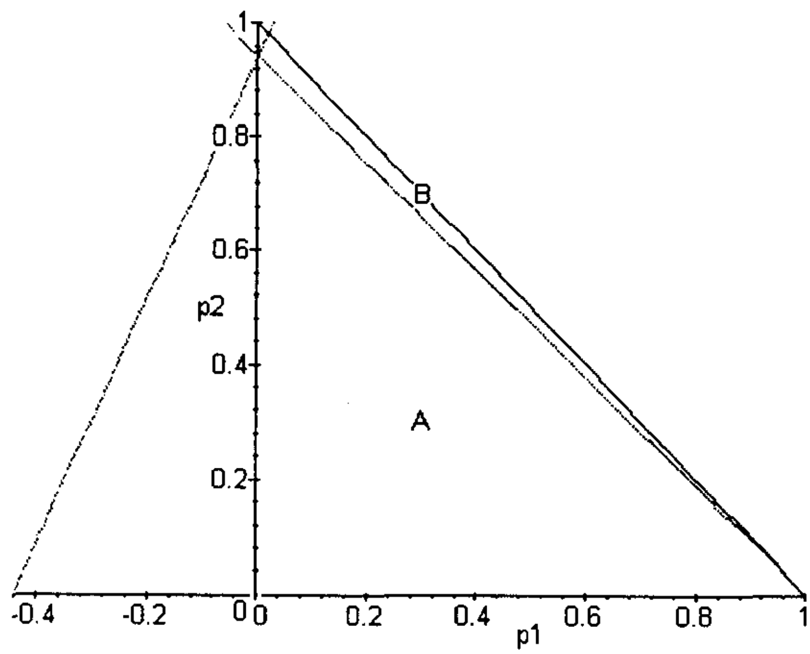
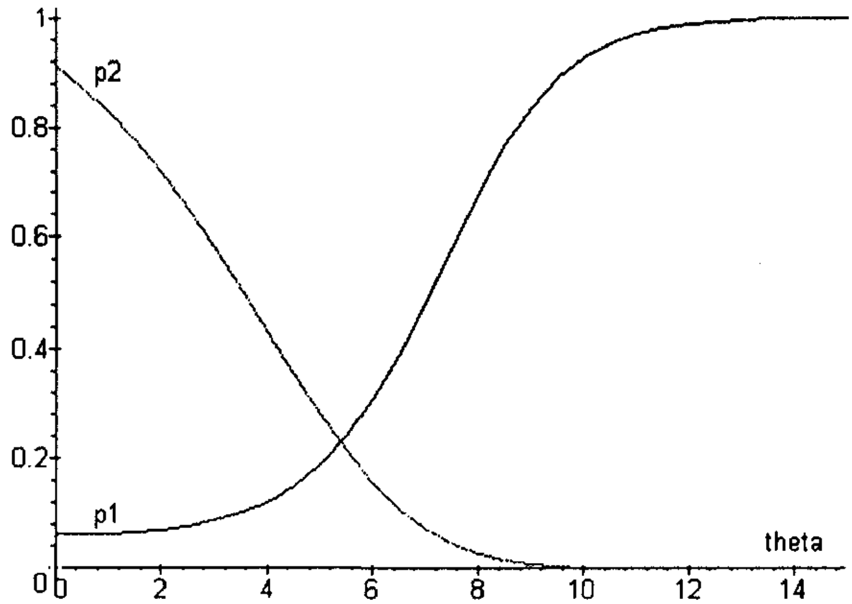


Figura III.6.20b: Curvas solucion



Análisis del ítem 22

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos pertenecientes al caso 6. Puesto que b_1 en este caso es casi igual a 1, las opciones de respuesta resultan similares para el sujeto que responde y puesto que a_1 es cercano a uno, puede considerarse que la opción correcta es medianamente accesible para el sujeto. Como el parámetro de interacción b_1 es casi igual a uno la competencia juega un papel importante ya que para el sujeto que responde ambas opciones parecen igualmente plausibles. Además, las condiciones iniciales se encuentran en la región del espacio fase correspondiente a la indicada con la letra B en la figura III.6.22a, lo cual indica que la gráfica para a este ítem, que se muestra en la figura III.6.22b, permite reconocer el intervalo de la variable independiente (0,0.83) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso.* El valor de la variable independiente θ para el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2, a partir de la cual puede considerarse que los sujetos responderán al ítem correctamente es de 3.35.

Figura III.6.22a: Regiones en plano fase

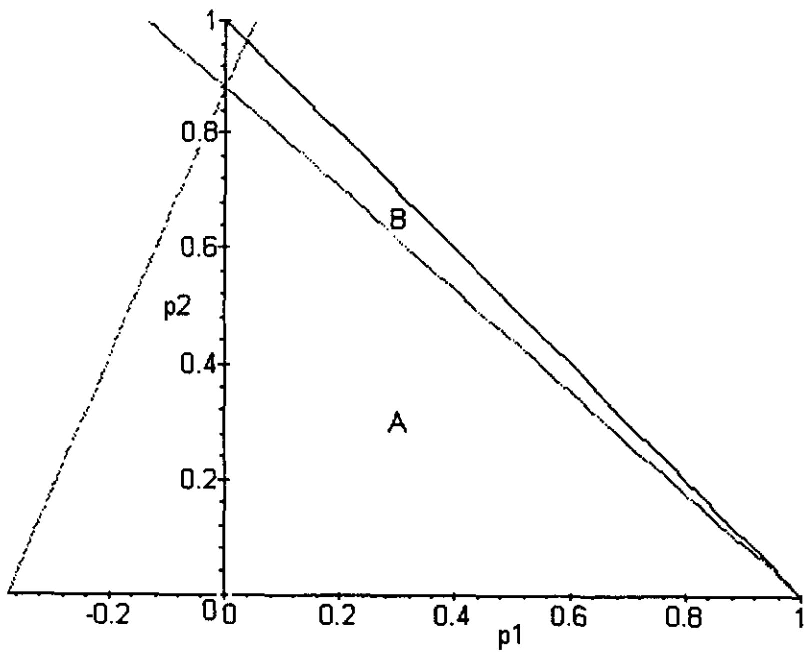
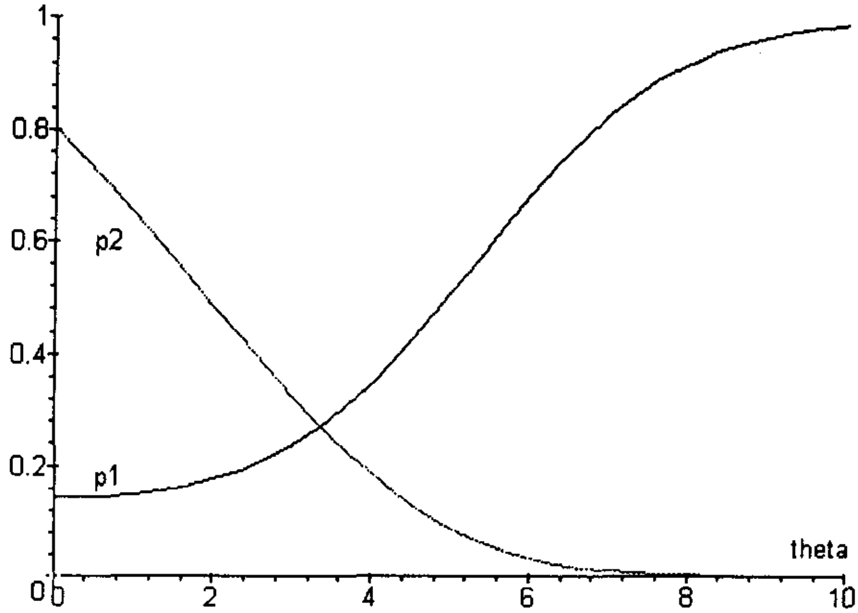


Figura III.6.22b: Curvas solucion



Análisis del ítem 23

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos correspondientes al caso 6. En este caso b_1 es mayor a uno, por lo que puede decirse que la segunda opción de respuesta interactúa fuertemente para el sujeto con la primera opción y puesto que a_1 es mayor a uno, puede considerarse que la opción correcta es accesible para el sujeto. Las condiciones iniciales en este caso se encuentran en la región del espacio fase que se indica con la letra A en la figura III.6.23a, lo cual muestra que la gráfica para este ítem, que se muestra en la figura III.6.23b, no permite reconocer el intervalo de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso* en la comprensión de la *variable como numéro general.* El valor de la variable independiente para el que la probabilidad de elección de la opción de respuesta correcta domina es θ = 2.231.

Figura III.6.23a: Regiones en plano fase

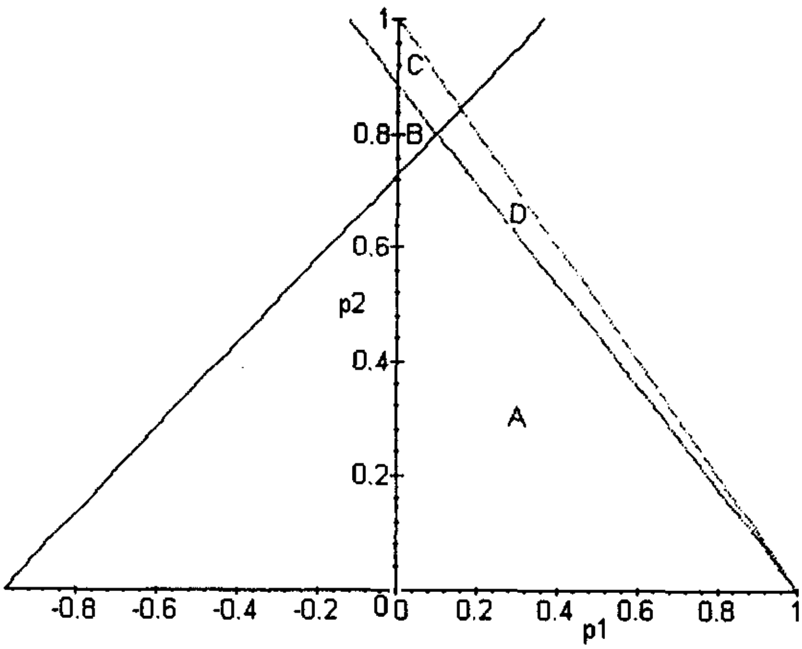
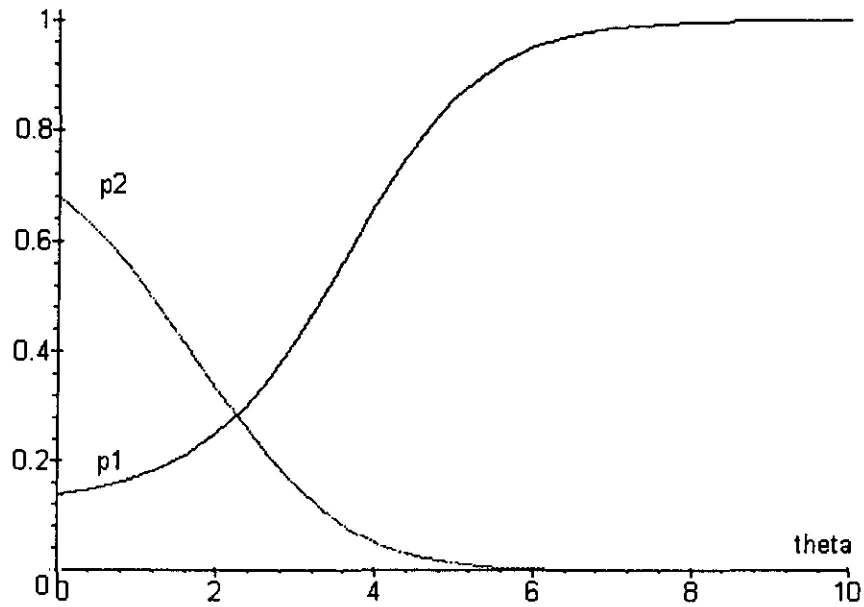


Figura III.6.23b: Curvas solucion



Análisis del ítem 29

Los parámetros de este ítem lo sitúan entre los reactivos incluidos en el caso 6. El parámetro a_1 es mayor que uno lo que indica que la opción correcta es muy accesible para el sujeto que responde al ítem. Puesto que b_1 en este caso es también mayor a uno, la opción 2 interactúa fuertemente con la opción 1. Por otra parte, las condiciones iniciales indican que la región del espacio fase a la que esta pregunta pertenece es la región B de la figura III.6.29a, por lo que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.29b, permite reconocer el intervalo de la variable independiente (0, 1.25) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso.* El valor de la variable independiente θ para el que la probabilidad de elección de la opción 1 domina sobre la probabilidad de elección de la opción 2, a partir de la cual puede considerarse que los sujetos responderán al ítem correctamente es de 5.71.

Figura III.6.29a: Regiones en plano fase

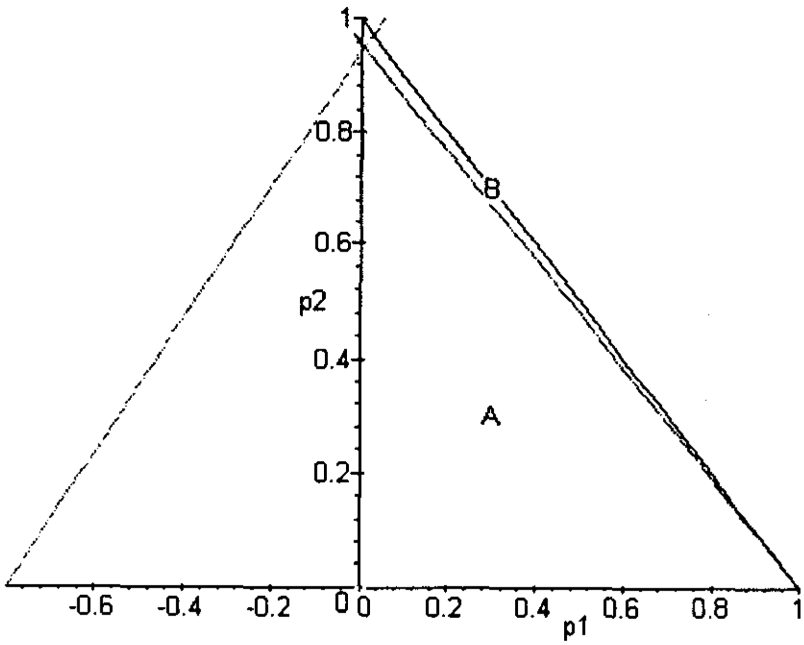
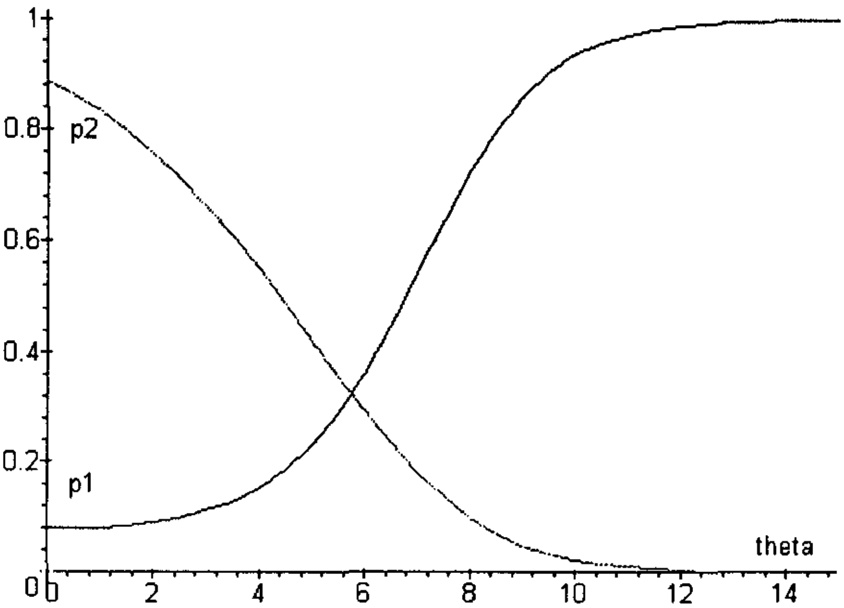


Figura III.6.29b: Curvas solucion



Ítems correspondientes al caso 7

Análisis del ítem 3

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos del caso 7 en los que a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a uno, la opción 2 interactúa fuertemente con la opción 1. Por otra parte, a_1 es muy pequeña lo que indica que la opción correcta no es accesible para el sujeto. Esto implica que, en esta pregunta, la opción correcta no es fácil de reconocer y que además la interacción juega un papel importante. Las condiciones iniciales muestran que la región del espacio fase a la que esta pregunta pertenece es la región B que se presenta en la figura III.6.3a, lo cual indica que la gráfica que describe las probabilidades para este ítem, que se muestra en la figura III.6.3b, permite reconocer el intervalo de la variable independiente (0,3.48) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso.* La probabilidad de elección de la opción correcta domina cuando la variable independiente toma el valor de 3.35; se espera que los sujetos caracterizados por valores de θ mayores a este valor responderán correctamente este ítem.

Figura III.6.3a: Regiones en plano fase

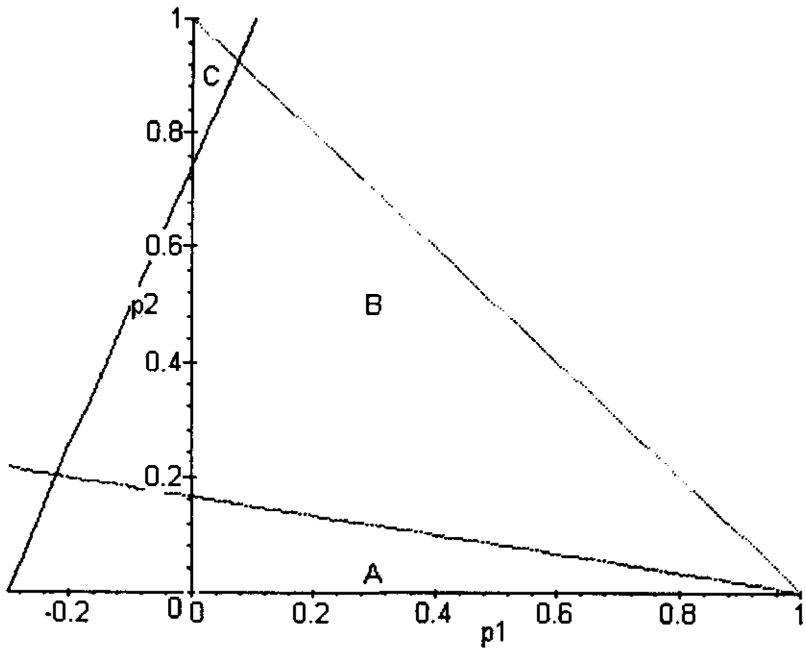
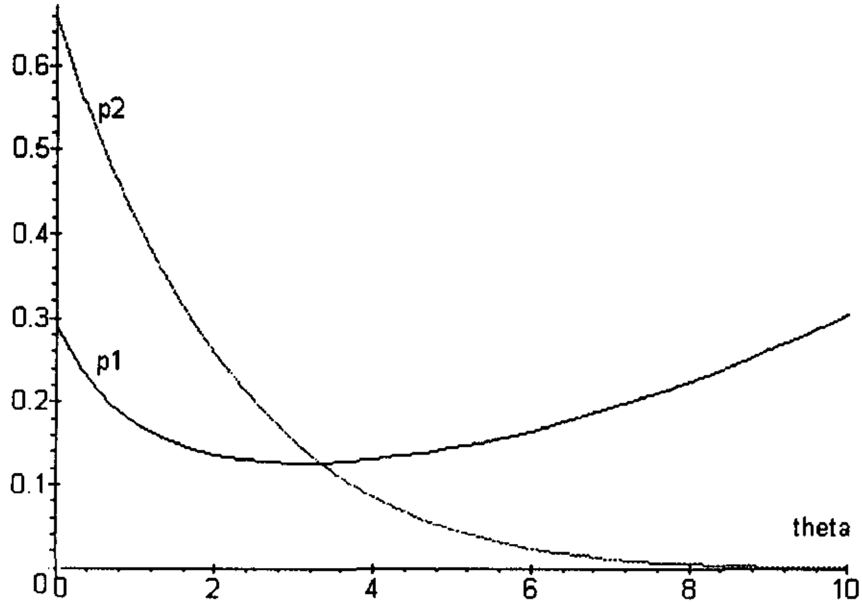


Figura III.6.3b: Curvas solucion



Análisis del ítem 4

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos correspondientes al caso 7 en los que a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a uno, la opción de respuesta incorrecta interactúa fuertemente sobre el sujeto que responde y puesto que a_1 es muy pequeña, la opción correcta resulta difícil de reconocer para el sujeto. Las condiciones iniciales indican que la región del caso 7 a la que esta pregunta pertenece es la región que se marca con la letra B en la figura III.6.4a, razón por la cual la gráfica correspondiente a este ítem, que se muestra en la figura III.6.4b, permite reconocer el intervalo de la variable independiente (0,2.07) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. La probabilidad de elección de la opción correcta domina cuando la variable independiente toma el valor de 1.37; se

Figura III.6.4a: Regiones en plano fase

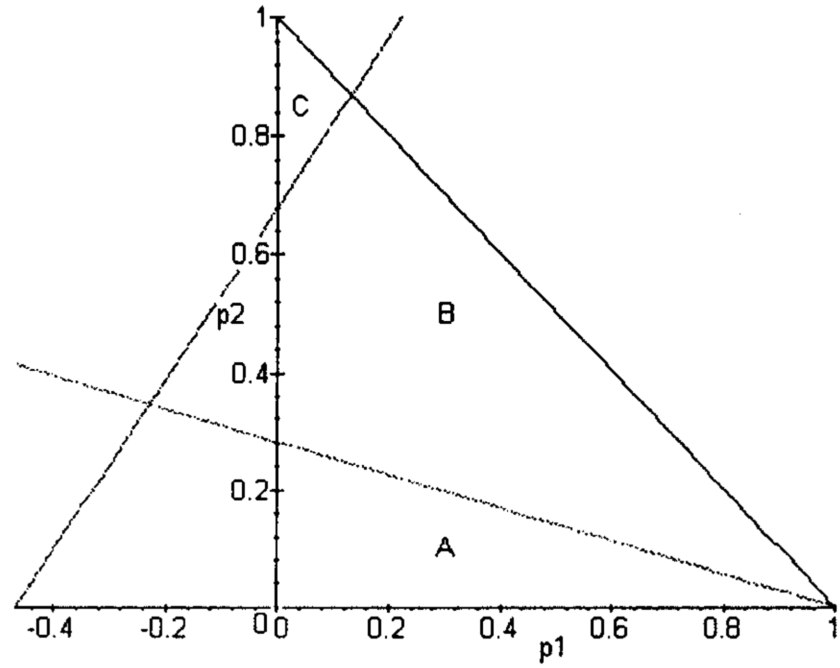
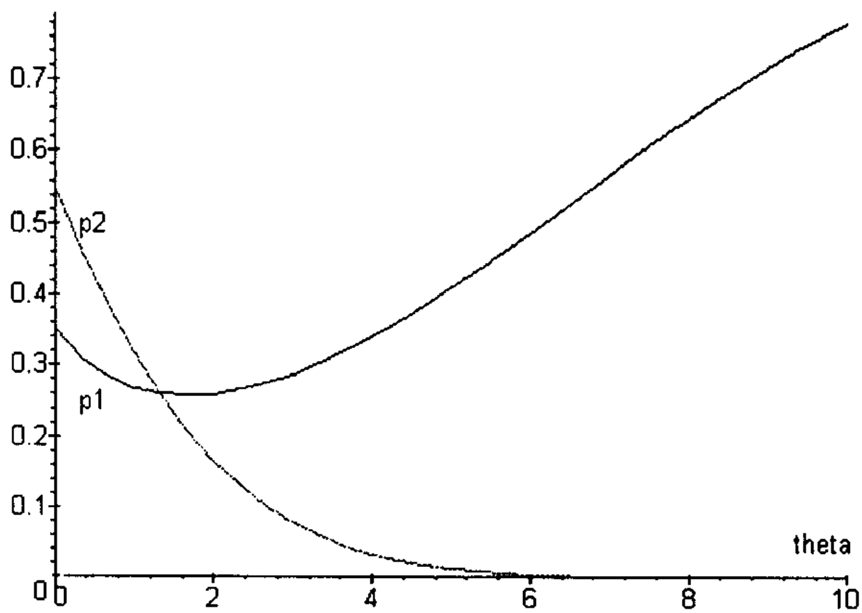


Figura III.6.4b: Curvas solucion



espera que los sujetos caracterizados por valores de θ mayores a este valor responderán correctamente este ítem.

Análisis del ítem 5

Los parámetros para este ítem lo sitúan entre los reactivos correspondientes al caso 7 para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, a_1 es 0.66 por lo que se puede decir que la opción correcta es medianamente accesible y dado que b_1 es casi igual a 1, las opciones de respuesta resultan similares para el sujeto que responde. En este ítem, la opción correcta no es muy fácil de reconocer y la interacción juega un papel importante. Además, las condiciones iniciales indican que la región del caso 7 a la que esta pregunta pertenece es la región A de la figura III.6.5a, lo cual permite inferir que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.5b, no permite reconocer el intervalo de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 0.78. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.5a: Regiones en plano fase

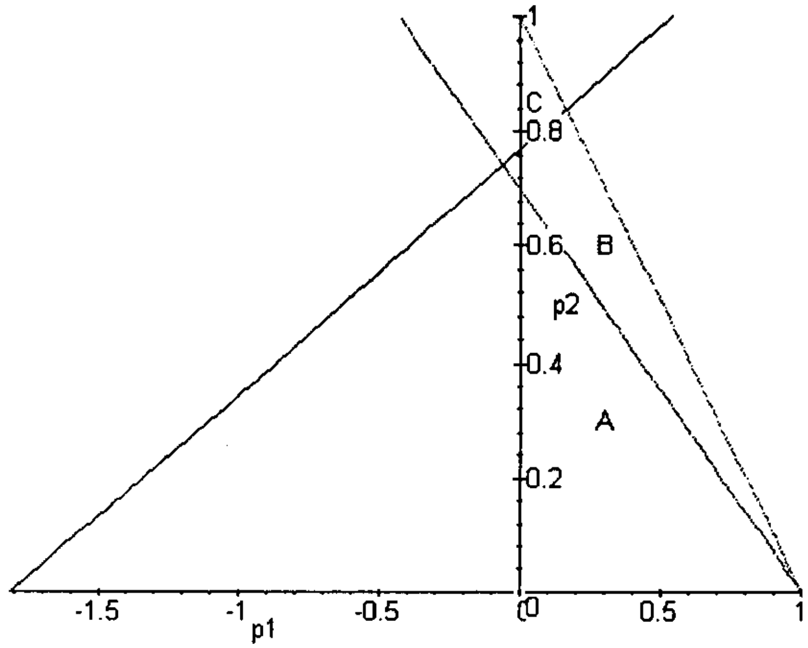
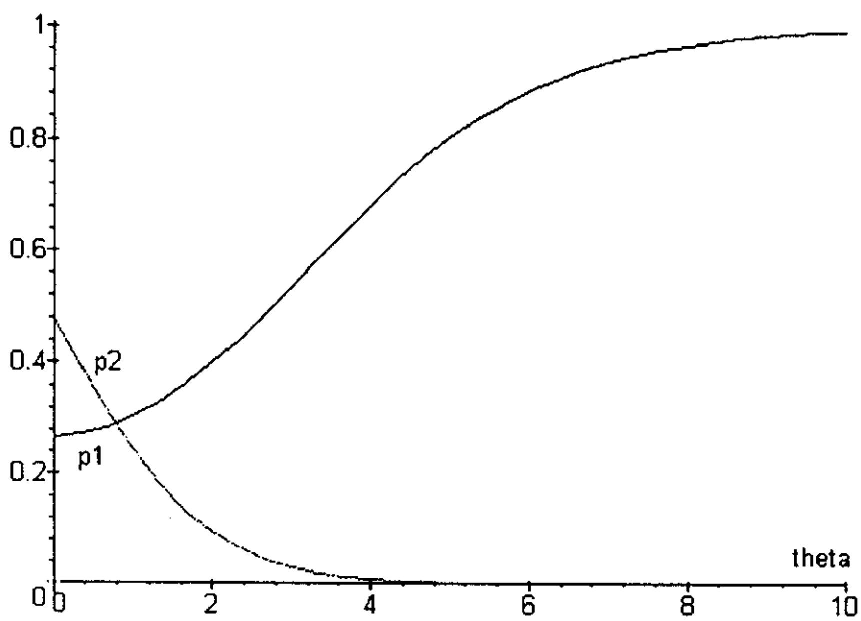


Figura III.6.5b: Curvas solucion



Análisis del ítem 9

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos del caso 7 en los que a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es muy cercana a uno, las opciones de respuesta resultan similares para el sujeto que responde y puesto que a_1 es 0.734 la accesibilidad de la opción correcta es mediana. Esto implica que en esta pregunta la opción correcta no es muy fácil de reconocer y que, además, la interacción juega un papel importante. Las condiciones iniciales indican que la región del caso 7 a la que esta pregunta pertenece es la región B de la figura III.6.9a, lo cual implica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.9b, permite reconocer el intervalo de la variable independiente (0,1.164) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.4. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.9a: Regiones en plano fase

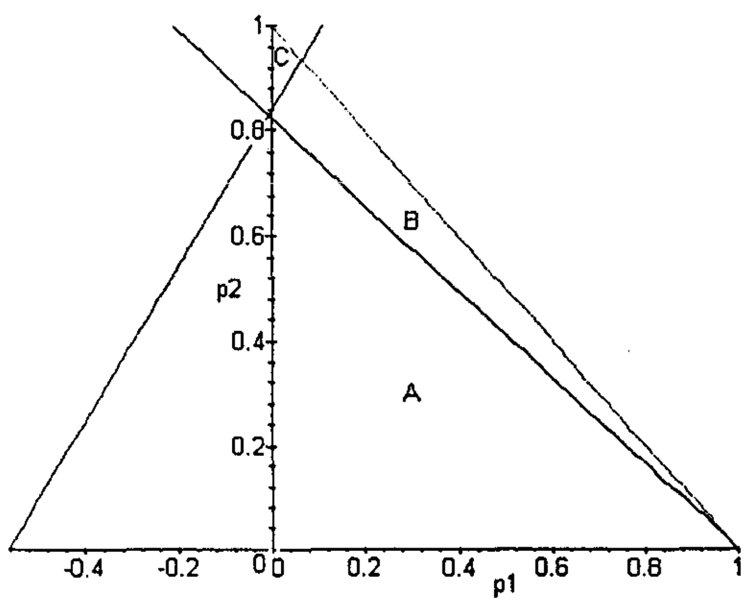
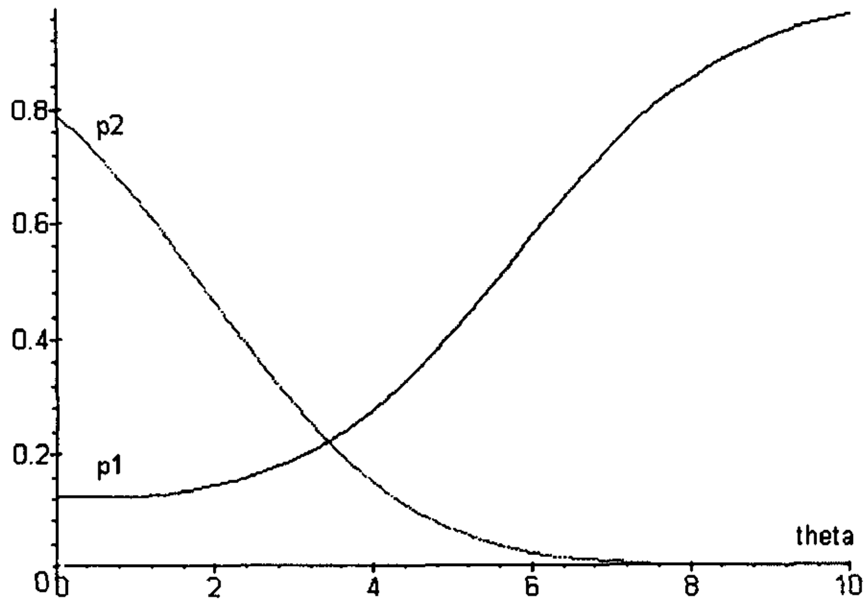


Figura III.6.9b: Curvas solucion



Análisis del item10

Los parámetros de este ítem lo sitúan entre los reactivos incluidos en el caso 7 que corresponde a aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a uno, la interacción que ejerce la opción incorrecta es grande, es decir la opción de respuesta incorrecta es atractiva para el sujeto que responde y puesto que a_1 es 1.006 la accesibilidad de la opción correcta es grande. Esto implica que en esta pregunta la opción correcta es fácil de reconocer pero que la respuesta incorrecta interactúa fuertemente. Las condiciones iniciales indican que la región del caso 7 a la que esta pregunta pertenece es la región B del espacio fase que se muestra en la figura III.6.10a, por lo cual la gráfica correspondiente a este ítem, que se muestra en la figura III.6.10b, permite reconocer el intervalo de la variable independiente (0, 0.731) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 2.257. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.10a: Regiones en plano fase

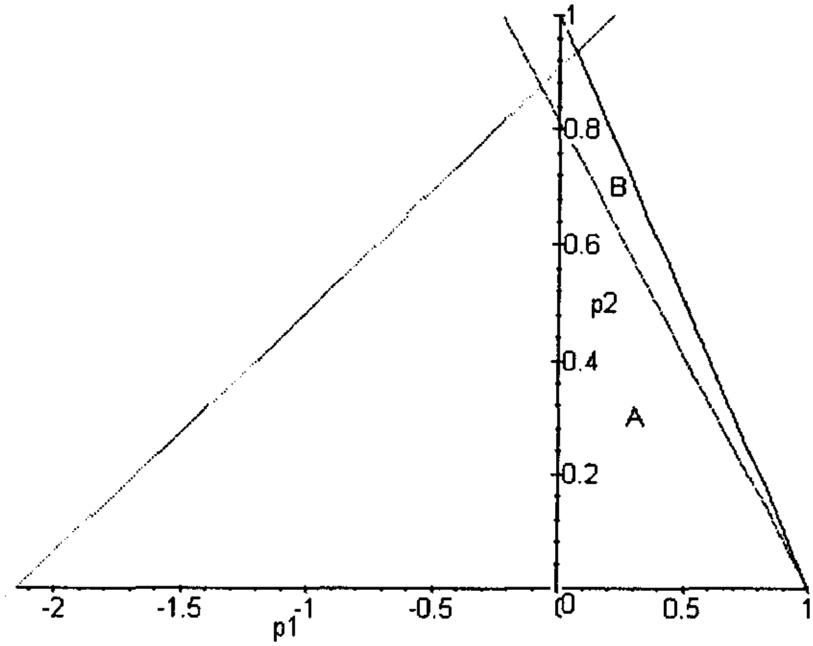
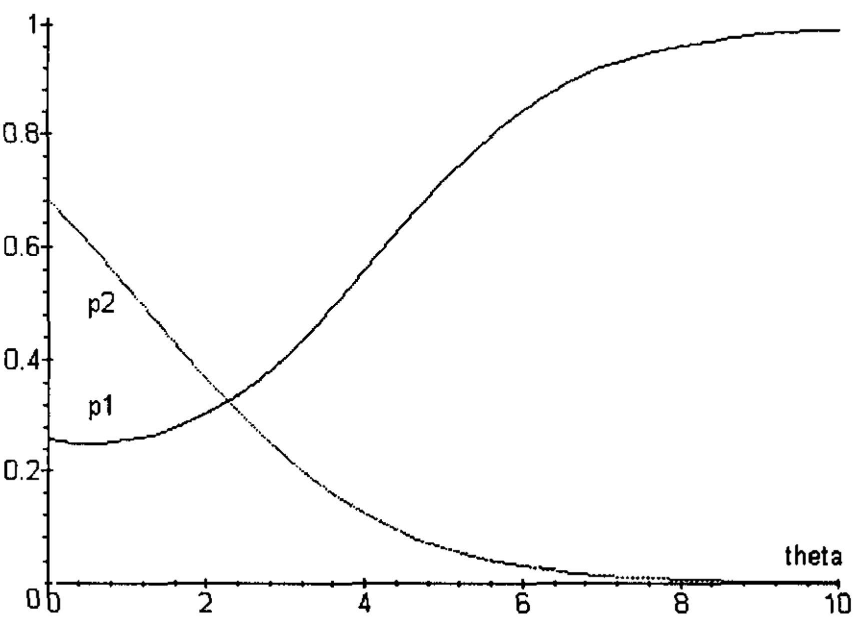


Figura III.6.10b: Curvas solucion



Análisis del ítem 12

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos pertenecientes al caso 7 en los que a_2 es negativa y que son tales que a_1/b_1 es menor a k_2. El hecho de que el parámetro de accesibilidad sea en este caso menor a uno y el parámetro de interacción b_1 es también menor a uno indican que, si bien la opción 1 es difícil de reconocer como la opción correcta, la interacción entre las opciones de respuesta no juega un papel muy importante. Dado que las condiciones iniciales se encuentran en la región del espacio fase indicada por la letra B en la figura III.6.12a, la gráfica correspondiente a este ítem, que se muestra en la figura III.6.12b, permite reconocer el intervalo de la variable independiente (0,1.21) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.19. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.12a: Regiones en plano fase

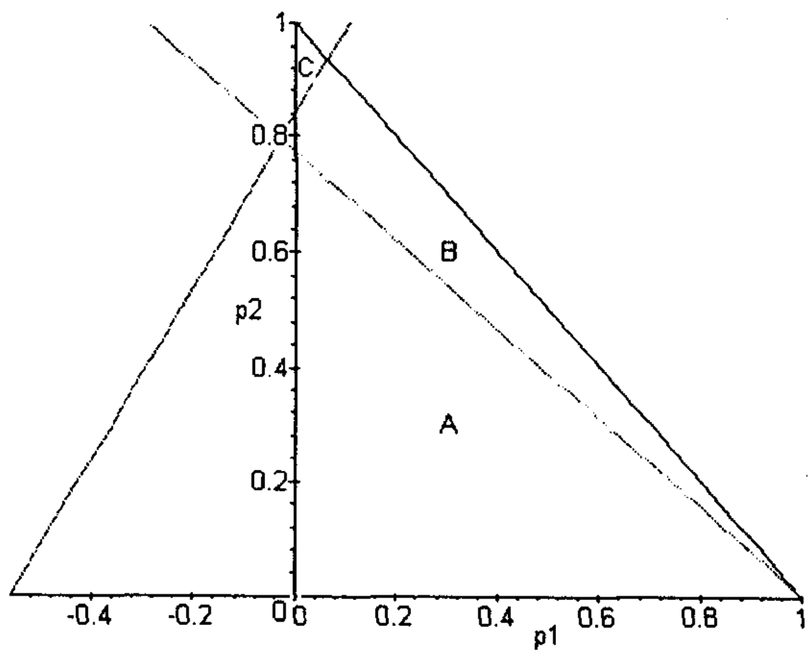
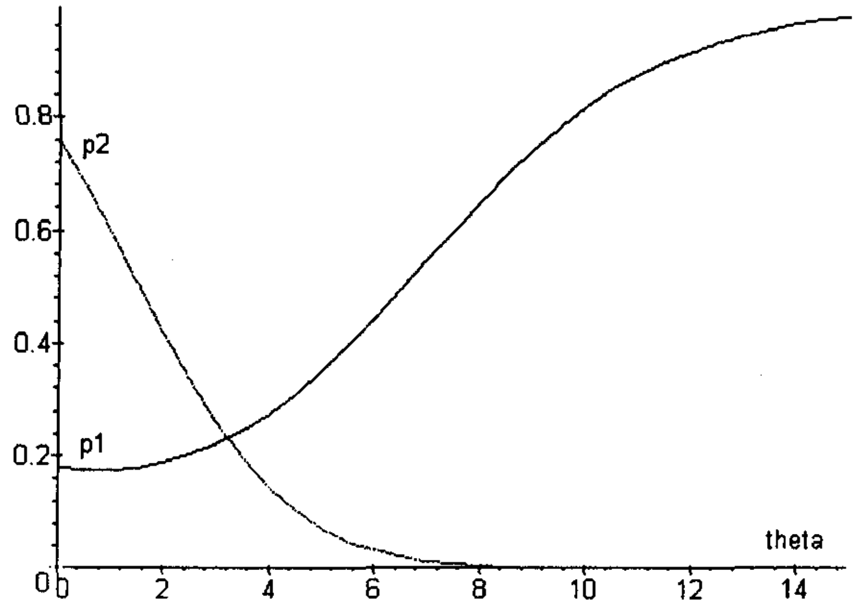


Figura III.6.12b: Curvas solucion



Análisis del ítem 15

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos que pertenecen al caso 7 en los que a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. El parámetro de accesibilidad para este ítem es de 0.421, lo que significa que la opción correcta no es fácil de reconocer y puesto que b_1 es menor a uno, la interacción de la opción 2 no es muy fuerte. Las condiciones iniciales indican que la región del espacio fase correspondiente a esta pregunta pertenecen a la región B de la figura III.6.15a, por lo cual la gráfica correspondiente a este ítem, que se muestra en la figura III.6.15b, permite reconocer el intervalo de la variable independiente (0, 3.04) para el que los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.72. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.15a: Regiones en plano fase

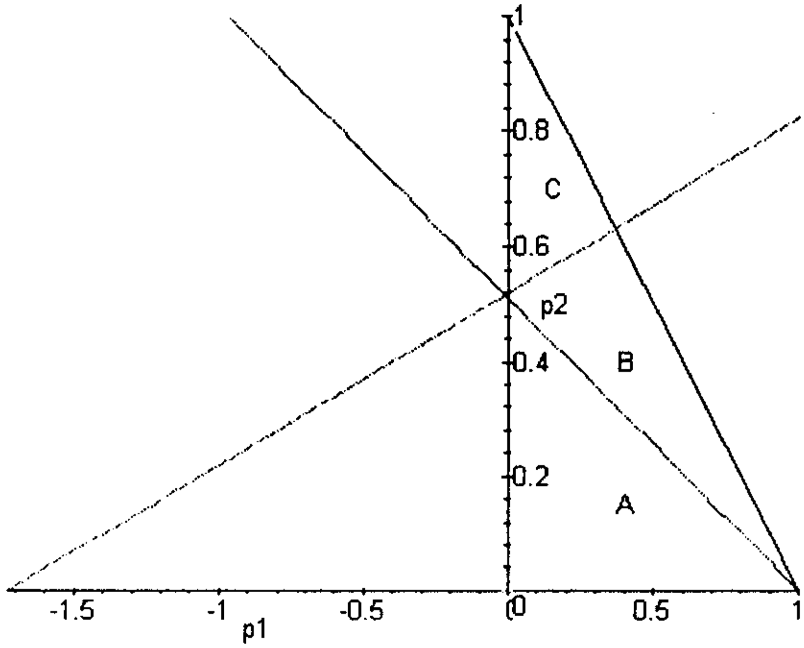
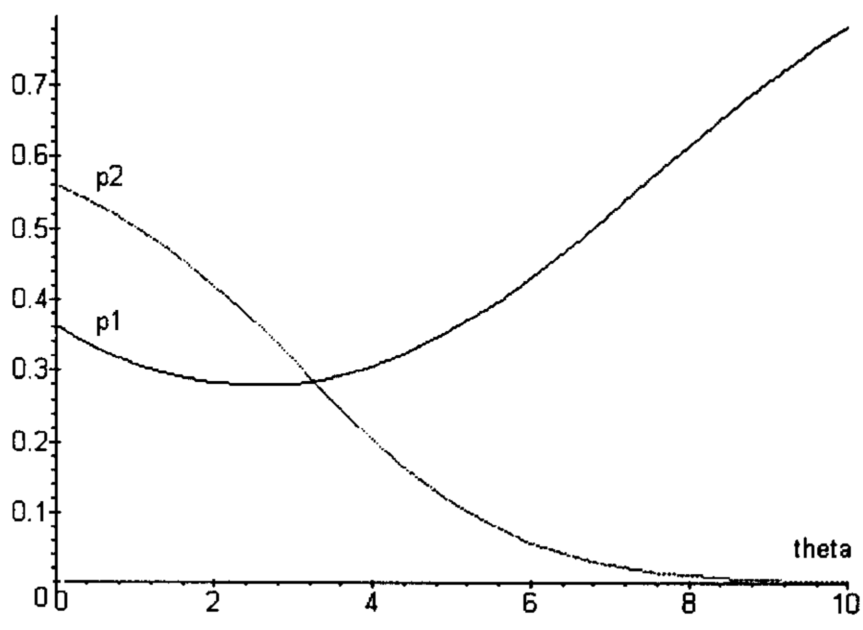


Figura III.6.15b: Curvas solucion



Análisis del ítem 16

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos pertenecientes al caso 7 que corresponde a aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a uno, la opción 2 interactúa fuertemente con la opción de respuesta correcta. Dado que a_1 toma el valor de 0.324 la accesibilidad de la opción correcta es pequeña. Esto implica que en esta pregunta la opción correcta no es fácil de reconocer y que además la interacción juega un papel importante. Además, las condiciones iniciales indican que la región del espacio fase a la que esta pregunta pertenece es la región B de la figura III.6.16a, lo cual indica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.16b, permite reconocer el intervalo de la variable independiente (0,2.28) para el cual los sujetos se encuentran posiblemente en transición entre le *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.5. Este valor indica que los sujetos caracterizados por valores mayores de éste, responderán correctamente esta pregunta.

Figura III.6.16a: Regiones en plano fase

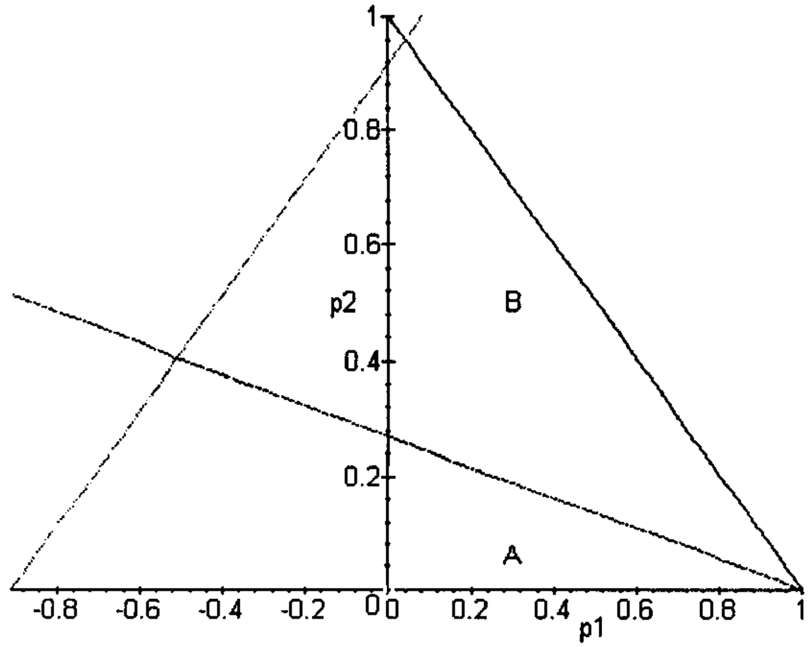
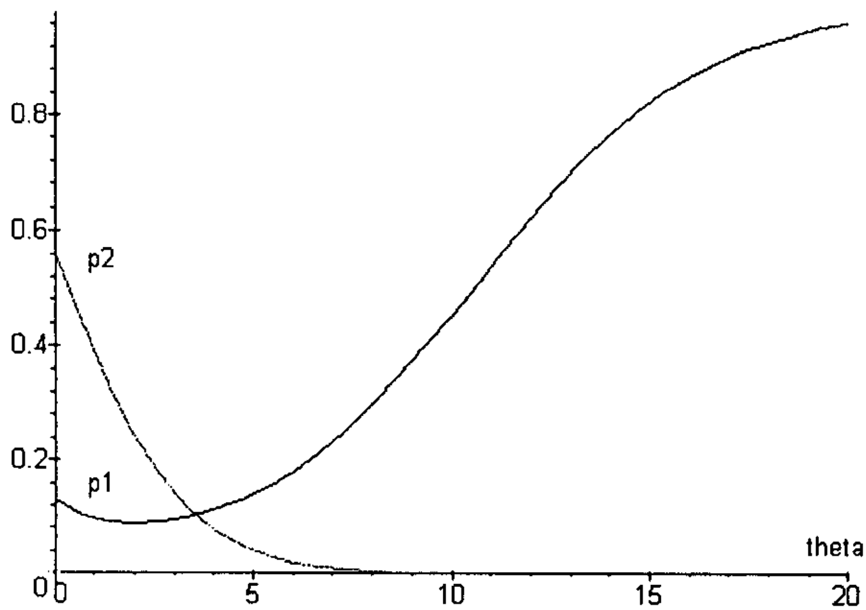


Figura III.6.16b: Curvas solucion



Análisis del ítem 18

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos pertenecientes al caso 7 en el que los ítems se carcterizan por que a_2 es negativa y porque a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es menor a uno, la interacción que ejerce la opción incorrecta no es grande y puesto que a_1 es menor a uno la accesibilidad de la opción correcta es pequeña para el sujeto que responde. Esto implica que en esta pregunta la opción correcta no es fácil de reconocer independientemente del papel de la interacción con la otra opción de respuesta que se está considerando. Además, las condiciones iniciales indican que la región del espacio fase a la que corresponden los datos de este ítem es la denotada por B en la figura III.6.18a, por lo cual la gráfica correspondiente a este ítem, que se muestra en la figura III.6.18b, permite reconocer el intervalo de la variable independiente (0,0.75) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. En el valor θ = 0.75 se cruzan las curvas que describen la probabilidad de elección de cada una de las opciones. Puede considerarse que los sujetos cuya variable independiente sea mayor a este valor responderán correctamente este ítem.

Figura III.6.18a: Regiones en plano fase

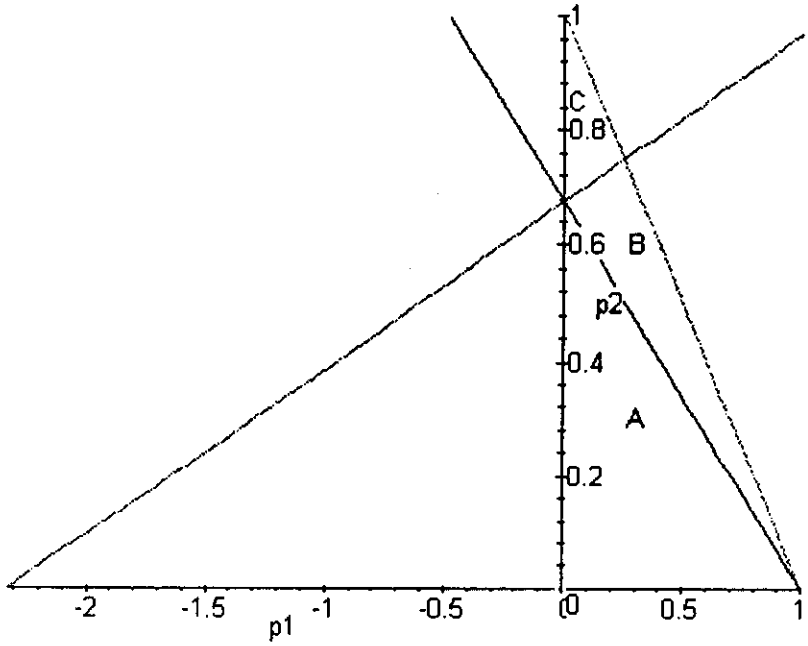
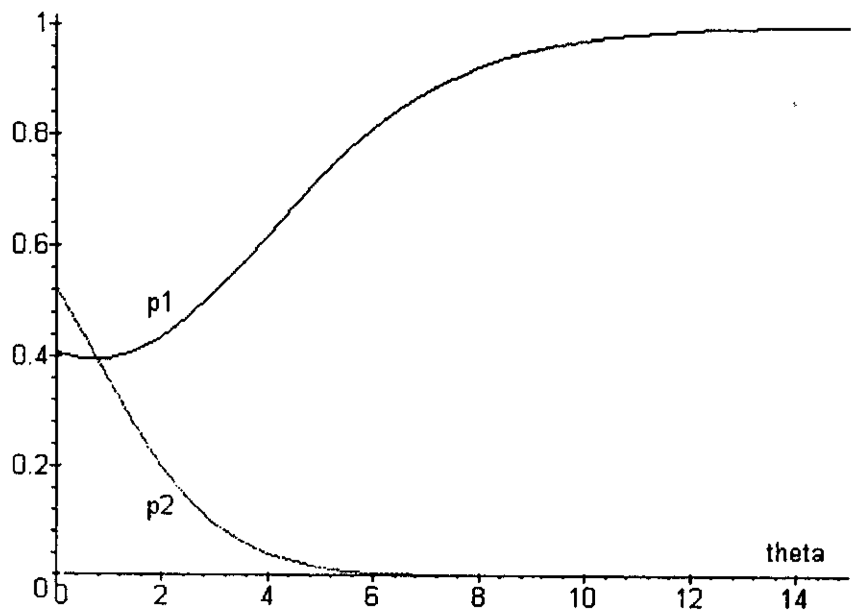


Figura III.6.18b: Curvas solucion



Análisis del ítem 21

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos correspondientes al caso 7 para los que a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. La opción correcta es poco accesible ya que a_1 es 0.431. Por otra parte b_1 es mayor a 1, es decir, la interacción que ejerce la opción incorrecta es grande, la opción de respuesta incorrecta resulta atractiva para el sujeto que responde. Las condiciones iniciales se encuentran en la región del espacio fase que se indica con la letra B en la figura III.6.21a, lo cual indica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.21b, permite reconocer el intervalo de la variable independiente (0,4.24) para el cual los sujetos se encuentran posiblemente en transición entre le *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 7.53. Este valor indica que los sujetos caracterizados por valores mayores de éste, responderán correctamente esta pregunta.

Figura III.6.21a: Regiones en plano fase

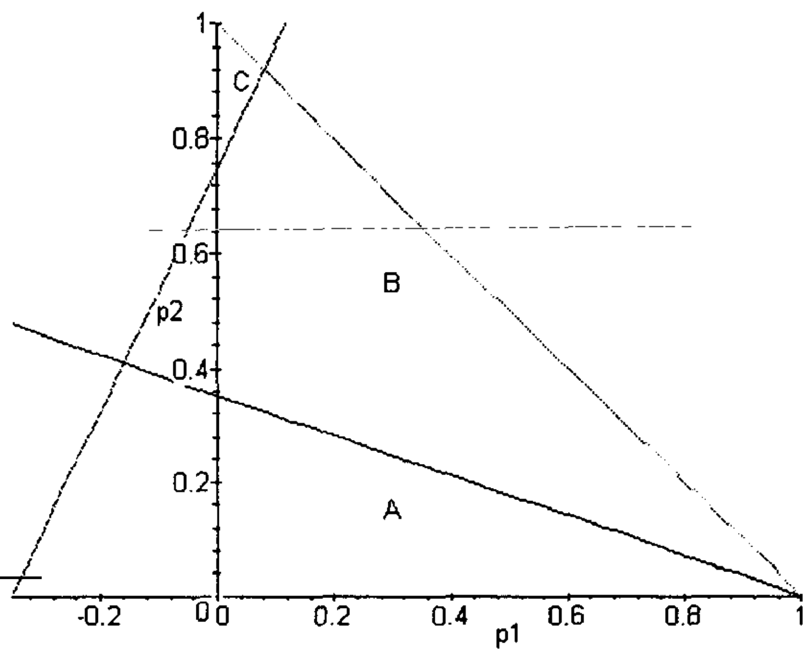
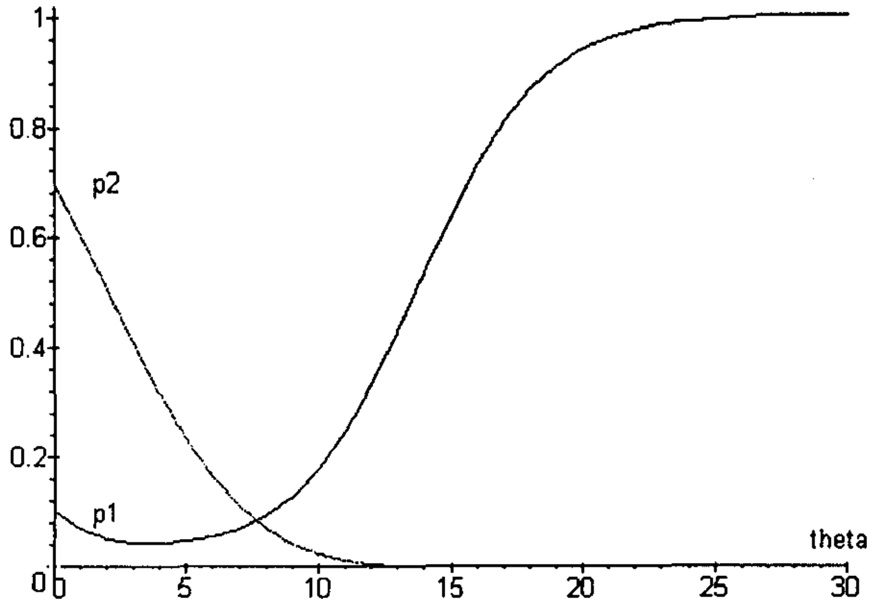


Figura III.6.21b: Curvas solucion



Análisis del ítem 24

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos del caso 7 que corresponde a aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a 1, la interacción que ejerce la opción incorrecta es grande, es decir la opción de respuesta incorrecta es atractiva para el sujeto que responde y puesto que a_1 es 0.651 la accesibilidad de la opción correcta es pequeña para el sujeto. Esto implica que en esta pregunta la opción correcta no es fácil de reconocer y además la opción de respuesta incorrecta es atractiva. Por otra parte, las condiciones iniciales indican que la región del espacio fase correspondiente a esta pregunta pertenece a la región B que se muestra en la figura III.6.24a, por ello, la gráfica correspondiente a este ítem, que se muestra en la figura III.6.24b, permite reconocer el intervalo de la variable independiente (0,3.445) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel* -

proceso y el *nivel - objeto*. En el valor $\theta = 5.27$ se cruzan las curvas que describen la probabilidad de elección de cada una de las opciones. Puede considerarse que los sujetos cuya variable independiente sea mayor a este valor responderán correctamente este ítem.

Figura III.6.24a: Regiones en plano fase

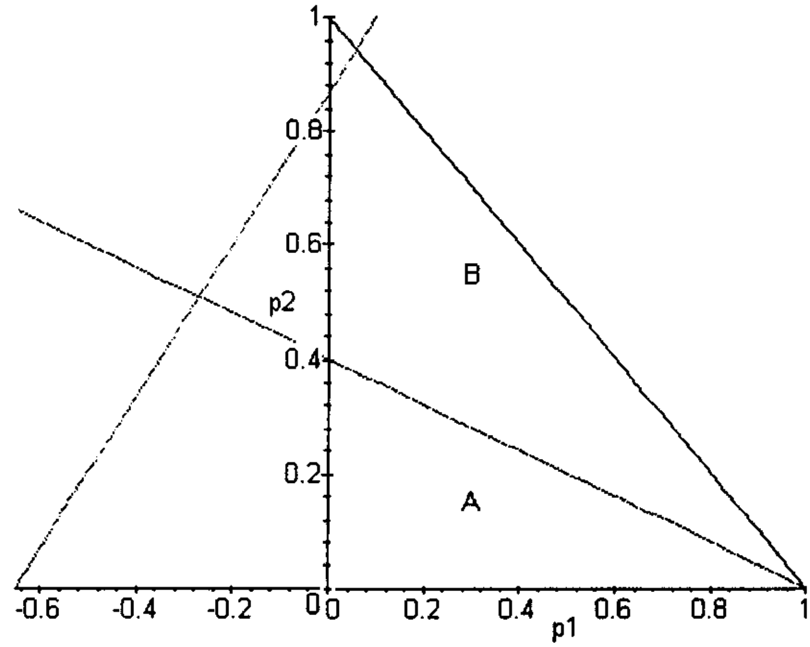
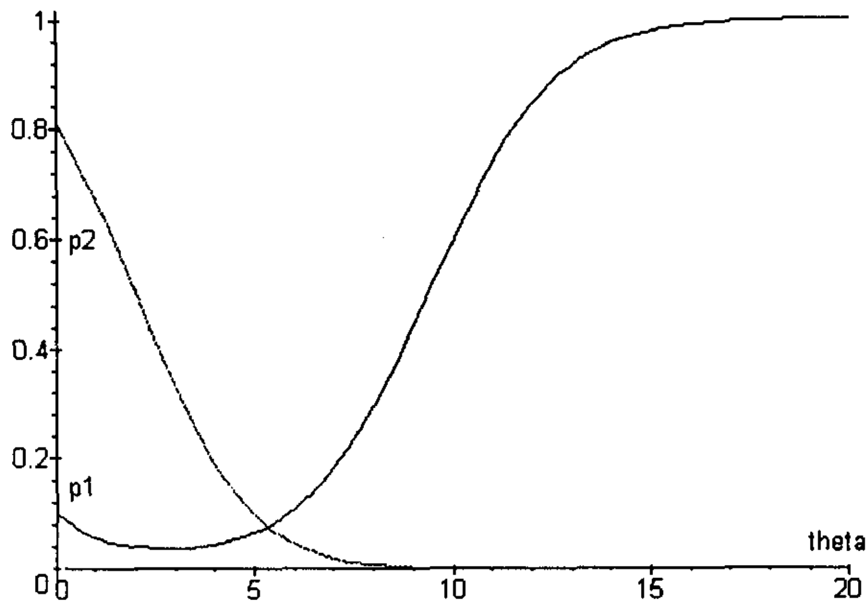


Figura III.6.24b: Curvas solucion



Análisis del ítem 26

Los parámetros que se obtienen para este ítem lo sitúan entre los reactivos pertenecientes al caso 7 que corresponde a aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. Para este ítem la opción correcta es medianamente accesible, ya que su valor es de $a_1 = 0.679$. Por otra parte, como b_1 es igual a uno, la interacción que ejerce la opción incorrecta es grande, es decir la opción de respuesta incorrecta es atractiva para el sujeto que responde. Además, las condiciones iniciales implican que la región del espacio fase a la que esta pregunta pertenece es la región denotada por la letra B en la figura III.6.26a, lo cual indica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.26b, permite reconocer el intervalo de la variable independiente (0, 1.583) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 4.108. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.26a: Regiones en plano fase

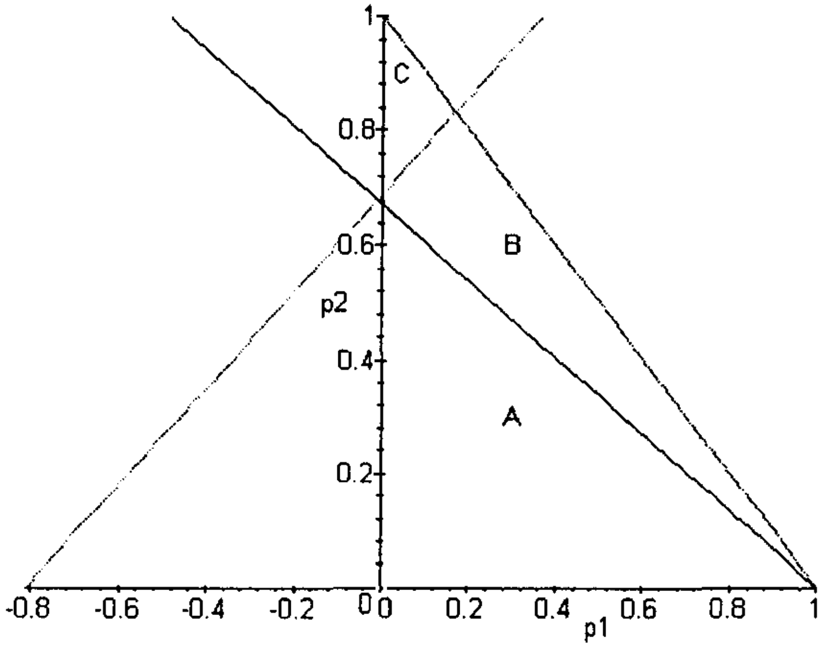
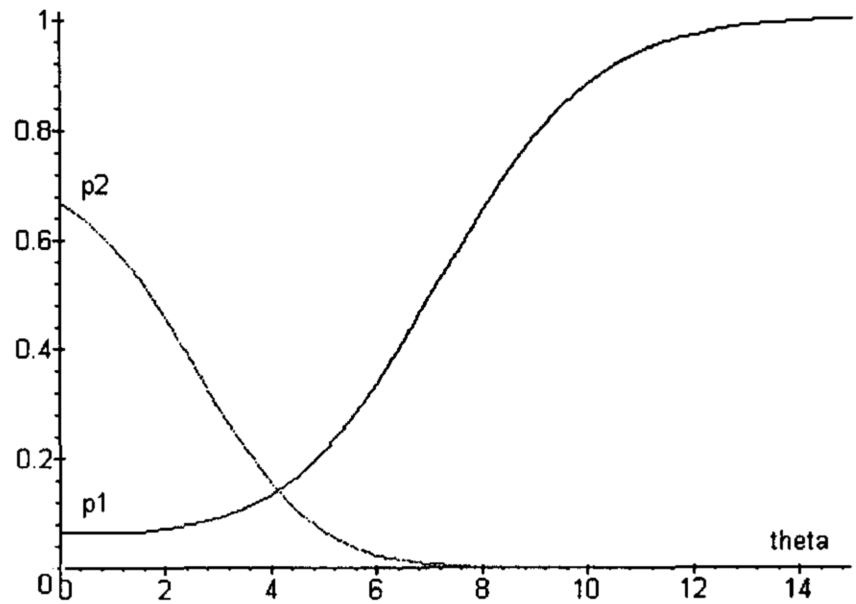


Figura III.6.26b: Curvas solucion



Análisis del ítem 30

Los parámetros correspondientes a este ítem lo sitúan entre los reactivos del caso 7 en el que se encuentran aquellos ítems para los cuales a_2 es negativa y que son tales que a_1/b_1 es menor que k_2. En este caso, puesto que b_1 es mayor a 1, la interacción que ejerce la opción incorrecta es grande, es decir la opción de respuesta incorrecta es atractiva para el sujeto que responde y puesto que a_1 es asimismo mayor a uno, la accesibilidad de la opción correcta es grande. Esto implica que en esta pregunta la opción correcta es fácil de reconocer pero que la respuesta incorrecta interactúa fuertemente con la correcta. Las condiciones iniciales indican que la región del espacio fase a la que esta pregunta pertenece es la región marcada con la letra B en la figura III.6.30a, por lo que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.30b, permite reconocer el intervalo (0, 0.78) de la variable independiente para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - acción* y el *nivel - proceso*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.04. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.30a: Regiones en plano fase

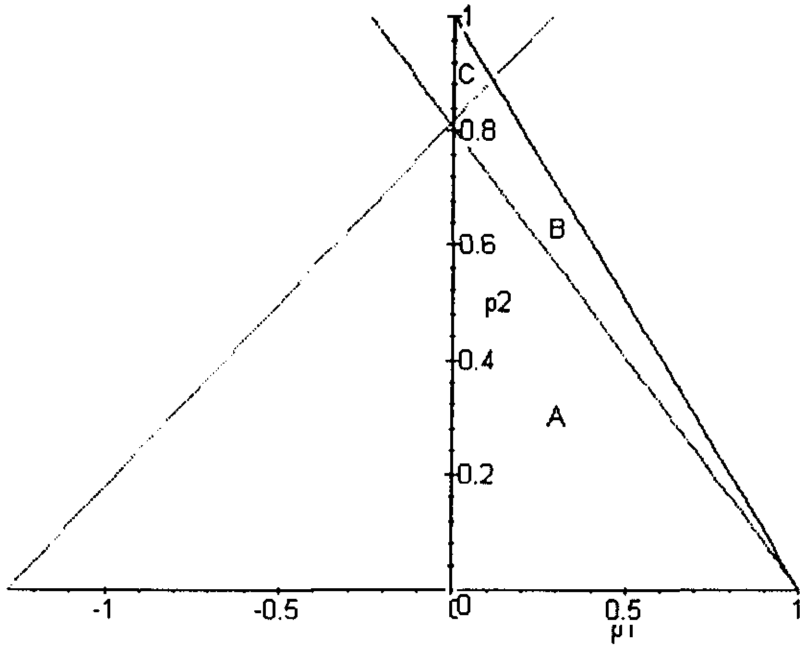
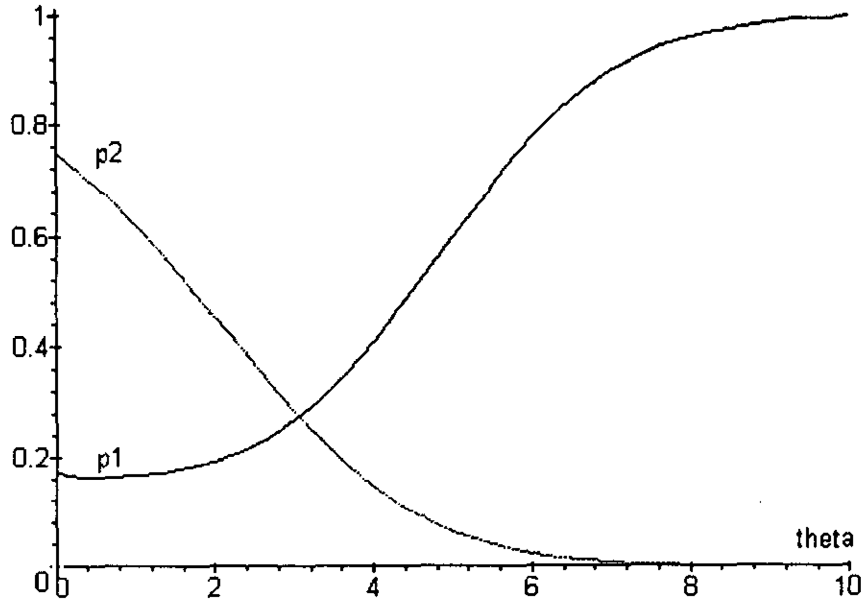


Figura III.6.30b: Curvas solucion



Ítems correspondientes al caso 8

Análisis del ítem 8

Los parámetros obtenidos para este ítem lo sitúan entre los reactivos correspondientes al caso 8 en los que a_2 es negativa y que son tales que a_1/b_1 es casi igual a k_2. En este caso, puesto que b_1 es mayor a uno, la opción 2 interactúa fuertemente con la opción de respuesta correcta. El parámetro de accesibilidad a_1 es 0.874 e indica que la opción correcta es medianamente accesible para el sujeto. Esto implica que en esta pregunta la opción correcta no es muy fácil de reconocer pero que la opción incorrecta representa un factor importante de interacción. Además, las condiciones iniciales sitúan a la curva del ítem en la región del caso 8 que se marca con la letra B de la figura III.6.8a, por lo cual la gráfica correspondiente a este ítem, que se muestra en la figura III.6.8b, permite reconocer el intervalo de la variable independiente (0, 1.01) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel* -

proceso y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 4.31. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura II.6.8a: Regiones en plano fase

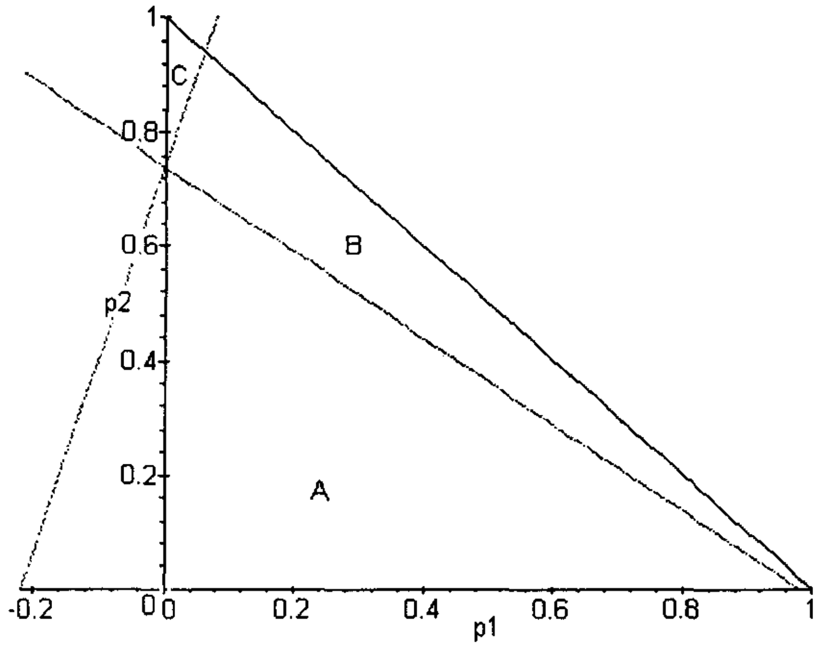
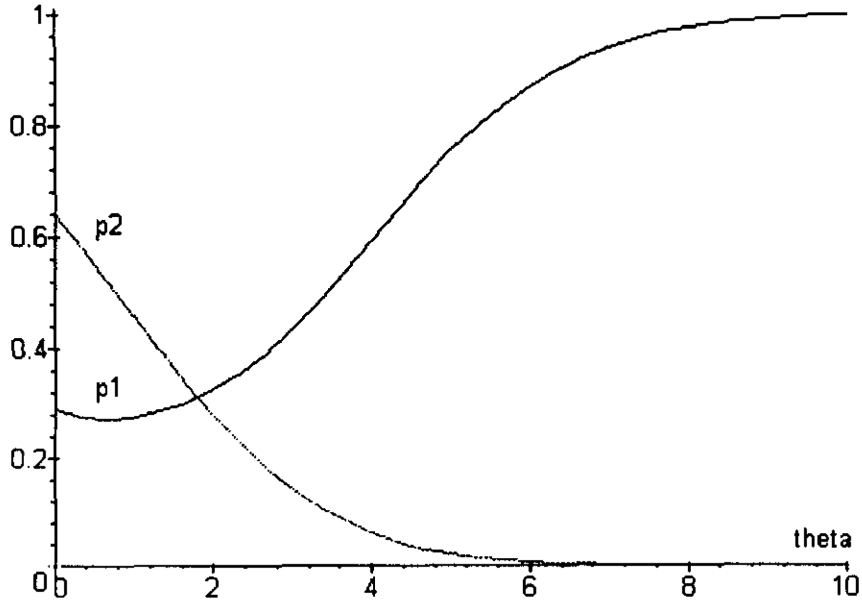


Figura III.6.8b: Curvas solucion



Análisis del ítem 14

Los parámetros estimados para este ítem lo sitúan entre los reactivos correspondientes al caso 8 en los que a_2 es negativa y a_1/b_1 es casi igual a k_2. Para este ítem la accesibilidad a_1 es 0.74, lo cual significa que la opción correcta es medianamente accesible para el sujeto. En este caso, puesto que b_1 es casi igual a 1, la interacción entre las opciones de respuesta es pareja, así es que en esta pregunta la opción correcta no es muy fácil de reconocer pero la opción incorrecta representa un factor importante de interacción. Además, las condiciones iniciales indican que la región del espacio fase a la que esta pregunta pertenece se encuentra en la sección denominada B en la figura III.6.14a, lo cual implica que la gráfica correspondiente a este ítem, que se muestra en la figura III.6.14b, permite reconocer el intervalo de la variable independiente (0,0.83) para el cual los sujetos se encuentran posiblemente en transición entre el *nivel - proceso* y el *nivel - objeto*. Las curvas de las probabilidades de elección de las dos opciones se cruzan cuando θ vale 3.72. Este valor indica que los sujetos caracterizados por valores mayores a éste, responderán correctamente esta pregunta.

Figura III.6.14a: Regiones en plano fase

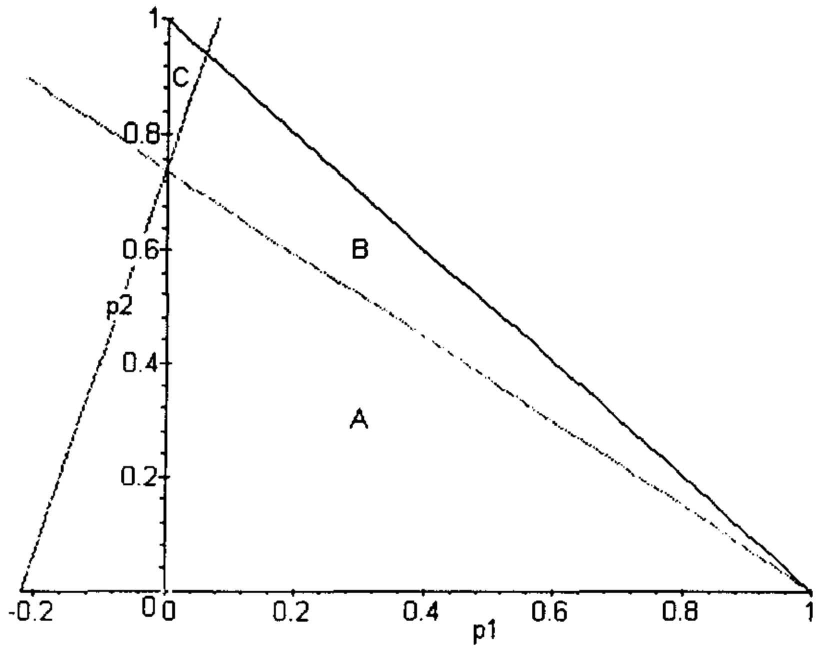
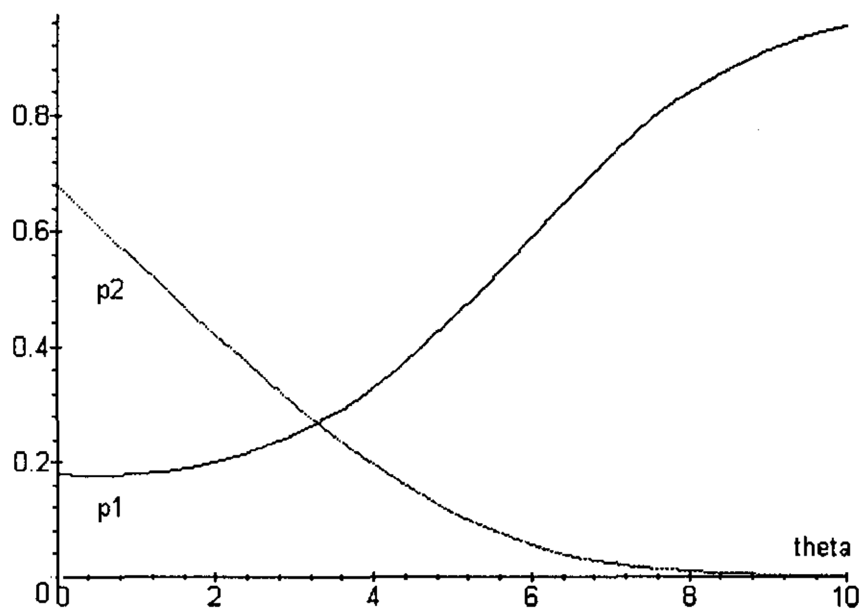


Figura III.6.14b: Curvas solucion



En el caso de los ítems 6 y 27 el parámetro correspondiente a la interacción de la opción 2 con la opción 1, b_1 resultó negativo. Estos dos ítems se analizan conjuntamente.

Análisis de los ítems cuyos parámetros de interacción son negativos

Análisis de los ítems 6 y 27

Los parámetros correspondientes a estos ítems los sitúan entre los reactivos del caso 13 para los cuales a_2 es negativa y que son tales que a_1/b_1 es mayor que k_2. En este caso el parámetro b_1 es negativo lo que indica que la presencia de la opción de respuesta 2 contribuye a aumentar la probabilidad de respuesta de la opción 1. En estos casos el distractor no cumple su papel por lo que habría que rediseñar los ítems o sustituirlos por otros en el test. Las figuras correspondientes al espacio fase y a las soluciones de estos dos ítems se muestran respectivamente en las figuras III.6.6a y b y III.27a y b.

Figura III.6.6a: Regiones en plano fase

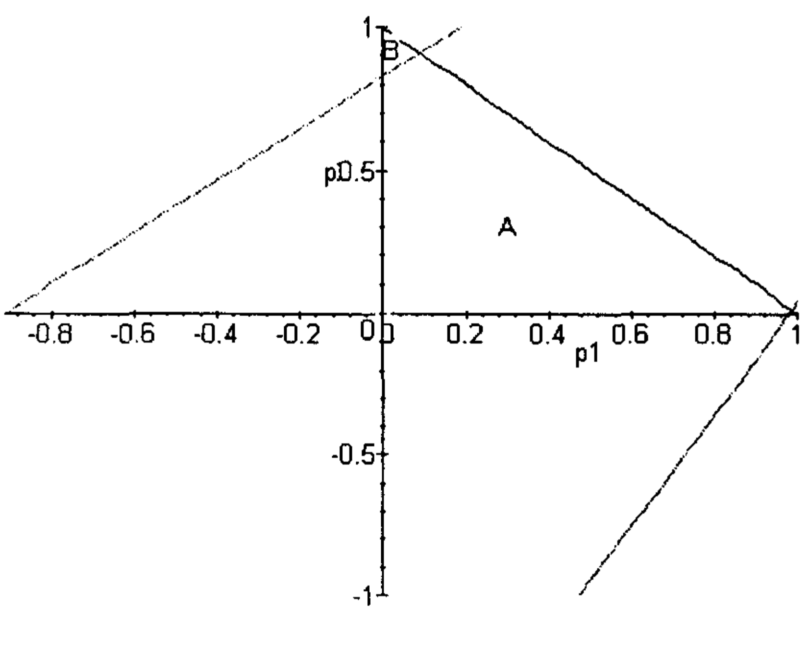


Figura III.6.6b: Curvas solucion

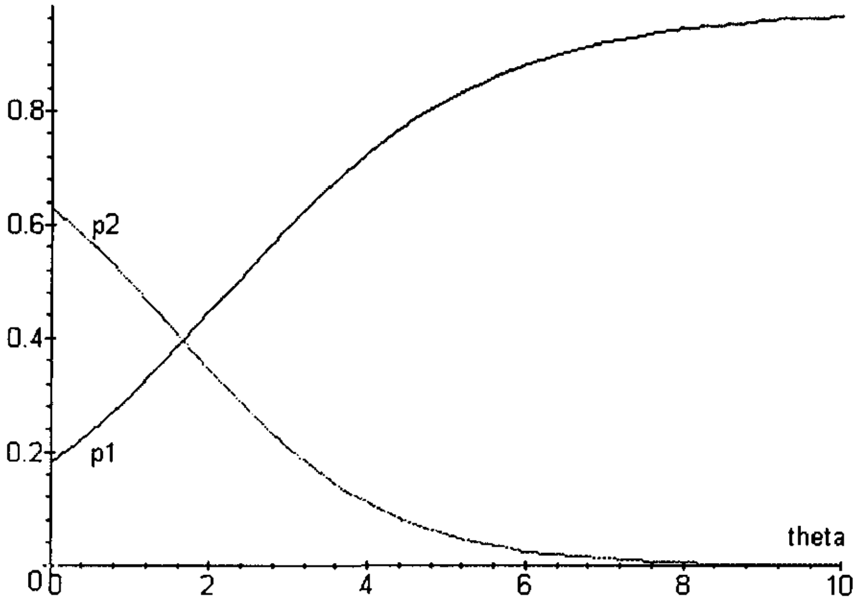


Figura III.6.27a: Regiones en plano fase

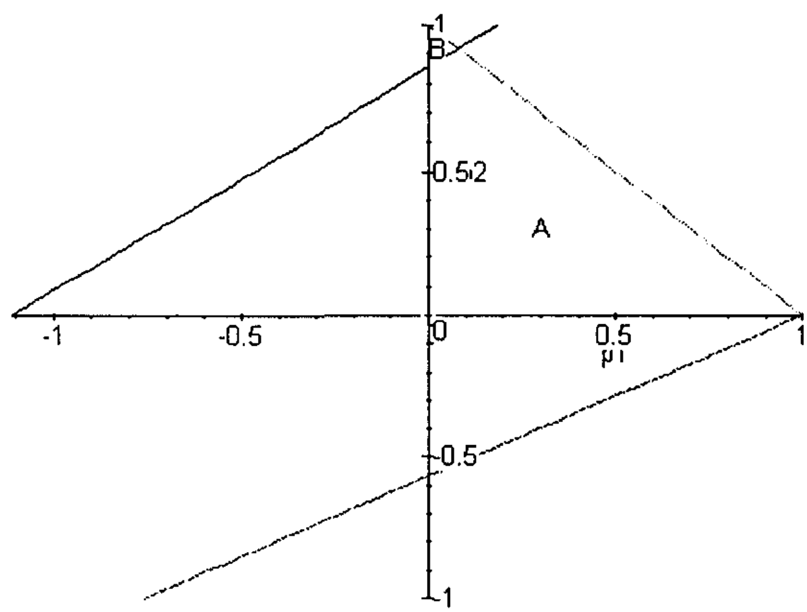
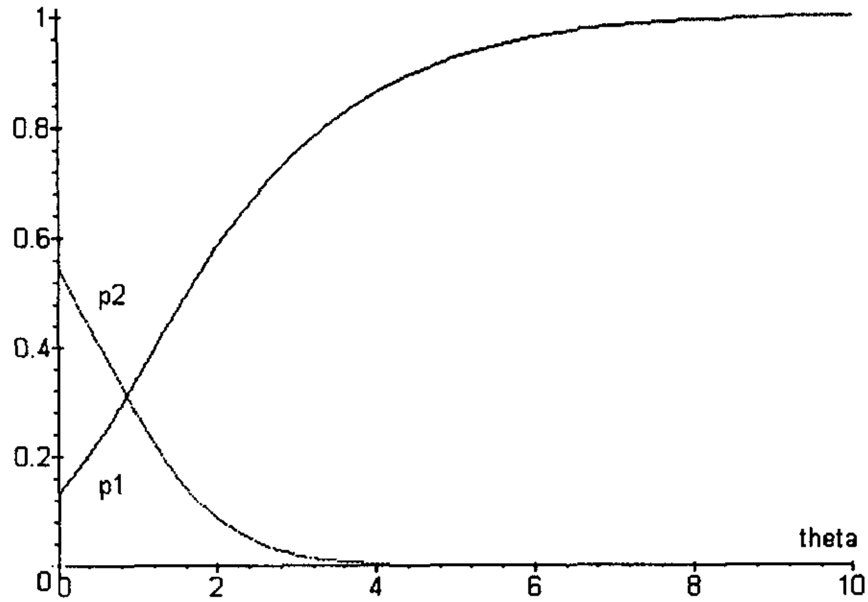


Figura III.6.27b: Curvas solucion



Algunas conclusiones

Al introducir el estudio de las componentes del ítem a través de un parámetro de accesibilidad de la opción se está incluyendo un cierto modelo de complejidad cognitiva en el ítem. Las opciones pueden diseñarse de manera que esten fundamentadas en estudios cognitivos que les confieran un determinado estatus en cuanto a la jerarquía que ocupan dentro del desarrollo del concepto. El parámetro de interacción complementa la información que proporciona la medida de accesibilidad de la opción al proporcionar una medida del nivel de interferencia que la presencia de una opción viable, dentro de ciertos límites de desarrollo conceptual del estudiante, presenta a quien responde.

A diferencia del modelo linear logístico de Fisher (1973) que se ha descrito con anterioridad, en el modelo que aquí se presenta no se presupone ningún submodelo en cuanto a la forma en la que las complejidades cognitivas de los ítems influyen en el procesamiento de la información. Las opciones pueden calibrarse independientemente en cuanto a la dificultad que involucran a través del análisis de las respuestas de los estudiantes a ítems abiertos o más simples en los que se utilice el modelo logístico de dos parámetros en el análisis de la información.

A lo largo del análisis de los ítems empleados para probar el modelo es posible darse cuenta de que aún cuando el método de estimación de los parámetros es perfectible, los resultados obtenidos son satisfactorios.

El modelo permite describir a los ítems a través del análisis de los distractores y modela adecuadamente la interacción que puede darse en la mente del sujeto que responde cuando se le presentan preguntas de opción múltiple.

En la mayor parte de los ítems aquí analizados es posible localizar un intervalo de valores de la variable independiente para la cual las probabilidades de elección de las dos opciones en estudio decrecen, lo cual puede interpretarse, como ya se había mencionado anteriormente, como una posible transición entre un nivel de conocimiento y otro, cuando los ítems se analizan a la luz de una teoría cognitiva.

El punto en el que las curvas de probabilidad de elección de las opciones en estudio se cruzan, provee también información interesante. Este punto podría asociarse a la dificultad del ítem en términos de este modelo.

Las gráficas asociadas a la probabilidad de elección de cada uno de los distractores, conjuntamente con la representación de la curva de la variación de estas probabilidades en el espacio fase, permiten darse una idea muy clara del comportamiento de dichas probabilidades y de la posible existencia de un estado de transición.

La mayor limitación del modelo es, sin duda, la no linearidad de las ecuaciones diferenciales que lo constituyen y la necesidad de desarrollar un método de ajuste de los parámetros diseñado específicamente para él. ◆

Referencias

ABRAHAMOWITZ, M. y RAMSAY, J.O. (1992). Multicategorical spline model for item response theory. Psychometrika, 57, 5-27.

ARZARELLO, F., BAZZINI, L. y CHIAPPINI, G. (1993). Cognitive processes in algebraic thinking: towards a theoretical framework, Proc. PME XVII, Tsukuba, Japan, I, 138-145.

ARZARELLO, F., BAZZINI, L. y CHIAPPINI, G. (1994). The process of naming in algebraic problem solving, Proc. PME XVIII, Lisbon, II, 40-47.

ASIALA A., BROWN A., DEVRIES D., DUBINSKY E., MATHEWS D. Y THOMAS K. (1996). A Framework for Research and Curriculum Development in Undergraduate Mathematics Education, en Kaput J., Shoenfeld A. y Dubinsky E. (Eds.) Research in Collegiate Mathematics Education II, pp. 1-32.

ASIALA, M., COTTRILL, J., DUNINSKY, E. y SCHWINGENDORF, K. (1997). The development of students' graphical understanding of the derivative. Journal of Mathematical Behavior, 16, 399-431.

AVILA A., GARCIA F. y ROJANO T. (1990). Algebraic Syntax Errors: A Study with Secondary School Children. Proceedings of the Fouteenth PME Conference, México pp. 11-18.

BEDNARZ, N. y JANVIER, B. (1994). The emergence and development of algebra in a problem solving context: A problem analysis, Proc. PME XVIII, Lisbon, II, 64-71.

BEDNARZ, N., JANVIER, B., MARY, C. y LEPAGE, A. (1992). L'algèbre comme outil de résolution de problèmes: une réflexion sur les changements nécessaires dans le passage d'un mode de traitement arithmétique à un mode de traitement algébrique, Recueil des textes du Colloque du programme de Recherche sur l'emergence de l'algèbre, 17-31. Montréal, CIRADE, Université du Quebec à Montreal.

BELL, A. y MALONE, J. (1993). Learning the language of algebra, Proc. PME XVII, Tsukuba, Japan, I, 130-137.

BEDNARZ N. y DUFOUR-JANVIER B. (1991). A Study of External Representations of Change developed by Young Children. Proceedings of the Thirteenth PME-NA Conference, Blacksburg, Virginia, pp. 140-146.

BETH, E.W. y PIAGET, J. (1966). Mathematical Epistemology and Psychology. Reidel Dordrecht.

BOLEA, P., BOSCH, M. y GASCÓN, J., (1998a). The role of algebraization in the study of a mathematical organization, CERME-1, Osnabrueck, Germany.

BOLEA, P., BOSCH, M. y GASCÓN, J. (1998b). Le caractère problématique du processus d'algébrisation. Proporcionnalité et grandeurs dans l'enseignement obligatoire, Actes de la IXème école d'été de didactique des mathématiques, ARDM, 153-159.

BOLEA, P., BOSCH, M. y GASCÓN, J. (1998c). El proceso de algebrización de las matemáticas escolares, (pendiente de publicación).

BOOTH L. (1984). Algebra: Children's Strategies and Errors, NFER-NELSON, Windsor.

BRUST, M. (1997). Solución de ecuaciones de segundo grado: un análisis comparativo. Tesis para obtener el grado de licenciatura. ITAM. México.

CHEVALLARD, Y. (1989). Arithmétique, Algèbre, Modélisation. Étapes d'une recherche. Publications nº 16 de l'IREM Aix-Marseille.

CLARK, J.M., CORDERO, F., COTRILL, J., CZARNOCHA, B., DEVRIES, D.J., ST JOHN, D., TOLIAS, G., VIDAKOVIC, D. (1997). Constructing a schema: The case of the chain rule. Journal of Mathematical Behavior, 16, 345-364.

CLEMENT, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception, Journal for Research in Mathematics Education, 13, 16-30.

CORTÉS, A. (1992). Invariants operatoires dans le traitement des equations, Actes du colloque du programme cogniscience du CNRS ECCOS'92; Orsay.

DROUHARD, J. P. (1992). Les Écritures Symboliques de l'Algèbre Elémentaire, Tesis Doctoral, Université Denis Diderot, París 7.

DUBINSKY, E, y LEWIN, P . (1986). Reflective abstraction and mathematics education. The journal of Methematical Behaviour, 5, pp. 125-250.

DUBINSKY, E. (1985). The Constructive Aspects of Reflective Abstraction. Internal Publication. Purdue University.

DUBINSKY, E. (1989). The Case Against Visualization in School and University Mathematics. PME 13. Paris.

DUBINSKY, E. (1991) Reflective Abstraction in Advanced mathematical Thinking en Tall D. (Ed.) Advanced Mathematical Thinking. Kluwer Academic Publishers.

DUBINSKY, E. (1991). Constructive aspects of reflective abstraction in advanced mathematics. En L.P. Steffe (Ed.) Epistemological foundations of mathematical experience, pp. 160 – 202. New York. Springer Verlag.

DUBINSKY, E. (1996). El aprendizaje de los conceptos abstractos de la matemática avanzada. En Memorias de la X Reunión centroamericana y del Caribe sobre Formación de Profesores e investigación en Matemática Eduactiva, Puerto Rico, pp. 1-9.

DUBINSKY, E. y HAREL, G. (1992). The nature of the process conception of function. En G. Harel y E. Dubinsky (Eds.) The concept of function. Washington.

DUBINSKY, E. y SHWINGENDORF K. (1993). Constructing calculus concepts: Cooperation in a computer laboratory. Internal communication. Purdue University.

ELIAS CALLES P. (1997). Evaluación de algunos conceptos del álgebra en estudiantes universitarios a través de tres instrumentos. Tesis para obtener el grado de licenciatura. ITAM. México.

ENGLISH L. D. Y SHARRY P.V. (1996). Analogical Reasoning and the Development of Algebraic Abstraction, en Educational Studies in Mathematics, 30, pp. 135-157.

FILLOY E. and ROJANO T. (1984). From an Arithmetical to an Algebraic Thought. En Proceedings of the Sixth PME-NA Conference, Wisconsin, USA, PP. 51-56.

FILLOY E. and ROJANO T. (1989). Solving Equations: The Transition from Arithmetic to Algebra. For the Learning of Mathematics 9, 2, pp. 19-25.

FISHER, G. H. (1973). The linear logistic test model as an instrument in educational research. Acta Psychologica, 37, 359 –374.

GARCÍA NIETO, N. (1994). Apuntes del Curso de Doctorado. México. 1994

GARCÍA NIETO, N. (1987) Proyecto Docente e Investigador. Universidad Complutense. Madrid.

GASCÓN J. (1993). Desarrollo del conocimiento matemático y análisis didáctico: del patrón de análisis-síntesis a la génesis del lenguaje algebraico, Recherches en Didactique des Mathématiques, 13(3), 295-332.

GASCÓN J. (1994). Un nouveau modèle de l'algèbre élémentaire comme alternative à l'"arithmétique généralisée", Petit x, 37, 43-63.

GASCÓN J. (1998). La naturaleza prealgebraica de la matemática escolar, Educación matemática (por aparecer).

GINSSBURG H. y OPPER S. (1979). Piaget's Theory of Inetellectual Development: an Introduction. Prentice Hall, Engelwood Cliff.

GOMEZ, S y MORALES, J.L. (1997). A truncated method for solving parameter identification problems. Documento de trabajo. Depto. De Matemáticas, ITAM, México.

HATWELL, Y. (1992). A propósito de las nociones de asimilación y acomodación en los procesos cognoscitivos. En Ajuriaguerra, A. Psicología y Epistemología Genéticas. Nociones. México.

HERSCOVICS N. and LINCHEVSKI L. (1991). Pre-algebraic Thinking: Range of Equations and Informal Solution Processes Used by Seventh Graders Prior to any Instruction. En Proceedings of the Fifteenth PME Conference, Assisi, Italy, pp. 173-180.

KAPUT, J. (1983). Errors in translations to algebraic equations: Roots and implications, Focus on Learning Problems in Mathematics, 5, 63-78.

KAPUT, J. (1987). Algebra papers: A representational framework, en N. Bergeron, N. Herscovics y C. Kieran (eds.), Proceedings of the eleventh International Conference for the Psychology of Mathematics Education, pp.345-354.

KAPUT, J. (1996). ¿Una línea de investigación que sustente la reforma del álgebra? I y II, UNO. Revista de Didáctica de las Matemáticas, 9, 85-97 y 10, 89-103.

KIERAN C. (1984). Cognitive Mechanisms Underlying the Equation-Solving Errors of Algebra Novices, Proceedings of the VIII PME Conference, pp. 70-77.

KIERAN C. (1992). Multiple solutions to problems: The role of non-linear functions and graphical representations as catalysts in changing students' beliefs. Seminar presented at CINVESTAV, México.

KIERAN, C. y FILLOY, E. (1989). El aprendizaje del álgebra escolar desde una perspectiva psicológica, Enseñanza de las Ciencias, 7 (3), 229-240.

KUCHEMANN D. (1980). The understanding of generalized arithmetic (algebra) by secondary school children, PhD Thesis, University of London.

LABINOWITZ, E. (1987). Introducción a Piaget. Pensamiento, aprendizaje, enseñamza. Addison Wesley Iberoamericana. Delaware.

LÓPEZ, A. (1996). Construcción de la noción de variable algebraica en alumnos de nivel medio superior. Ms Sc. Thesis. Universidad Autónoma de Querétaro, México.

LOZANO D. (1998). El Concepto de Variable: Evolución a lo largo de la Instrucción Matemática. Tesis de licenciatura. ITAM, México.

MARKOVITS Z., EYLON B.S. and BRUCKHEIMER M. (1986). Functions Today and Yesterday. For the Learning of Mathematics 6, 2, pp. 18-28.

MARTINEZ, A. (1990). Problemas y pole´micas en torno al diagnóstico. Bordón, 42,1. Sociedad Española de Pedagogía.

MATZ M. (1982). Towards a Process Model for High School Algebra Errors. En Sleeman, D. y Brown, J.S. (Eds.). Intelligent Tutoring Systems, London, Academic Press.

PIAGET, J. (1970). L' épistemologie génetique. Paris. Presses Universitaires Francaises.

PIAGET, J. (1973). Introduction a l' épistemologie génetique (Vol I y II). Paris. Presses Universitaires Francaises.

PIAGET, J. (1975). L'équilibration des structures cognitives. Paris. Presses Universitaires de France.

PIAGET, J. (1977). Recherches sur l'abstraction réflechissante. Paris. Presses Universitaires de France.

PIAGET, J. (1980). Adaptation and Intelligence. University of Chicago Press. Chicago.

PIAGET, J. y GARCÍA R. (1982). Psicogénesis e Historia de la Ciencia. Siglo XXI Editores.

QUINTERO, R.; REYES, A.; TRIGUEROS, M. y URSINI, S. (1994). Misconceptions of variable, Proceedings of the XVI Annual Meeting of the PME-NA, p. I-185.

QUINTERO, R.; URSINI, S., REYES, A. y TRIGUEROS, M. (1995). Students Approaches to Different Uses of Variable. Proceeding of the IXX PME International Conference.

RADFORD, L. y GRENIER, M. (1996). Entre les choses, les symboles et les idées... une séquence d'enseignement d'introduction à l'algèbre, Revue des sciences de l'education, XXII/2, 253-276.

REGGIANI M. (1994). Generalisation as a Basis for Algebraic Thinking: Observations with 11-12 Year Old Pupils. Proceedings of the XVIII PME International Conference, pp. IV-97-104.

ROJANO, T. (1994). La matemática escolar como lenguaje. Nuevas perspectivas de investigación y enseñanza, Enseñanza de las Ciencias, 12 (1), 45-56.

RUÍZ, L. y RODRIGUEZ, J.L. (1994). La modelización funcional algebraica en la Enseñanza Secundaria: incidencia de la parametrización, Proyecto de Investigación (no publicado), Universidad de Jaén.

SFARD A. (1989), 'Transition from Operational to Structural Conception: The Notion of Function Revisited', Proceedings of the Thirteenth PME Conference, Paris, pp.151-158.

SFARD A. (1994). The Gains and the Pitfolls of Reification - The Case of Algebra, en Educational Studies in Mathematics, 26, pp. 191-228.

SFARD, A. y LINCHEVSKI, L. (1994): The gains and the pitfalls of reification. The case of algebra, Educational Studies in Mathematics, 26, 191-228.

SUTHERLAND R. (1987). A Longitudinal Study of the Development of Pupils' Algebraic Thinking in a Logo Environment, PhD Thesis, University of London Institute of Education.

THOMAS M. y TALL D. (1991). Encouraging Veratile Thinking in Algebra Using the Computer, Mathematics Education Research Center, University of Warwick, Coventry, U.K.

THORNDIKE E.L., COBB M.V., ORLEANS J.S., SYMONDS P.M., WALD E. y WOODYARD E. (1923). The Psychology of Algebra, The Macmillan Company, New York. Institute of Educational Research Teachers College, Columbia University.

TRIGUEROS M., REYES A., URSINI S. Y QUINTERO R. (1996). Diseño de un Cuestionario de Diagnóstico acerca del Manejo del Concepto de Variable en el Álgebra, en Enseñanza de las Ciencias, 14 (3), pp. 351-363.

TRIGUEROS M. y CANTORAL R. (1992). On Students Understanding of Variation. Proceedings of the PME-NA Conference Blacksburg, Virginia, pp. 176-182.

TRIGUEROS, M. y URSINI, S. (1998). Dificultades de los Estudiantes Universitarios Frente al Concepto de Variable, in Hitt, F. (Ed.), Investigaciones en Matemática Educativa II, Grupo Editorial Iberoamérica.

URSINI S. (1996). Creación de un Potencial para Trabajar con la Noción de Variable, en Hitt F. (Ed.), Investigaciones en Matemática Educativa, Grupo Editorial Iberoamérica, pp. 423-440.

URSINI S. y TRIGUEROS M. (1997). Understanding of different uses of variable: A study with starting college students. Proceedings of the XXI PME International Conference, Finland.

URSINI, S. (1994). Pupils' Approaches to Different Characterisations of Variable in Logo, PhD Thesis, University of London Institute of Education.

USISKIN Z. (1988). Conceptions of School Algebra and Uses of Variables, in Coxford A.F. y Shulte A.P. (Eds.), The Ideas of Algebra K-12, pp. 8-19.

VAN EGEN H. (1953). The Formation of Concepts, the Learning of Mathematics, en Febr H.F. (Ed.) The Learning of Mathematics: It's Theory and Practice, (Twenty-first Yearbook), Washington, D.C., National Council of Teachers of Mathematics, p. 69-98.

WAGNER S. (1981). An Analytical Framework for Mathematical Variables, Proceedings of the Fifth PME Conference, Grenoble, Francia, pp. 165-170.

WAGNER S. (1983). What are These Things Called Variables Variables?, Mathematics Teacher, October 1983, pp. 474-479.

ZHU, C., BYRD, H., LU, P. y NOCEDAL, J. (1994). LBFGS-B: Fortram subroutines for large-scale bound constrained optimization, Report NAMII, EECS, Department, Northwestern University, USA.

CONCLUSIONES

El propósito del presente trabajo consistió en primer lugar, en desarrollar un modelo matemático de medida, basado sobre un modelo logístico de dos parámetros, que ampliara la información obtenida de otros modelos matemáticos a través de la posibilidad de análisis de las estrategias de solución de los ítems por parte de los sujetos. Por otra parte, el modelo desarrollado intentaba ser útil como una herramienta de diagnóstico del conocimiento.

Dado este doble propósito, las conclusiones obtenidas pueden distinguirse en dos rubros. Aquéllas relacionadas con las características del modelo de medida y las que se derivan de su posibilidad de aplicación al diagnóstico.

Resultados relacionados con el modelo matemático de medida

El modelo matemático de medida diseñado en este trabajo constituye la contribución central del mismo.

El modelo propuesto es una generalización dinámica del modelo logístico de dos parámetros. Cuando decimos que el modelo es una generalización dinámica del modelo logístico nos referimos a un elemento que consideramos de suma importancia: un cambio de paradigma en las bases matemáticas sobre las que se diseñan los modelos de medida. En el caso del modelo presentado en este trabajo, su diseño se enmarca dentro de una perspectiva diferente a la tradicionalmente utilizada dentro del marco de la Teoría de los Tests. En lugar de utilizar la Estadística como herramienta de diseño, se emplea aquí la teoría de Ecuaciones Diferenciales y de Sistemas Dinámicos. La atención se centra en el cambio de la probabilidad de respuesta a un ítem o a una opción del ítem al variar la variable que determina el rasgo a medir y, por ello, se considera de naturaleza dinámica.

Durante muchos años se pensó que los modelos dinámicos eran más apropiados para estudiar fenómenos determinísticos y que por ello no tenían cabida en estudios estocásticos de la naturaleza de aquellos involucrados en la Teoría de Medida; sin embargo, al pasar de los años, se ha encontrado que la diferencia entre los fenómenos que se consideraban determinísticos y los no determinísticos es sutil y que el estudio de los fenómenos desde diferentes perspectivas puede aportar nuevos elementos de análisis y resultados importantes que desde otra perspectiva serían difíciles de encontrar. Los métodos de análisis de la teoría de las Ecuaciones Diferenciales y de los Sistemas Dinámicos permiten acercarse a las curvas solución del modelo desde ángulos nuevos que hacen posible encontrar información proveniente de las curvas de respuesta al ítem diferente de aquélla que se obtiene utilizando otros paradigmas.

La aplicación de una nueva perspectiva en la construcción del modelo que se presenta en este trabajo probó ser efectiva. El modelo provee resultados interesantes y nuevas formas de análisis útiles en la aplicación de la Teoría de Medida que se detallan a continuación.

En el modelo propuesto se modelan las curvas para la probabilidad de elección de las distintas opciones de respuesta al ítem. Se introducen tres parámetros para cada una de ellas.

Un parámetro que permite relacionar la elección de la opción con la accesibilidad de la misma para quien responde y que puede ser considerado como una medida de la complejidad cognitiva de la opción.

Un parámetro de máxima probabilidad de respuesta que permite distinguir entre opciones que se eligen con mayor frecuencia y aquéllas que se eligen con menor frecuencia, y que puede relacionarse con el conocimiento de los sujetos en el rasgo a medir en el caso de ítems diseñados con este fin, o con niveles de preferencia de elección en otros casos.

Un parámetro de interacción que puede considerarse como una medida de qué tanto afecta la presencia de las distintas alternativas de respuesta en la elección de cada una de ellas. El parámetro de interacción complementa la información que proporciona la medida de accesibilidad de la opción al proporcionar una medida del nivel de interferencia que la presencia de diferentes opciones viables causa en quien responde. Este parámetro puede tomar valores positivos o negativos; ello permite obtener información de suma importancia acerca del diseño mismo de las opciones de respuesta propuestas para los ítems.

En el modelo que aquí se presenta no se presupone ningún submodelo en cuanto a la forma en la que las complejidades cognitivas de los ítems influyen en el procesamiento de la información por parte de los sujetos que responden. Las opciones pueden calibrarse independientemente, en cuanto a la dificultad que involucran, a través del análisis de las respuestas de los estudiantes a ítems abiertos o más simples en los que se utilice el modelo logístico de dos parámetros en el análisis de la información. El valor del parámetro de accesibilidad así obtenido se incorpora al modelo y, a partir de las respuestas obtenidas al ítem completo, que contiene las opciones previamente calibradas, se puede ajustar el parámetro de interacción y el de las habilidades de los estudiantes en la variable a medir.

La relación con la complejidad cognitiva de las posibles opciones de respuesta al ítem no es única a este modelo. Existen, como se ha mencionado en el capítulo I del presente trabajo, modelos que incluyen la complejidad cognitiva de las alternativas de respuesta. La forma en la que esto se hace en el modelo que aquí se presenta es diferente a la empleada en otros modelos, ya que la complejidad cognitiva no se introduce de manera externa al modelo, mediante la asignación de pesos específicos, por ejemplo, para la elección de una u otra opción. En este modelo la complejidad cognitiva se toma en consideración de manera intrínseca y, por ello, es la misma estimación de los parámetros la que la determina.

El parámetro de máxima probabilidad de respuesta es algo único que diferencia a este modelo de otros existentes. Este parámetro proporciona información útil. A través de él es posible distinguir entre diversas poblaciones con las que es posible utilizar el ítem y es posible, por otra parte, distinguir categorías de ítems en términos de esa probabilidad máxima de elección de cada opción. Ambas posibilidades resultan importantes cuando se diseñan pruebas estandarizadas y en el diseño y aplicación de pruebas a la medida.

El parámetro de interacción es también único entre los modelos de medida. Este parámetro permite reconocer el efecto que la presencia de distintas alternativas de respuesta produce en quien responde y tiene implicaciones importantes. Desde el punto de vista cognitivo o psicológico, este parámetro permite establecer una relación más estrecha y natural con las teorías cognitivas, didácticas o psicológicas relacionadas con el dominio del cual provienen los ítems. Nuevamente esta relación no es exógena al modelo, sino que está incluida dentro del diseño mismo del ítem. Por otra parte, desde el punto de vista de la aplicación del modelo, al analizar la información proveniente del test, este parámetro puede relacionarse con estrategias de respuesta al ítem y con el nivel de solidez que tienen los conocimientos de los sujetos en el momento de resolver los tests diseñados para medirlos. La posibilidad de medir de alguna manera las dudas que surgen en quien responde al enfrentar una pregunta determinada permite adentrarse en la comprensión de la forma en la que los sujetos aprenden y en las estrategias que emplean al poner en práctica su conocimiento o al mostrar sus habilidades o intereses. Finalmente, desde la perspectiva misma del diseño de los ítems, este parámetro juega también un papel muy importante pues permite decidir acerca de qué tan pertinentes son las opciones seleccionadas para cada ítem o si alguna de ellas debe cambiarse.

El hecho de que el modelo que aquí se presenta pueda emplearse para validar el diseño de las opciones de respuesta al ítem hace de este modelo algo único. No existen en la literatura formas de determinar si el diseño de los ítems es adecuado y un paso en esta dirección es fundamental cuando se desea contar con ítems que proporcionen información efectiva acerca del rasgo a medir.

La solución al modelo propuesto proporciona diversas formas para las curvas de probabilidad de elección de las opciones de respuesta, sin necesidad de establecer un modelo específico para cada una de ellas ni de incluir información externa al modelo.

El análisis de la solución general del modelo permite distinguir dos familias de submodelos que juegan papeles diferentes y que son ambas importantes en su aplicación a situaciones educativas.

En una de ellas, los parámetros de accesibilidad de las opciones son todos positivos, indicando una tendencia al crecimiento para valores pequeños de la variable que representa el rasgo a medir. En este caso, el modelo es adecuado para situaciones en las que se pretenden estudiar las estrategias que los sujetos siguen al resolver el problema que el ítem plantea o para estudiar el comportamiento de las actitudes o preferencias de los sujetos que responden al ítem. Las curvas de elección de cada opción en este submodelo pueden cruzarse para un cierto valor del rasgo a medir, proporcionando información acerca del valor del mismo para el cual las estrategias o las actitudes cambian. Las curvas correspondientes a cada opción pueden crecer o decrecer en distintos intervalos de la variable que representa el rasgo a medir, dando cuenta de una gran diversidad de comportamientos posibles. En estos modelos se encuentra además la posibilidad de alcanzar puntos de equilibrio; es decir, valores estables para las probabilidades de elección de las distintas opciones de respuesta al ítem, lo cual proporciona información acerca de posibles situaciones en las que, una vez alcanzado el punto de equilibrio, aunque la variable que representa el rasgo a medir aumente, las probabilidades de elección no cambian. En términos de estrategias de solución de problemas, un punto de esta naturaleza podría indicar el valor de la variable independiente para el cual los sujetos alcanzan madurez en el conocimiento adquiriendo flexibilidad suficiente para tomar decisiones estables respecto a la estrategia a seguir. Dependiendo de la naturaleza del problema, en el caso de tests diseñados para medir actitudes o preferencias, estos puntos pueden dar información para distinguir entre poblaciones con preferencias o actitudes claramente definidas.

El otro submodelo que surge del análisis general del modelo se obtiene cuando uno de los parámetros de accesibilidad del ítem es positivo y el otro negativo. Este caso modela con mayor precisión situaciones en las que los ítems miden niveles de conocimiento en algún dominio específico. El punto donde las curvas de elección de las opciones se cruzan, cuando esto sucede, podría interpretarse como una medida de la dificultad del ítem, ya que indica para qué valores de la variable que representa el rasgo a medir, la probabilidad de elección de la opción correcta prevalece sobre la probabilidad de elección de la opción incorrecta. En este caso, es posible obtener curvas que para ciertos intervalos de la variable dependiente son crecientes y para otros son decrecientes, lo que proporciona asimismo información útil para el diagnóstico. En particular, cabe mencionar que el modelo predice la aparición de situaciones en las que tanto la probabilidad de elección de la opción correcta como la de la opción incorrecta decrecen para un intervalo dado de valores de la variable independiente. Esta posibilidad, que desde el punto de vista educativo puede relacionarse con situaciones predichas por algunas teorías cognitivas, indica intervalos en los que la duda o la confusión son prevalentes y que son muy útiles para las aplicaciones con fines de diagnóstico.

La solución completa del modelo involucra la necesidad de encontrar las condiciones iniciales para las cuales los datos obtenidos de la aplicación del modelo se ajustan mejor a una de las posibles curvas de probabilidad de elección de cada opción. Las condiciones iniciales, que resultan a consecuencia de la estimación de los parámetros del modelo, complementan la información que se obtiene de los parámetros antes mencionados ya que permiten caracterizar a la población con la que se utiliza un test dado en términos del rasgo a medir. Las condiciones iniciales estimadas serán diferentes en el caso de poblaciones cuyo conocimiento, por ejemplo, sea débil en lo que se desea medir, que las correspondientes a poblaciones más adelantadas o con un conocimiento más sólido. Esta información es igualmente de gran valor cuando el modelo se emplea con fines de diagnóstico o cuando se pretende usar los ítems en test estandarizados o en pruebas a la medida.

El tipo de estimación de los parámetros que se propone en este modelo es similar a la que se utiliza en los modelos multicomponente de rasgo latente. Sin embargo, como ya se mencionó, en estos últimos, es necesario calibrar la dificultad y la habilidad para cada componente del modelo desde el punto de vista cognitivo, utilizando información externa al mismo. Al igual que el modelo multicomponente, el modelo que se presenta en este trabajo requiere del conocimiento de las respuestas al ítem total y de las respuestas a las opciones por separado para estimar los parámetros. Sin embargo, a diferencia de estos modelos, en el que aquí se propone, no se requiere de la definición de una dependencia secuencial de las componentes de la tarea a medir, sino de principios más generales de abstracción, ligados a la forma en la que se trabaja en una disciplina específica.

La naturaleza dinámica del modelo es la fuente de su riqueza, de las diferentes posibilidades de interpretación de los parámetros, de la flexibilidad en la obtención de diferentes formas para las curvas de elección de cada opción de respuesta y de la diversidad de situaciones a las que el modelo se puede aplicar. La misma naturaleza dinámica del modelo es asimismo el origen de algunas de las dificultades que presenta su aplicación. El modelo que se presenta en este trabajo es un modelo dinámico no lineal. No existe, por ello, la posibilidad de solución del modelo en forma analítica y, por tanto, su aplicación debe hacerse en forma diferencial, para obtener después una solución numérica que represente los datos obtenidos.

Dada la forma del modelo y nuevamente por su naturaleza no lineal, la estimación de los parámetros del modelo requiere de la introducción de métodos numéricos en los que se utilizan conceptos matemáticos avanzados. Esta es quizás la mayor limitación del modelo, el diseño de un método específico para la estimación de los parámetros del modelo es una tarea pendiente para investigaciones futuras. La forma en la que se hizo la estimación de los parámetros en este trabajo es laboriosa y poco eficiente, pero a través del estudio de la literatura correspondiente a la estimación de los parámetros de modelos del tipo del presentado en este trabajo, es posible sugerir que sería más apropiado desarrollar una estrategia de estimación de parámetros basada en el uso de métodos inversos desde el inicio del proceso de estimación. Una estrategia de esta naturaleza haría posible desarrollar formas más accesibles y eficientes para encontrar los parámetros, eliminando, incluso, la necesidad de calibrar los parámetros mediante una doble aplicación de los ítems.

En síntesis, el modelo de medida propuesto en este trabajo parece proveer una alternativa matemática dentro de la Teoría de los Tests, útil particularmente en el estudio del diseño de las opciones de respuesta para ítems de opción múltiple y en el diseño de tests en los que las diferentes alternativas de respuesta proporcionen información diagnóstica para una gran diversidad de situaciones educativas.

Resultados relacionados con la posibilidad de aplicación del modelo matemático de medida a situaciones educativas

Con el fin de poner a prueba el modelo, de mostrar la posibilidad de utilizarlo conjuntamente con una teoría cognitiva y de mostrar su utilidad en el marco del diagnóstico, en la tercera parte del presente trabajo se introduce una discusión acerca del papel del diagnóstico en la educación y se presenta un modelo educativo, basado en las ideas de Piaget, que es adecuado para el diagnóstico del aprendizaje de conceptos matemáticos. El modelo matemático se utiliza en conjunción con el modelo educativo en el análisis de la información obtenida a partir de las respuestas de los sujetos a un test específico.

La puesta a prueba del modelo matemático de medida resultó de interés en sí misma y no únicamente en términos de la validación del modelo de medida. A continuación se discuten algunos de los resultados de la aplicación del modelo a una situación educativa, tanto desde la perspectiva de la validación del modelo, cuanto de la situación didáctica en sí misma, que es conveniente resaltar.

Uno de los objetivos del presente trabajo consistió, como se ha mencionado, en el diseño de un modelo matemático de medida que pudiera ser empleado conjuntamente con alguna teoría cognitiva o educativa que diera un sentido específico a la interpretación de los resultados obtenidos de la aplicación del modelo de medida. La aplicación del modelo desarrollado a una situación específica en el contexto del aprendizaje de las matemáticas permite constatar que, efectivamente, la información obtenida a partir de los parámetros del modelo es susceptible de una interpretación en términos de un modelo de aprendizaje.

El análisis de la información en términos de la teoría de aprendizaje se realiza en este modelo a partir del análisis de las características de las respuestas de los sujetos a los ítems que componen el test, pero, a diferencia de otros modelos, de los que se ha hablado en el capítulo I de este trabajo, en este caso, es posible obtener mayor información, puesto que se cuenta con la posibilidad de interpretar la elección de las alternativas de respuesta de cada uno de los ítems. Así, las respuestas de los sujetos pueden organizarse en clases o conjuntos que pueden emplearse con mayor seguridad en las aplicaciones de índole diagnóstica, en la evaluación en términos generales, en el diseño de la instrucción en general y en el diseño de la instrucción personalizada.

En este trabajo, el análisis de la información obtenida del cuestionario se llevó a cabo, por una parte, a través de la interpretación de los valores de los parámetros para las opciones de respuesta de cada ítem de opción múltiple, obtenidos de la aplicación del modelo de medida y, por otra, de la conjunción de estos resultados con un modelo diseñado específicamente sobre las bases de una teoría del aprendizaje matemático, para entender la forma en la que los estudiantes construyen el concepto de variable en el álgebra elemental.

Aunque ya se ha mencionado anteriormente conviene repetir que si bien el tipo de estimación de los parámetros que se propone en este modelo es similar a la que se utiliza en los modelos multicomponente de rasgo latente, en estos últimos, es necesario calibrar los parámetros de dificultad y de habilidad para cada componente del modelo desde el punto de vista cognitivo, utilizando información externa al modelo. Esta misma información provee las bases para interpretar los resultados de su aplicación. En el modelo considerado en este trabajo no es necesario incluir la teoría cognitiva para calibrar ninguno de los parámetros. El modelo cognitivo se utiliza en el diseño de las diferentes opciones de respuesta alternativas para cada uno de los ítems y en el análisis de los resultados que se obtienen de la aplicación del test. Análogamente al modelo multicomponente, el modelo de medida que aquí se presenta requiere de respuestas al ítem total y de las respuestas a las opciones por separado para calibrar los parámetros. A diferencia del modelo multicomponente, en el modelo que aquí se propone no se requiere de la hipótesis de la existencia de una dependencia secuencial, en términos cognitivos, de las componentes de la tarea a medir, sino que se parte de principios más generales de abstracción, ligados a la forma en la que se trabaja en una disciplina dada. En este mismo aspecto, el modelo que se presenta en este trabajo es diferente al modelo general de rasgo latente de Embreston.

El modelo de medida que se diseñó en este trabajo permite describir a los ítems a través del análisis de los distractores y de la opción de respuesta correcta. La aplicación del modelo a una situación particular permite concluir que el modelo da cuenta de manera adecuada la interacción que puede darse en la mente del sujeto que responde cuando se le presentan preguntas de opción múltiple.

En la mayor parte de los ítems que fueron analizados en términos del modelo de medida fue posible localizar un intervalo de valores de la variable independiente, que representa el rasgo a medir, para el cual las probabilidades de elección de las dos opciones en estudio decrecen, lo cual puede interpretarse, como ya se había mencionado anteriormente, como una posible transición entre un nivel de conocimiento y otro en términos del modelo de aprendizaje. Este resultado es de interés desde el punto de vista de la aplicación de los modelos de medida a situaciones educativas, dado que varias teorías cognitivas y de aprendizaje postulan la ocurrencia de períodos de incertidumbre en la construcción del conocimiento. La posibilidad de detectarlos explícitamente abre muchas posibilidades a la aplicación de estos modelos al diagnóstico.

El punto en el que las curvas de probabilidad de elección de las opciones en estudio para cada ítem se cruzan, puede asociarse, desde la perspectiva de la educación a un indicador de la dificultad del contenido del ítem global. Esta información puede emplearse en el diseño de tests estandarizados o en el diseño de tests dirigidos a la medición de rasgos muy específicos en poblaciones con características predeterminadas.

Otro resultado de la aplicación del modelo que conviene destacar es que su aplicación al cuestionario diagnóstico pone en evidencia la interacción entre las alternativas. Este resultado es importante en el contexto del debate sobre el formato de los ítems. Se cuestiona, en este ámbito, la posibilidad de obtener información compleja a partir de los resultados que se obtienen de los ítems de opción múltiple y el hecho de que el proporcionar a los sujetos la respuesta correcta, en el caso de pruebas de conocimiento, provee información a quien responde y oscurece el análisis de los resultados. En la aplicación del modelo de medida que se diseñó en este trabajo al cuestionario acerca del concepto de variable se obtiene evidencia que refuta estos argumentos. Se muestra que los ítems de opción múltiple pueden diseñarse para tareas de diversa naturaleza y que la presencia de la respuesta correcta dentro del ítem no necesariamente induce su elección. Si los distractores del ítem se diseñan correctamente, la aplicación del modelo al cuestionario provee evidencias de que, incluso sujetos que en un formato abierto responden correctamente a un ítem, cuando enfrentan alternativas plausibles y su conocimiento de la materia en cuestión no es sólido, eligen la opción incorrecta o dejan de responder a causa de las dudas que les surgen. Esto sugiere que el formato de opción múltiple puede proporcionar incluso información que sería imposible obtener de otro tipo de formatos: la posibilidad de confusión de quien responde cuando sus conocimientos no son tan firmes como es deseable. Esta información es de gran valor cuando la medida se lleva a cabo con fines de diagnóstico.

En la aplicación del cuestionario diagnóstico se pone de manifiesto asimismo que las características particulares del diseño del ítem, es decir el diseño de las posibles opciones de respuesta y la forma en la que estas se redactan pueden influir en la dificultad del ítem de dos maneras. Por una parte en cuanto al nivel de conceptualización involucrado en la respuesta del ítem y reflejado en el tipo de abstracción que es necesaria para elección de la opción correcta, cuando los ítems intentan medir desempeño académico, o en la estrategia empleada en otros casos. Por otra parte, en cuanto al nivel de interacción que las diferentes alternativas de respuesta que se presentan en el ítem ejercen sobre quien lo responde.

Una ilustración de lo primero se encuentra en ítems que versan sobre el conocimiento acerca del concepto de variable en el álgebra. Un ítem en el que se pide encontrar la solución de la ecuación $3x+4=2$, que puede ser resuelto mediante procedimientos aritméticos o algebraicos, corresponde a un nivel de conceptualización inferior a otro que involucra la solución de una ecuación como $7x-5=4+2x$, que involucra la manipulación algebraica y una mayor profundización en lo que el concepto de variable significa.

En cuanto a lo segundo, el uso de las respuestas más comunes que pueden asociarse a estrategias específicas de razonamiento, registradas en investigaciones y en entrevistas con alumnos, proporcionan distractores adecuados con los cuales se puede medir el nivel de interacción que la presencia de los distractores origina.

La aplicación del modelo de medida a un cuestionario específico pone de manifiesto la posibilidad de utilizar el modelo en situaciones educativas concretas. Si bien el modelo fue probado únicamente en la situación específica relacionada con el conocimiento de los estudiantes del concepto de variable, es posible inferir que las posibilidades de aplicación del modelo no están limitadas a este concepto; ni siquiera están limitadas al contexto de las matemáticas. El modelo puede tener un ámbito de aplicación mayor. Sin embargo, es importante destacar que la interpretación de los resultados del modelo está ligada a la teoría cognitiva o al modelo de aprendizaje específicos y por ello, la aplicación del modelo de medida a otras situaciones o contextos requería, en cada caso, de la utilización de modelos educativos o de teorías cognitivas desarrolladas específicamente para el dominio de conocimiento en el que se desea aplicar el modelo.

El análisis de los ítems empleados para probar el modelo muestra que los resultados obtenidos de la aplicación del modelo al cuestionario acerca de la variable son satisfactorios. Esto permite un cierto optimismo en cuanto a las posibilidades de aplicación del modelo a diversas situaciones.

La aplicación del modelo requirió del diseño de un marco teórico pertinente para el análisis de datos relativos a la construcción del conocimiento matemático, misma que fue aplicada al caso particular del concepto de variable en el álgebra elemental. El diseño del marco teórico puede considerarse una aportación importante de este trabajo al estudio de la didáctica de las matemáticas.

El modelo de diagnóstico desarrollado con base en el trabajo de Dubinsky y sus colaboradores proporciona una estructura jerárquica y anidada en la que el análisis de las tareas a resolver permite la identificación del tipo de conceptualización de los alumnos frente a un concepto en particular y la interpretación de las respuestas equivocadas y de las estrategias utilizadas. Al incorporar esta información al modelo matemático que se ha propuesto en este trabajo, es posible identificar, mediante la información proporcionada por los parámetros de elección de las opciones de cada uno de los ítems y de la variable que representa el rasgo a medir, si un alumno determinado se encuentra en un nivel de la jerarquía, en otro e o en un estado de transición.

La descomposición del concepto de variable probó ser una herramienta de utilidad en el diseño del cuestionario y en el análisis de la información proporcionada por la respuesta a los ítems. A través del análisis de los datos obtenidos de la aplicación del cuestionario se obtiene información interesante, no únicamente en términos de lo que se puede inferir a través de la interpretación de los parámetros y las variables de los sujetos obtenidas, sino en términos de la conceptualización del concepto de variable de los sujetos que resolvieron el test. Puede inferirse de los resultados obtenidos que la capacidad de manipulación, interpretación y simbolización de los diferentes usos del concepto de variable en el álgebra es pobre a pesar de que la muestra de sujetos que se utilizó en el estudio es de estudiantes que inician los estudios universitarios. El análisis de cada uno de los ítems muestra que la mayor parte de los estudiantes involucrados en este estudio responde a las preguntas del cuestionario a un nivel de acción. Por otra parte, puesto que la interacción de las alternativas incorrectas de respuesta es grande para muchos de los ítems, puede concluirse que el conocimiento de estos estudiantes es además poco sólido.

Es cierto que a partir de la aplicación del modelo a un sólo instrumento de medida, como el que aquí se presenta, y a una muestra de sujetos pequeña, no es posible concluir en forma contundente el nivel en el que se encuentra cada uno de los estudiantes ni llegar a conclusiones categóricas en cuanto a su conocimiento. Los resultados obtenidos sugieren una cierta tendencia en las respuestas de los estudiantes. Por ello sería interesante aplicar el modelo en situaciones diversas y con muestras mayores de respondentes para aumentar la precisión en los resultados.

La aplicación del modelo a la enseñanza del concepto de variable resultó interesante en sí misma. El uso de la descomposición genética, como apoyo al diseño de instrumentos de diagnóstico, resultó efectiva. La tarea involucrada en la descomposición de un concepto matemático o de un conjunto de conceptos es, sin embargo, una tarea ardua que demanda de investigación y de trabajo en equipo. Una pregunta que se plantea es si en otras áreas del conocimiento sería posible desarrollar herramientas de esta naturaleza para sustentar el diseño de instrumentos sobre bases teóricas sólidas.

En cuanto a las posibles aplicaciones que pueden hacerse del modelo de medida puede decirse que el modelo que se diseñó en este trabajo es versátil. Sus posibilidades de aplicación son múltiples. La aplicación a la que en este trabajo se le ha dado un papel preponderante es al diagnóstico del conocimiento matemático, pero también podría usarse para probar distractores en cuanto a su complejidad cognitiva, a su dificultad y a su nivel de interacción para estudiar la posibilidad de utilizarlos en pruebas para diferentes edades o en situaciones diversas de instrucción. Puede, como ya se mencionó, utilizarse además, para probar la efectividad de los distractores. Esta aplicación, que es de gran utilidad en el diseño de tests, hace al modelo único entre los modelos actuales: ninguno de los modelos existentes permite esta posibilidad de aplicación. El modelo puede aplicarse a diferentes situaciones educativas y psicológicas así como a la creación de bancos de ítems con distintas posibilidades de elección de distractores, es decir, en los que los distractores tengan características específicas determinadas por los parámetros de cada una de las alternativas posibles elegidas como opciones de respuesta.

En la solución matemática del modelo presentado en este trabajo surgen, como se ha mencionado, dos familias o submodelos que se distinguen por el signo del parámetro de accesibilidad de las opciones. En este trabajo se aplicó el modelo de medida únicamente a la medida del conocimiento de los sujetos, es decir se hizo uso exclusivamente de uno de los posibles submodelos: aquél en el que los parámetros de accesibilidad de los ítems tienen signos diferentes. Sería muy interesante hacer aplicaciones del modelo a situaciones en las que el otro submodelo sea pertinente, es decir, a situaciones en las que el objeto de estudio sea el análisis de las estrategias de solución de problemas, de actitudes o de preferencias de los sujetos. Estas aplicaciones complementarían la información actual acerca de las posibilidades de aplicación del modelo que se han tratado en este trabajo y de su versatilidad.

Podemos concluir, a partir de las consideraciones anteriores, que el modelo de medida propuesto en este trabajo es original y de gran riqueza tanto en términos de las aplicaciones en las que se puede emplear, cuanto a la información que se puede obtener de su aplicación.

Prospectivas de investigación

El modelo presentado en este trabajo contribuye a la investigación en educación en el área de la Teoría de Medida. A lo largo del trabajo, como se ha mencionado hasta aquí, pueden encontrarse aportaciones relevantes e interesantes al diseño de ítems y al análisis de la información que se obtiene de la aplicación del modelo a tests dentro del ámbito educativo.

Las contribuciones de un trabajo de investigación no consisten únicamente en los resultados directos que éste aporta al área dentro de la cual se inscribe, sino en la posibilidad de apertura de nuevas preguntas y posibilidades de investigación que en el presente trabajo no fueron tratadas.

El diseño del modelo de medida propuesto en este trabajo y su aplicación en el ámbito de la educación pone de manifiesto las siguientes prospectivas de investigación:

En el diseño de cualquier modelo de medida, la estimación de los parámetros juega un papel muy importante. En el trabajo que aquí se presenta se sugiere un método para estimar los parámetros que no es muy eficiente. Es necesario hacer investigación en este aspecto y desarrollar métodos numéricos adecuados a la naturaleza del modelo y que hagan más fácil su aplicación. Esta investigación queda enmarcada dentro del ámbito de las matemáticas. En particular, dentro del área de los métodos numéricos. Es posible que la dirección que deba tomarse sea el partir desde el inicio de la estimación del uso de métodos inversos y dejar de lado, si es posible, la primera calibración de los parámetros del modelo.

En el capítulo III del trabajo se muestra que los ítems de opción múltiple pueden diseñarse para tareas de diversa naturaleza y que la presencia de la respuesta correcta dentro del ítem no necesariamente induce su elección. Si los distractores del ítem se diseñan correctamente, la aplicación del modelo a cuestionarios específicos puede proveer evidencias del tipo de interacción que la presencia de opciones incorrectas ejerce en quien responde. De aquí surge la posibilidad de emplear el modelo de medida propuesto, conjuntamente con algún otro modelo de medida para hacer un estudio crítico y riguroso acerca de las posibilidades y limitaciones de los diferentes formatos de diseño de los ítems. Este tipo de investigación contribuiría a aclarar si efectivamente los ítems de opción múltiple pueden proporcionar información que sería imposible obtener de otro tipo de formatos y a profundizar en los puntos relevantes en la discusión sobre el formato de los ítems que se mencionó en el capítulo I.

Como se ha reiterado en varias ocasiones, de la solución del modelo de medida propuesto en este trabajo surgen dos submodelos que son aplicables a situaciones diferentes. Sería conveniente analizar las posibilidades de aplicación del modelo a casos en los que los parámetros de accesibilidad de las opciones de respuesta sean todos positivos, así como estudiar críticamente las aportaciones del modelo en este tipo de situaciones.

En el presente trabajo se aplicó el modelo únicamente al análisis de la información de dos alternativas de respuesta para el ítem. Sin embargo, en el capítulo II se presentó el modelo de forma general y es posible considerar su aplicación para el análisis de la información proporcionada por cada una de las opciones de respuesta a un ítem particular. Sería pertinente aplicar el modelo a ítems que incluyan un número mayor de opciones de respuesta para obtener, por una parte más información acerca de las respuestas de los sujetos y por otra para analizar las posibilidades y limitaciones del modelo en este contexto.

Las posibilidades de aplicación del modelo de medida son muy diversas. En este ámbito las prospectivas de investigación son múltiples. Entre ellas pueden mencionarse la aplicación del modelo al aprendizaje de otros conceptos matemáticos, la aplicación del modelo al aprendizaje de otras disciplinas y la aplicación del modelo en el ámbito de la psicología. Todas estas aplicaciones requerirían, además de la puesta en práctica del modelo, del desarrollo de marcos teóricos cognitivos específicos a cada una de ellas.

La aplicación del modelo al diseño de tests cuya finalidad no sea el diagnóstico, por ejemplo tests estandarizados o tests diseñados a la medida constituye otra área en la que la investigación es pertinente.

Por último, sería interesante llevar a cabo un estudio comparativo entre el modelo que se presenta en este trabajo y otros modelos existentes utilizando los mismos ítems y el mismo modelo cognitivo para analizar con mayor profundidad las posibilidades de obtención de información y las limitaciones de éste y otros modelos de medida. ◆

ANEXO I

RESULTADOS DE ALGUNAS ITERACIONES DEL ALGORITMO DE AJUSTE DE PARÁMETROS

JLMSG#52931/

Iteration no. 1

row	a	b	c	d	xo	Yo
1	1.472	1.845	-1.157	2.682	0.252	0.682
2	0.875	0.450	-0.635	0.174	0.050	0.589
3	1.511	2.001	-1.049	1.950	0.490	0.469
4	0.953	1.987	-1.381	1.714	0.350	0.599
5	2.500	1.865	-2.500	1.032	0.265	0.590
6	1.566	-0.867	-0.994	1.213	0.330	0.679
7	2.500	0.997	-1.326	2.336	0.031	0.872
8	1.676	1.763	-1.117	2.876	0.135	0.699
9	1.192	1.684	-1.270	1.999	0.243	0.703
10	1.667	1.734	-1.016	0.982	0.541	0.384
11	1.377	1.998	-1.713	1.001	0.134	0.753
12	1.747	2.000	-0.999	1.214	0.436	0.528
13	1.749	0.879	-1.562	2.002	0.122	0.711
14	1.826	1.431	-1.277	2.050	0.546	0.384
15	1.423	1.674	-1.218	1.998	0.532	0.230
16	1.283	2.674	-0.887	1.551	0.572	0.261
17	1.566	2.998	-2.500	-2.543	0.000	0.980
18	0.828	1.897	-1.675	1.123	0.341	0.650
19	1.637	0,568	-0.833	1.998	0.332	0.623
20	1.417	2.453	-1.570	2.643	0.437	0.476
21	1.569	1.864	-1.504	2.209	0.157	0.770
22	1.304	1.692	-0.928	1.998	0.052	0.865
23	1.558	2.955	-1.558	2.000	0.555	0.370
24	1.555	1.689	-1.267	1.879	0.092	0.882
25	1.817	2.557	-2.500	2.978	0.001	0.992
26	1.529	1.784	-1.472	3.001	0.352	0.607
27	1.158	-1.592	-1.326	2.102	0.335	0.585
28	0.867	-2.756	-2.129	-2.221	0.023	0.907
29	1.396	2.003	-1.006	1.638	0.130	0.730
30	1.449	2.106	-1.396	1.793	0.369	0.531

JLMSG#52931/

Iteration no. 50

row	a	b	c	d	Xo	Yo
1	1.440	1.712	-0.873	2.493	0.284	0.405
2	0.800	0.731	-0.617	0.438	0.223	0.500
3	1.372	0.998	-0.729	1.960	0.421	0.409
4	0.921	1.362	-0.964	1.701	0.301	0.472
5	1.222	1.752	-1.227	0.882	0.257	0.582
6	1.000	-0.860	-0.792	1.004	0.289	0.666
7	2.173	0.799	-2.028	2.121	0.047	0.805
8	1.001	1.672	-0.990	2.330	0.156	0.692
9	0.932	1.042	-1.923	1.883	0.241	0.752
10	1.517	1.696	-1.005	0.954	0.502	0.397
11	1.293	1.855	-0.777	0.556	0.279	0.645
12	1.020	1.046	-0.829	1.207	0.120	0.475
13	1.020	0.805	-0.624	1.743	0.422	0.709
14	1.573	1.352	-0.963	1.950	0.515	0.400
15	0.921	1.489	-1.502	1.672	0.500	0.271
16	0.900	2.000	-1.231	1.507	0.513	0.303
17	1.530	2.000	-0.744	2.488	0.125	0.799
18	0.771	1.569	-0.886	1.006	0.349	0.572
19	1.213	0.588	-2.174	1.471	0.301	0.600
20	1.402	2.005	-1.002	2.329	0.427	0.510
21	1.213	1.471	-0.751	2.000	0.197	0.760
22	1.210	1.684	-1.552	1.927	0.059	0.863
23	1.432	2.032	-1.300	1.051	0.430	0.442
24	1.023	1.725	-0.883	1.523	0.088	0.442
25	1.674	1.782	-1.348	3.121	0.215	0.723
26	1.413	1.627	-1.253	2.548	0.176	0.614
27	0.782	-1.587	-1.150	2.044	0.223	0.585
28	1.000	-2.542	1.979	-2.054	0.107	0.797
29	1.396	1.884	-0.442	1.606	0.114	0.869
30	1.443	2.060	-1.067	1.584	0.365	0.545

JLMSG#52931/

Iteration no. 100

row	a	b	c	d	xo	yo
1	1.372	1.666	-0.629	1.402	0.266	0.521
2	0.779	0.993	-0.582	0.598	0.180	0.618
3	1.041	1.872	-0.838	1.647	0.401	0.541
4	0.954	1.256	-0.841	1.628	0.331	0.557
5	1.021	0.686	-0.500	0.813	0.258	0.521
6	0.722	-0.515	-0.866	0.972	0.251	0.603
7	1.850	0.681	-0.790	1.851	0.147	0.791
8	1.000	1.533	-1.537	2.142	0.163	0.614
9	0.950	0.962	-0.862	1.802	0.200	0.754
10	1.462	1.620	-1.051	0.371	0.305	0.412
11	1.217	1.173	-0.877	0.497	0.295	0.644
12	1.056	0.875	-0.650	1.101	0.352	0.502
13	0.679	0.782	-0.431	0.883	0.117	0.699
14	1.333	1.260	-0.379	1.042	0.372	0.513
15	0.920	1.205	-0.700	1.612	0.443	0.337
16	0.654	1.884	-0.986	0.828	0.257	0.412
17	1.183	2.275	1.437	-2.937	0.327	0.601
18	0.742	1.222	-0.880	0.739	0.398	0.555
19	0.928	0.564	-1.063	1.921	0.105	0.746
20	1.387	1.726	-0.997	1.987	0.050	0.618
21	1.177	1.422	-0.636	1.553	0.210	0.646
22	1.043	1.612	-0.738	1.652	0.098	0.825
23	1.204	1.607	-1.005	0.862	0.323	0.498
24	1.005	1.600	-0.827	1.437	0.079	0.844
25	1.872	1.835	-1.021	2.021	0.347	0.618
26	0.973	1.585	-1.003	1.977	0.124	0.667
27	0.780	-1.638	-1.072	1.023	0.208	0.560
28	0.543	-2.103	-1.787	1.552	0.301	0.862
29	1.383	1.522	-0.627	0.793	0.092	0.726
30	1.440	1.499	-0.698	0.873	0.257	0.602

JLMSG#52931/

Iteration no. 150

row	a	b	c	d	xo	yo
1	1.352	1.567	-0.497	1.137	0.259	0.722
2	0.684	1.007	-0.325	0.787	0.181	0.801
3	0.423	1.320	-0.556	1.622	0.372	0.591
4	0.770	1.223	-1,061	1.346	0.367	0.548
5	0.821	1.230	-0.881	0.600	0.294	0.512
6	0.651	-0.30	-0.672	0.491	0.186	0.785
7	1.230	0.623	-0.755	1.231	0.144	0.792
8	0.852	1.428	-0.923	1.412	0.304	0.518
9	0.785	0.956	-0.628	1.446	0.176	0.700
10	1.321	1.412	-0.921	0.251	0.244	0.705
11	0.437	1.707	-0.650	0.475	0.311	0.652
12	0.993	0.661	-0.514	1.006	0.196	0.724
13	0.665	0.367	-0.332	0.798	0.102	0.701
14	1.312	1.258	-0.227	1.157	0.317	0.612
15	0.918	1.374	-0.968	0.733	0.430	0.554
16	0.388	1.767	-0.811	0.724	0.147	0.572
17	1.523	1.003	-1.732	-0.921	0.405	0.489
18	0.740	0.899	-0.893	0.572	0.427	0.530
19	0.650	1.003	-0.827	1.610	0.090	0.847
20	1.221	0.998	-0.647	1.447	0.078	0.893
21	0.524	1.406	-0.988	1.206	0.308	0.559
22	0.922	1.230	-0.621	1.278	0.110	0.807
23	1.300	1.500	-0.852	0.809	0.286	0.673
24	0.663	1.703	-0.820	1.425	0.053	0.815
25	1.197	1.822	-1.003	1.784	0.349	0.600
26	0.900	1.172	-0.999	1.534	0.078	0.697
27	0.681	-1.479	-1.652	0.992	0.199	0.592
28	1.725	-2.198	1.584	1.031	0.393	0.507
29	1.372	1.501	-0.676	0.513	0.107	0.886
30	1.372	1.576	-0.752	0.684	0.239	0.734

JLMSG#52931/

Iteration no. 200

row	a	b	c	d	xo	yo
1	1.253	1.563	-0.517	0.918	0.230	0.691
2	0.722	1.006	-0.300	0.990	0.116	0.820
3	0.242	1.349	-0.527	1.721	0.331	0.662
4	0.391	1.215	-0.577	1.200	0.400	0.592
5	0.628	0.992	-1.000	0.586	0.342	0.567
6	0.539	-0.150	-0.982	0.485	0.092	0.791
7	0.772	0.562	-0.945	1.005	0.144	0.849
8	0.874	1.291	-0.501	0.812	0.321	0.655
9	0.781	0.931	-0.601	1.077	0.145	0.810
10	1.049	1.292	-0.514	0.240	0.236	0.700
11	1.063	1.550	-0.310	0.399	0.432	0.550
12	0.437	0.628	-0.512	0.905	0.200	0.786
13	0.482	0.344	-0.247	0.852	0.091	0.700
14	1.015	1.245	-0.215	1.012	0.235	0.652
15	0.500	0.827	-0.879	0.402	0.412	0.570
16	0.386	1.386	-0.472	0.721	0.152	0.566
17	2.630	-.002	-1.237	1.200	0.412	0.470
18	0.728	0.893	-0.770	0.351	0.456	0.393
19	0.594	0.563	-0.852	1.290	0.110	0.861
20	1.113	0.998	-0.421	1.000	0.115	0.806
21	0.474	1.371	-0.435	1.182	0.300	0.591
22	0.917	1.157	-0.532	1.022	0.075	0.832
23	1.314	1.494	-0.983	0.805	0.220	0.680
24	0.712	1.699	-0.884	1.172	0.020	0.800
25	1.568	1.021	-0.889	1.022	0.151	0.800
26	0.895	1.052	-0.823	1.113	0.096	0.721
27	0.700	-1.146	-1.004	0.952	0.192	0.500
28	1.368	1.123	-1.023	1.005	0.412	0.532
29	1.323	1.472	-0.521	0.500	0.100	0.795
30	1.078	1.336	-0.799	0.568	0.237	0.682

JLMSG#52931/

Iteration no. 243

row	a	b	c	d	xo	Yo
1	1.210	1.520	-0.500	0.850	0.200	0.600
2	0.690	0.980	-0.231	0.985	0.103	0.677
3	0.210	1.270	-0.450	1.500	0.290	0.660
4	0.330	1.170	-0.540	1.150	0.351	0.550
5	0.660	0.931	-0.970	0.535	0.265	0.477
6	0.480	-0.250	-0.410	0.450	0.130	0.780
7	0.678	0.508	-0.745	0.927	0.144	0.849
8	0.874	1.230	-0.521	0.789	0.290	0.640
9	0.734	0.891	-0.570	1.010	0.124	0.789
10	1.006	1.237	-0.574	0.267	0.257	0.684
11	1.030	1.151	-0.223	0.330	0.340	0.641
12	0.442	0.571	-0.487	0.871	0.189	0.760
13	0.248	0.304	-0.247	0.785	0.089	0.680
14	0.740	1.002	-0.220	0.998	0.180	0.680
15	0.421	0.827	-0.643	0.372	0.363	0.561
16	0.342	1.270	-0.410	0.665	0.130	0.560
18	0.539	0.799	-0.745	0.320	0.41	0.526
19	0.546	0.563	-0.778	1.278	0.030	0.886
20	0.932	0.998	-0.436	0.987	0.062	0.910
21	0.431	1.220	-0.401	1.141	0.312	0.520
22	0.884	1.003	-0.374	0.999	0.145	0.803
23	1.290	1.455	-0.761	0.777	0.140	0.680
24	0.651	1.627	-0.730	1.120	0.020	0.813
26	0.679	1.010	-0.792	0.976	0.660	0.670
27	0.678	-1.125	-0.941	0.845	0.133	0.544
29	1.349	1.410	-0.363	0.454	0.080	0.890
30	1.022	1.270	-0.651	0.510	0.170	0.750

End.. ˜10-6
Time expired
Failure rows 17 25 28

JLMSG#52931/

Independent variable t: 284

0.251	-0.664	-1.239	0.660
0.575	-0.057	-1.463	0.451
-0.657	-0.445	0.180	0.277
-1.670	0.897	0.658	0.633
0.157	0.258	-2.732	-0.642
0.531	0.686	-0.707	0.705
0.073	0.701	0.011	1.218
0.288	-0.194	-0.194	0.248
-1.306	0.788	0.137	0.888
-1.236	0.252	-1.677	-1.033
0.130	-0.355	-0.711	-1.029
-0.704	0.171	0.075	0.249
-1.022	-0.145	0.176	-1.462
0.088	0.582	-0.709	0.083
-0.142	0.923	0.577	-2.108
-1.030	0.919	0.135	-1.035
0.653	-2.642	0.178	0.256
0.359	0.880	0.493	-0.442
0.173	0.160	-0.149	-0.555
-0.054	-2.254	-1.673	1.287
0.517	-0.551	0.002	-1.459
0.005	0.291	0.010	-0.641
-1.457	-0.662	0.095	0.522
-0.196	-0.202	-0.358	0.133
-0.636	0.538	0.342	0.245
0.699	0.080	0.704	0.295
0.243	-1.025	-1.313	-0.638
0.487	-0.059	-0.449	0.323
0.433	-0.660	-0.190	0.492
0.918	-0.147	0.439	0.656
-0.439	0.489	0.534	-0.061
-0.550	-0.059	0.164	-1.459
-2.104	0.921	0.364	-1.037
-2.632	-0.552	0.164	0.155
0.154	0.293	-2.642	-2.732
0.322	0.422	0.332	0.438
-0.290	-0.100	0.435	-0.191
0.267	0.367	-1.675	-2.104
0.486	0.586	0.496	0.268
-0.454	0.580	-1.309	1.220
0.870	0.487	0.523	0.214
0.277	0.871	0.216	0.555
0.040	-0.965	0.740	0.284
1.287	0.536	-2.124	1.387
-0.444	0.280	0.878	-0.944
0.743	0.793	0.519	0.744
0.449	1.296	-1.311	-0.445
0.547	-0.462	0.040	-0.211

-0.964	-0.164	0.554	0.047
1.557	0.452	1.757	0.556
0.657	0.657	-0.065	-0.298
-0.062	0.078	0.957	-0.102
-0.227	-0.974	0.449	-0.267
0.682	0.090	0.182	1.577
0.395	0.132	0.995	0.445
0.975	0.301	0.095	-0.627
0.634	1.559	0.134	0.604
-0.291	0.749	0.657	1.072
0.788	-1.241	-1.027	0.365
-0.291	0.390	0.630	0.878
-0.192	-1.273	0.687	-0.292
0.165	-1.091	0.364	0.272
0.870	-0.191	0.162	0.491
1.052	-0.195	0.170	-0.196
1.061	0.485	-0.292	1.485
1.546	1.446	1.536	
-1.156	-1.155	-0.156	
0.602	0.656	0.613	
-1.691	1.991	-1.690	
0.997	-0.281	1.055	
0.879	1.461	0.987	
0.157	0.258	-2.732	
0.682	1.072	-0.291	

BIBLIOGRAFÍA

ABRAHAMOWITZ, M. y RAMSAY, J.O. (1992). Multicategorical spline model for item response theory. Psychometrika, 57, 5-27.

ACKERMAN, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. Journal of Experimental Psychology: General, 117, 288-318.

ALBANESE, M.A. (1992). Type K items. Educational Measurement: Issues and Practices, 12, 117-128.

ANDERSEN, E. B. (1973). Conditional inference and multiple choice questionnaires. British Journal of Mathematical and Statistical Psychology, 26, 31-44.

ANDERSEN, E. B. (1983). A general latent structure model for contingency table data. En H. Wainer y S.Messick (Eds.), Principles of modern psychological measurement (pp. 117-138). Hillsdale, NJ: Lawrence Erlbaum.

ANDERSEN, E.B. (1985). Estimating latent correlations between repeated testings. Psychometrika, 50, 3-16.

ANDERSON, J.R. (1990). The adaptive character of thought. Hillsdale, N.J., Lawrence Erlbaum.

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

ANDRICH, D. (1985). An elaboration of Guttman scaling with Rasch models for measurement. In N. Brandon-Tuma (Ed.), Scociological methodology (pp. 33-80). San Francisco: Jossey-Bass.

ANDRICH, D. (1995). Distinctive and incompatible properties of two common classes of IRT models for graded responses. Applied Psychological Measurement, 19, 101-119.

ANGER, G. (1990). Inverse problems in differential equations. Akademieverlag, Berlin.

ARZARELLO, F., BAZZINI, L. y CHIAPPINI, G. (1993). Cognitive processes in algebraic thinking: towards a theoretical framework, Proc. PME XVII, Tsukuba, Japan, I, 138-145.

ARZARELLO, F., BAZZINI, L. y CHIAPPINI, G. (1994). The process of naming in algebraic problem solving, Proc. PME XVIII, Lisbon, II, 40-47.

ASIALA A., BROWN A., DEVRIES D., DUBINSKY E., MATHEWS D. Y THOMAS K. (1996). A Framework for Research and Curriculum Development in Undergraduate Mathematics Education, en Kaput J., Shoenfeld A. y Dubinsky E. (Eds.) Research in Collegiate Mathematics Education II, pp. 1-32.

ASIALA, M., COTTRILL, J., DUNINSKY, E. y SCHWINGENDORF, K. (1997). The development of students' graphical understanding of the derivative. Journal of Mathematical Behavior, 16, 399-431.

AVILA A., GARCIA F. y ROJANO T. (1990). Algebraic Syntax Errors: A Study with Secondary School Children. Proceedings of the Fouteenth PME Conference, México pp. 11-18.

BABAD, E.Y. Y BUDOFF, M. (1974). Sensitivity and validity of training- Potential measurement in three levels of ability. Journal of Educational Psychology, 66, 439-447.

BANKS, H. y KUNISH, K. (1989). Parameter estimation techniques for distributed systems. Birkhauser, Boston.

BEDNARZ N. y DUFOUR-JANVIER B. (1991). A Study of External Representations of Change developed by Young Children. Proceedings of the Thirteenth PME-NA Conference, Blacksburg, Virginia, pp. 140-146.

BEDNARZ, N. y JANVIER, B. (1994). The emergence and development of algebra in a problem solving context: A problem analysis, Proc. PME XVIII, Lisbon, II, 64-71.

BEDNARZ, N., JANVIER, B., MARY, C. y LEPAGE, A. (1992). L'algèbre comme outil de résolution de problèmes: une réflexion sur les changements nécessaires dans le passage d'un mode de traitement arithmétique à un mode de traitement algébrique, Recueil des textes du Colloque du programme de

Recherche sur l'emergence de l'algèbre, 17-31. Montréal, CIRADE, Université du Quebec à Montreal.

BEGGS, F.L. y LEWIS, E.L. (1979). Evaluación del Proceso Educativo: Uso y Diseño de Tests Escolares. TEA, Madrid.

BELL, A. y MALONE, J. (1993). Learning the language of algebra, Proc. PME XVII, Tsukuba, Japan, I, 130-137.

BELL, E.T. (1945). The development of Mathematics, Ed. N. Y. - London

BENNETT, R.E., ROCK, D.A., BRAUN, H.I., FRYE, D., SPOHRER, J.C., y SOLOWAY, E. (1990). The relationship of constrained free-response to multiple-choice and open-ended items. Applied Psychological Measurement, 14, 151-162.

BENNETT, R.E., y WARD, W. C. (Eds.)(1993). Construction versus choice in cognitive measurement. Hillsdale N.J. Lawrence Erlbaum.

BEREITER, C. (1963). Some persisting dilemmas in the measurement of change. En C.W. Harris (Ed.) Problems in measuring change (pp.3-20). Madison. University of Wisconsin press.

BERGEN, M. P. F. (1992), Sequential sampling designs for the two parameter item response theory models. Psychometrika, 57, 161-177.

BETH, E.W. y PIAGET, J. (1966). Mathematical Epistemology and Psychology. Reidel Dordrecht.

BETH, E.W. y PIAGET, J. (1966). Mathematical Epistemology and Psychology. Reidel Dordrecht.

BIRNBAUM, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M.R. Novick, Statistical theories of mental test scores (pp. 17-20). Reading, MA: Addison Wesley.

BIRNBAUM, A. y TATSUOKA, K.K. (1987). Open ended versus multiple choice response formats- It does make a difference for diagnostic purposes. Applied Psychological Measurement, 11, 329-341.

BOCK, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. Psychometrika, 37, 29-51.

BOCK, R.D. (1975). Multivariante statistical methods in behavioral research, New York, NY: McGraw Hill.

BOCK, R. D. y LIEBERMAN, M. (1970). Fitting a response model for n dichotomously scored items. Psychometrika, 35, 179-197.

BOLEA, P., BOSCH, M. y GASCÓN, J., (1998a). The role of algebraization in the study of a mathematical organization, CERME-1, Osnabrueck, Germany.

BOLEA, P., BOSCH, M. y GASCÓN, J. (1998b). Le caractère problématique du processus d'algébrisation. Proporcionnalité et grandeurs dans l'enseignement obligatoire, Actes de la IXème école d'été de didactique des mathématiques, ARDM, 153-159.

BOLEA, P., BOSCH, M. y GASCÓN, J. (1998c). El proceso de algebrización de las matemáticas escolares, (pendiente de publicación).

BOOTH L. (1984). Algebra: Children's Strategies and Errors, NFER-NELSON, Windsor.

BORMOUTH, J.R. (1970) On a theory of achievement test items. Chicago, Universiy of Chicago Press.

BRENNAN, R.L. (1982). Elements of generalizability theory. Iowa City, Iowa: American College Testing Program.

BRENNAN, R.L. (1993, April). The context of context effects. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

BRIDGEMAN, B., & ROCK, D.A. (1993). Relationship among multiple-choice and open-ended analytical questions. Journal of Educational Measurement, 30, 313-329.

BRIDGEMEN, B. (1992). A comparison of quantitative questions in open-ended and multiple-choice formats. Journal of Educational Measurement, 29, 253-271.

BRUST, M. (1997). Solución de ecuaciones de segundo grado: un análisis comparativo. Tesis para obtener el grado de licenciatura. ITAM. México.

BUNT, L.N.H., JONES, P.S. AND BEDIENT, J.D. (1976). The Historical Roots of Elementary Mathematics, Prentice-Hall, Englewood Cliffs, New Jersey.

BURDEN, R. L. Y FAIRES, J. D. (1985). Numerical Analysis, 3rd Edition. Boston. Prindle, Weber y Schmidt.

BUTTER, R., DE BROECK, P. y VERHELST N, (1998). An item response model with internal restrictions on item difficulty. Psychometrika 63, 47-65.

CAJORI, F. (1980). A History of Mathematics, Chelsea Publishing Company New York.

CAMILLI, G. (1994) Origin of the scaling constant d=1.7 in Item Response Theory. Journal of Educational and Behavioral Statistics 47, 293-295.

CAMPIONE, J.C., BROWN, A.L., FERRARA, R.A., JONES, R. S., Y STEINBERG, E. (1985). Breakdown in flexible use of information: Inteligence related differences in transfer following equivalent learning performance. Intelligence, 9, 323-344.

CARLSON, J.S. y WEIDL, K.H. (1979). Towards a differential testing approach: Testing the limits employing the Raven's matrices. Intelligence, 3, 323-344.

CARROLL, J. B. (1976). Psychometric tests as cognitve tasks: A new structure of intellect. In L. B. Resnick (Ed.), The nature of intelligence (pp. 27-56). Hillsdale, NJ: Lawrence Erlbaum Associates.

CATTEL, R.B. y KLINE, P. (1982). El Análisis Científico de la Personalidad y la Motivación. Pirámide, Madrid.

CEDILLO, T. (1996). Exploring Algebra as a Language in Use: A Study with 11-12 year-olds using graphic calculators. Tesis Doctoral, Instituto de Educación, Universidad de Londres.

CHANG, H. H. (1996). The asymptotic posterior normality of the latent trait for politomous IRT Models. Psychometrika 61, 445-464.

CHEVALLARD, Y. (1989). Arithmétique, Algèbre, Modélisation. Étapes d'une recherche. Publications nº 16 de l'IREM Aix-Marseille.

CHOPPIN, B. (1983). A two-parameter latent trait model. (CSE Report No. 197). Los Angeles, CA: University of California, Center for the Study of Evaluation, Graduate School of Education.

CLARK, J.M., CORDERO, F., COTRILL, J., CZARNOCHA, B., DEVRIES, D.J., ST JOHN, D., TOLIAS, G., VIDAKOVIC, D. (1997). Constructing a schema: The case of the chain rule. Journal of Mathematical Behavior, 16, 345-364.

CLEMENT, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception, Journal for Research in Mathematics Education, 13, 16-30.

COLE, N.S. (1990). Conceptions of educational achievement. Educational Researcher, 19, 2-7.

CONFREY J. (1993). The role of Technology in Reconceptualizing Funcions. En Proceeding of the Fifteenth PME-NA Conference, San José State University, USA, Vol. 1, pp. 47-74.

CORTÉS, A. (1992). Invariants operatoires dans le traitement des equations, Actes du colloque du programme cogniscience du CNRS ECCOS'92; Orsay.

CROCKER, L. Y ALGINA, J. (1986). Introduction to calssical and modern test theory. Harcourt Brace Jovanovich College Publishers. Florida USA.

CRONBACH, L.J. (1951): Coefficient alpha and the internal structure of tests. Psycometrika, 16, 297-334.

CROMBACH, J.L.(1977). Fundamentos de Exploración Psicológica. México, Trillas.

CLARK, J. M., COTTRILL, J., CZARNOCHA, B., DEVRIES, D.J., ST. JOHN, D., TOLIAS, G., Y VIDAKOVIC, D. (1997). Constructing a schema: The case of the chain rule. Journal of Mathematical Behavior, 16, 345-364.

DAWSON-SAUNDERS, B., RESHETAR, R., SHEA, J.A., FIERMAN, C.D., KANNGILASKI, R., y PONIATOWSKI, P.A. (1992, April). Alterations to item text and effects on item difficulty and discrimination. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

DAWSON-SAUNDERS, B., RESHETAR, R., SHEA, J.A., FIERMAN, C.D., KANNGILASKI, R., y PONIATOWSKI, P.A. (1993, April). Changes in difficulty and discrimination related to altering item text. Paper presented at the annual meeting of the National Council on Measurement in Education, Atlanta.

DEMBO, R.S. y STEINWAY, T. (1983) Truncated Newton methods for large scale unconstrained optimization. Math Progf, 26, 190-212.

DEUFLHARD, P. HAIRER, E. (Eds.). (1983). Numerical treatement of inverse problems in differential and integral equations, Birkhauser, Boston.

DIBELLO, L.V., ROUSSOS, L.A., & STOUT, W.F. (1993, April). Unified cognitive/psychometric diagnosis foundations and application. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

DOWNING, S.M. (1992). True-false and alaternate-choice item formats: A review of research. Educational Measurement: Issues and Practices, 11, 27-30.

DREYFUS T. y EISENBERG T. (1981). Function Concepts: Intuitive Baseline. En Proceedings of the Fifth PME Conference, Grenoble, Francia, pp. 183-188.

DROUHARD, J. P. (1992). Les Écritures Symboliques de l'Algèbre Elémentaire, Tesis Doctoral, Université Denis Diderot, París 7.

DUBINSKY, E. (1985). The Constructive Aspects of Reflective Abstraction. Internal Publication. Purdue University.

DUBINSKY, E. (1989). The Case Against Visualization in School and University Mathematics. PME 13. Paris.

DUBINSKY, E. (1991a) Reflective Abstraction in Advanced mathematical Thinking en Tall D. (Ed.) Advanced Mathematical Thinking. Kluwer Academic Publishers.

DUBINSKY, E. (1991b). Constructive aspects of reflective abstraction in advanced mathematics. En L.P. Steffe (Ed.) Epistemological foundations of mathematical experience, pp. 160 – 202. New York. Springer Verlag.

DUBINSKY, E. (1992) A learning thory approach to calculus, symbolic calculation. En Z.A. Karian (Ed.) Undergraduate Mathematics Education, MAA Notes No. 24, the Mathematical Association of America, pp. 43- 55.

DUBINSKY, E. (1994). A theory and Practice of learning college mathematics. En A. Schoenfeld (Ed.) Mathematical Thinking and Problem Solving, pp. 221 – 243. Hillsdale, N.J. Erlbaum.

DUBINSKY, E. (1996). El aprendizaje de los conceptos abstractos de la matemática avanzada. En Memorias de la X Reunión centroamericana y del Caribe sobre Formación de Profesores e investigación en Matemática Eduactiva, Puerto Rico, pp. 1-9.

DUBINSKY, E. y HAREL, G. (1992). The nature of the process conception of function. En G. Harel y E. Dubinsky (Eds.) The concept of function. Washington.

DUBINSKY, E, y LEWIN, P . (1986). Reflective abstraction and mathematics education. The journal of Methematical Behaviour, 5, pp. 125-250.

DUBINSKY, E. y SHWINGENDORF K. (1993). Constructing calculus concepts: Cooperation in a computer laboratory. Internal communication. Purdue University.

EBEL, R.L., & FRISBIE, D.A. (1991). Essential of educational measurement (th de.). Englewood Cliffs, NJ: Prentice-Hall.

ELIAS CALLES P. (1997). Evaluación de algunos conceptos del álgebra en estudiantes universitarios a través de tres instrumentos. Tesis para obtener el grado de licenciatura. ITAM. México.

EMBRESTON, S. E. (1980). Multicomponent latent trait models for ability tests. Psychometrika, vol. 45, pp 479- 494.

EMBRESTON, S. E. (1984). A general latent trait model for response processes. Psychometrika, vol 49, pp. 175-186.

EMBRESTON, S.E. (1983a). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93, 179-197.

EMBRESTON, S.E. (1983b) An incremental fit index for the linear logistic latent trait model. Paper presented at the annual meeting of the Psychometric Society, Los Angeles, CA.

EMBRESTON, S.E. (1985a). Psychometric models for learning and cognitive processes. En Test design: Developments in psychology and psychometrics (pp.195-218) Orlando, FL. Academic Press.

EMBRESTON, S.E. (1985b). Multicomponent latent trait models for test design. En S.E. Embreston (Ed.) Test design: Developments in psychology and psychometrics (pp.195-218) Orlando, FL. Academic Press.

EMBRESTON, S. E. (1985c). Introduction to the problem of test design. In S. E. Embreston (Ed.) Test design: Developments in psychology and psychometrics. New york, Academic Press.

EMBRESTON, S.E. (1993). Psychometric models for learning and cognitive processes. En Frederiksen, N., Mislevy, R.J., & Bejar, Y. (Eds.). Test theory for a new generation of tests. Hillsdale, NJ: Lawrence Erlbaum Associates.

EMBRESTON, S.E. y WAXMAN, M. (1989). Models for processing and individual differences in spatial folding. Unpublished manuscript.

EMBRESTON, S.E., SHNEIDER, L. Y ROTH, D.L.(1986). Multiple processing strategies and the construct validity of verbal reasoning tests. Journal of Educational Measurement, 23, 13-22.

ENGL, H. W., (1994). Inverse problems. Notas para curso. Comunicación personal.

ENGL, H. W., HANKE, M, NEUBAUER, A. (1990). Tikhonov regulatization of nonlinear differential-algebraic equations, en Sabatier, P.C. (Ed.) Inverse metods in action, Springer, Berlin, pp 92-105.

ENGL, H. W., KUNISCH, K.,NEUBAUER, A., (1989). Convergence rates for Tikhonov regularization of ill-posed problems. Inverse problems 5, 1-28.

ENGL, H.W. (1993). Inverse problems. Johanes Kepler Universitat. Linz, Austria.

ENGLISH L. D. Y SHARRY P.V. (1996). Analogical Reasoning and the Development of Algebraic Abstraction, en Educational Studies in Mathematics, 30, pp. 135-157.

ESCARENO S.F., LIMA S.A. y NORIEGA C.B.M. (1979). Matemáticas por Objetivos 1, Trillas, Mexico.

FELDT, L. S. (1965): "The aproximate sampling distribution of Kuder-Richardson reliability coefficient twenty", Psycometrika, 30, 357-370.

FEUERSTEIN, R. (1979). The dynamic assessement of retarded performers: The learning potential assessement device, theory, instruments and techniques. Baltimore. University Park Press.

FILLOY E. and ROJANO T. (1984). From an Arithmetical to an Algebraic Thought. En Proceedings of the Sixth PME-NA Conference, Wisconsin, USA, PP. 51-56.

FILLOY E. and ROJANO T. (1989). Solving Equations: The Transition from Arithmetic to Algebra. For the Learning of Mathematics 9, 2, pp. 19-25.

FISHER, G. H. (1973). The linear logistic test model as an instrument in educational research. Acta Psychologica, 37, 359 –374.

FISHER, G. H. (1983). Logistic latent trait models with linear constraints. Psychometrika 48, 3-26.

FISHER, G.H. (1995). Some neglected problems in IRT. Psychometrika, 60, 459-485.

FISHER, G.H. y LAMMING, D. (Eds.)(1994). Contributions to mathematical psychology, sychometrics and methodology. Hillsdale, N.J. Lawrence Erlbaum.

FISHER, G. H. y MOLENAAR, I. W. (1996). (Eds.) Rasch models: Ffundations, recent dvelopments and applications. Hillsdale, N.J. Erlbaum.

FLANAGAN, J.C. , (1937). A note on calculating the standard error of measurement and reliability coefficients with the test scoring machine. Journal of Applied Psychology, 23, 529.

FREDERIKSEN, N. (1984) The real test bias: influences on testing on teaching and learning. American Psychologist 39, 193-202.

FREDERIKSEN, MILEVY, R.J. y BEJAR, (1993). Test theory for a new generation of tests. Hillsdale, N.J. Lawrence Erlbaum.

FRISBIE, D.A. (1992). The status of multiple true-false testing. Educational Measurement: Issues and Practices, 5, 21-26..

FROBERG C. E. (1985). Numerical mathematics. Theory and computer applications. The Benjamin Cummings Publishing Co., Menlo Park, California, USA.

GAGNÉ, R.M. (1968). Learning hyerarchies. Educational Psychologist, 6, 1-9.

GARCÍA NIETO, N. (1994). Apuntes del Curso de Doctorado. México. 1994

GARCÍA NIETO, N. (1987). Proyecto Docente e Investigador. Universidad Complutense. Madrid.

GASCÓN J. (1993). Desarrollo del conocimiento matemático y análisis didáctico: del patrón de análisis-síntesis a la génesis del lenguaje algebraico, Recherches en Didactique des Mathématiques, 13(3), 295-332.

GASCÓN J. (1994). Un nouveau modèle de l'algèbre élémentaire comme alternative à l'"arithmétique généralisée", Petit x, 37, 43-63.

GASCÓN J. (1998). La naturaleza prealgebraica de la matemática escolar, Educación matemática (por aparecer).

GINSSBURG, H. (1977). Children's arithmetic. The learning process. New York: Van Nostrand.

GINSSBURG H. y OPPER S. (1979). Piaget's Theory of Inetellectual Development: an Introduction. Prentice Hall, Engelwood Cliff..

GLAS, C.A. W., y VERHERLST, N.D. (1995). Testing the Rasch model. In G.H. Fischer & I. W. Molenaar, Rasch models: Their foundations, recent developments and applications (pp. 68-85). New York: Springer-Verlag.

GLASER, R. (1985). The integration of instruction and testing. Paper presented at the ETS Invitational Conference on the Redesign of testing for the 21th century, New york.

GLASER, R. (1993, April). Criterion-referenced tests: Origins and unfinished business. In Criterion-referenced measurement: A 30-year retrospective.

Symposium conducted at the annual meeting of the american Educational Research Association, Atlanta.

GLASER, R., LESGOLD, A. Y LAJOIE, S. (1987). Towards a cognitive theory for the measurement of achievement. En R. RONNING, J. GLOVER, J. C. CONOLEY y J. WITT (Eds.) The influence of cognitive psychology on testing and measurement, Vol. 3, 41-85. Hillsdale, NJ, Erlbaum.

GLONEK, G. F. V. Y MCCULLAGH, P. (1995). Multivariate Logistic Models. J. R. Statist. Soc. B, 57, pp 533-546.

GOMEZ, S y MORALES, J.L. (1997). A truncated method for solving parameter identification problems. Documento de trabajo. Depto. De Matemáticas, ITAM, México.

GROFF, G. (1987). Number Patterns 1: Simple Mapping. Deártment of Education, University of Stirling, Gran Bretaña.

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

HABACUC, L. (1980). Matemáticas, Primer Curso. Editorial Hierro S.A., México.

HABERMAN, S.J. (1973). The analysis of residuals in cross-classified tables. Biometrics, 29, 205-220.

HABERMAN, S.J. (1979). Analysis of qualitative data: New developments, Vol. 2. New York: Academic Press.

HAERTEL, E.H. (1984). An application of latent class models to assessment data. Applied Psychological Measurement 8, 336-346.

HAERTELL, E.H., y WILEY, D.E. (1993). Representations of ability structures: Implications for testing. In N. Frederiksen, R.J. Mislevy, & Y. Bejaar (Eds.), Test theory for a new generations of test (pp. 359-384). Hillsdale, NJ: Lawrence Erlbaum Associates.

HALADYNA, T.M. (1992a). Context dependent item sets. Educational Measurement: Issues and Practices, 11, 21-25.

HALADYNA, T.M. (1992b). The effectiveness of several multiple-choice formats. Applied Measurement in Education, 5, 73-88.

HALADYNA, T.M. (1994). Developing and validating multiple choice test items. Hillsdale, N.J. Lawrence Erlbaum.

HALADYNA T.M. y DOWNING, S.M. (1989). A taxonomy of multiple choice item writing rules. Applied Meassurement in Education 1, 51-78.

HALADYNA, T.M., y DOWNING, S.M. (1997). How many options is enough for a multiple-choice test item. Educational & Psychological Measurement.

HALADYNA, T.M., y SYMPSON, J.B. (1988, April). Empirically based polychotomous scoring of multiple-choice test items: A review. In New Development in Polychotomos Scoring. Symposium conducted at the annual meeting of the American Educational Research Association, New Orleans.

HALADYNA, M., THOMAS, (1994). Developing and validating multiple choice test items, Lawrence Erlbaum associates Pu., Hillsdale, New Jersey.

HAMBLETON, R.K. Y SWAMINATHAN, H.(1985) Item Response Theory. Principles and Applications. Kluwer - Nijhoff Publishing. Boston USA

HAMBLETON, R.K., y JONES, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. Educational Measurement: Issues and Practices, 12, 38-46.

HANCOCK, G.R. (1992, April). Impact of item complexity on the comparability of multiple-choice and constructed-response test formats. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

HANCOCK, G.R., THIEDE, K.W., y SAX, G. (1992, April). Reliability of comparably written two-option multiple-choice and true-false test items. Paper presented at the annual meeting of the Naational Council on Measurement in Education, Chicago.

HATWELL, Y. (1992). A propósito de las nociones de asimilación y acomodación en los procesos cognoscitivos. En Ajuriaguerra, A. Psicología y Epistemología Genéticas. Nociones. México.

HEID M.K. y KUNKLE D. (1988). Computer-generated Tables: Tools for Concept Development in Elementary Algebra. En Coxford A.F. y Shulte A.P. (Eds.). The Ideas of Algebra K-12, pp. 170-177.

HEMKER, B. T., SIJTSMA, K., MOLENAAR, I. W. Y TEN BERGER, J. M. F. (1996). Politomous IRT models and monotone likelihood ratio of the total score. Psychometrika 61, 679- 694.

HERSCOVICS N. and LINCHEVSKI L. (1991). Pre-algebraic Thinking: Range of Equations and Informal Solution Processes Used by Seventh Graders Prior to any Instruction. En Proceedings of the Fifteenth PME Conference, Assisi, Italy, pp. 173-180.

HESS, R. (1994, April). Using the Rasch model to calibrate a district-wide, curriculum-based mathematics assessment. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

HILLEL J. (1992). The Notion of variable in the Context of the Turtle Graphics. En C. Hoyles y R.Noss (Eds.), Learning Mathematics and LOGO. MIT Press, pp. 11-36.

Hillsdale, New Jersey 1994

HOIJTINK, H. y MOLENAAR I.W. (1997). A multidimensional item response model: constrained class analysis using the Gibbs sampler and position predictive checks. Psychometrika 62, 171-189.

HOLLAND, P.W. y WAINER, H. (1993). Differential item functioning. Hillsdale, N.J. Lawrence Erlbaum.

HOSKENS, M. y De BOECK, P. (1995). Componential IRT models for polytomous items. Journal of Educational Measurement, 32, 364-384.

HOYLES C. y NOSS R. (1987). Children Working in a Structured Logo Environment: From Doing to Understanding. Recherches en Didactique des Mathématiques, 8, 12, pp. 131-174.

HUYNH , H. (1996). Decomposition of a Rash partial credit item into independent binary and indecomposable trinary items. Psychometrika 61, 31-39.

HUYNH , H. Y FERRARA, S. F. (1994). A comparison of equal percentile and partial credit equatings for performance-based assessments comprised of free response items. Journal of Educational Measurement, 31, 125-141.

JANNARONE, R. J. (1986). Conjunctive item response theory kernels. Psychometrika, 51, 357-373.

JANSSEN, R, y DE BOECK, P. (1997). Psychometric modelling of componenetially designed synonym tasks. Applied Psychological Measurement, 21, 37-50.

KAPUT, J. (1983). Errors in translations to algebraic equations: Roots and implications, Focus on Learning Problems in Mathematics, 5, 63-78.

KAPUT, J. (1987). Algebra papers: A representational framework, en N. Bergeron, N. Herscovics y C. Kieran (eds.), Proceedings of the eleventh International Conference for the Psychology of Mathematics Education, pp.345-354.

KAPUT, J. (1996). ¿Una línea de investigación que sustente la reforma del álgebra? I y II, UNO. Revista de Didáctica de las Matemáticas, 9, 85-97 y 10, 89-103.

KEEVES, J.P. (1988). Questionnaires. Educational Research, Methodology and Measurement: an International Handbook, pp 478- 482. Pergamon Press, Oxford.

KELDERMAN, H. (1989). Item bias detection using loglinear IRT. Psychometrika 54, 681-697.

KELDERMAN, H. Y MACREADY, G.B. (1990). The use of loglinear models for assessing differntial item functioning across manifest and latent examinee groups. Journal of Educational Measurement, 27, 307-327.

KELLER, J. (1976). Inverse problems. Amer. Math. Monthly 83, 107-118.

KIERAN C. (1984). Cognitive mechanisms underlying the equation - solving errors of algebra novices. Proceedings of the Eighth PME Conference, pp. 70-77.

KIERAN C. (1992). Multiple solutions to problems: The role of non-linear functions and graphical representations as catalysts in changing students' beliefs. Seminar presented at CINVESTAV, México.

KIERAN, C. y FILLOY, E. (1989). El aprendizaje del álgebra escolar desde una perspectiva psicológica, Enseñanza de las Ciencias, 7 (3), 229-240.

KLEIN, J. (1968). Greek Mathematical Thought and the Origin of Algebra, M.I.T. Press.

KRISTOF, W. (1963). The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. Psychometrika 28, 221-238.

KÜCHEMANN, D. (1980). The Understanding of Generalised Arithmetic (Algebra) by Secondary School Children, PhD Thesis, University of London.

KUDER, G.F., y RICHARDSON, M.W. (1937). The theory of the estimation of test reliability. Psychometrika 2, 151-160.

LABINOWITZ, E. (1987). Introducción a Piaget. Pensamiento, aprendizaje, enseñamza. Addison Wesley Iberoamericana. Delaware.

LADUCA, A., STAPLES, W.I., TEMPLETON, B., y HOLZMAN, G.B. (1986). Item modelling procedure for constructing content-equivalent multiple-choice questions. Medical Education, 20, 53-56.

LADUCA, A., y DOWNING, S.M. (1985). Test development: Systemataic item writing and test construction. En J.C. Impara y J.C. Fortune (Eds.), Licensure examination, Lincoln, NE: Buos Institute of Mental Measurements.

LAWLEY, D. N. (1943). On problems connected with item selection and test construction. Proceedings of the Royal Society of Edinburgh, 62-A (Pt. I), 74-82.

LAZARO, A. (1991). La formalización de Indicadores de Evaluación. Bordón 43-44, 477-495.

LAZARO, A. (1994). Dossier para el curso Elaboración de Instrumentos de Diagnóstico y Evaluación.

LERON U. y ZAZKIS R. (1986). Functions and Variables. Proceeding of the Second International Conference for Logo and Mathematics Education, London, U.K., pp. 186-192.

LEVINE, M., y WILLIAMS, B. (1991). Nonparametric item response function estimation strategies. Paper presented at the ONR Conference on Model-Based Measurement, Princeton, New Jersey.

LOHMAN, D.F. (1989). Human intelligence: An introduction to advances in theory and research. Review of Educational Research, 59, 333-373.

LOHMAN, D.F., e IPPEL, M.J. (1993). Cognitive diagnosis: From statistically based assessment toward theory-based assessment. In N. Frederiksen, J.R. Mislevy, & Y. Bejar (Eds.), Test theory for a new generation of tests (pp. 41-71). Hillsdale, NJ: Lawrence Erllbaum Associates.

LÓPEZ, A. (1996). Construcción de la noción de variable algebraica en alumnos de nivel medio superior. Ms Sc. Tesis. Universidad Autónoma de Querétaro, México.

LORD, F.M. (1952). A theory of test scores. Psychometric Monograph, No. 7.

LORD, F.M. (1970). Item characteristic curves estimated without knowledge of their mathematical form: A confrontation of Birnbaum's logistical model. Psychometrika 35, 43-50.

LORD, F.M. (1975). Formula scoring and number right scoring. Journal of Educational Measurement, 12, 7-12.

LORD, F.M. (1977) Optimal number of choices per item-A comparison of four approaches. Journal of Educarional Measurement, 14, 33-38.

LORD, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale N.J. Lawrence Erlbaum.

LORD, F.M., y NOVICK, M.R. (1968). Statistical theories of mental test scores. Reading, MA. Addison Wesley.

LOZANO D. (1998). El Concepto de Variable: Evolución a lo largo de la Instrucción Matemática. Tesis de licenciatura. ITAM, México.

LUKHELE, R., THISSEN, D., y WAINER, H. (1992, April). On the relative value of multiple-choice, constructed-response, and examminee-selected items on two achievement tests. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

MARIS, E. (1995). Psychometric latent response models. Psychometrika 60, 523-547.

MARKOVITS Z., EYLON B.S. y BRUCKHEIMER M. (1986). Functions Today and Yesterday. For the Learning of Mathematics 6, 2, pp. 18-28.

MARKOVITS Z., EYLON B.S. y BRUCKHEIMER M. (1988). Difficulties Students Have with the Function Concept, en Coxford A.F. and Shulte A.P. (Eds.). The Ideas of Algebra K-12, pp. 43-60.

MARSHAL, S. P. (1989) Generating good items for diagnostic tests. En N. Frederiksen, R. Glaser, A. Lesgold y M. G. Shafto (Eds.). Diagnostic monitoring of skill and knowledge acquisition. Hillsdales, N. J. Erlbaum.

MARSHAL, S. P. (1993). Assesing schema knowledge. En N. Frederiksen (Ed.) Test theory for a new generation of tests. Hillsdale, Lawrence Erlbaum.

MARSHAL, S.P. (1990, April). What students learn (and remember) from word problem instruction. In Penetrating to the Mathematical Structure of Word Problems. Symposium conducted at the annual meeting of the American Educational Research Association, Boston.

MARSHAL, S. P, BARTHULI, K. E., BREWER y M. A. ROSE, F. E. (1989). Story problem solver: A schema- based system of instruction (Tech. Rep. 89-01). Center for Research in Mathematics and Science Education, San Diego, CA.

MARTINEZ, A. (1990). Problemas y polémicas en torno al diagnóstico. Bordón, 42,1. Sociedad Española de Pedagogía.

MASON J. (1989). Mathematical Abstraction as a Result of a Delicate Shift of Attention, en For the Learning of Mathematics, 9 (2), pp. 2-8.

MASON J., GRAHAM A., PIMM D. Y GOWAR N. (1985). Routes to/Roots of Algebra, The Open University Press, England.

MASTERS, G.N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

MASTERS, G.N. (1985). A comparison of latent-trait and latent-class analyses of Likert-type data. Psychometrika, 50, 69-82.

MASTERS, G.N. y MISLEVY R.J. (1993). New views of student learning: implications for educational measurement en N. Frederiksen (Ed.) Test theory for a new generation of tests. Hillsdale, Lawrence Erlbaum.

MATZ M. (1982). Towards a Process Model for High School Algebra Errors. En Sleeman, D. y Brown, J.S. (Eds.). Intelligent Tutoring Systems, London, Academic Press.

MAYER, R.E., LARKIN, J.H., & KADANE, J.B. (1984). A cognitive analysis of mathematical problem-solving ability. In R.J. Sternberg (Eds.), Advances in the psychology of human intelligence (pp. 231-273).

McCOLLAN, K.M. y EMBERSTEN, S. (1995). Applying the multicomponent latent trait model for measuring spatial processing components. 60th Annual Meeting of the Psychometric Society. June 15-18.

MELLENBERGH, G.J. (1995). Conceptual notes on models for discrete polytomous item responses. Applied Psychological Measurement, 19, 91-100.

MESSIK, S. (1984). The psychology of educational measurement. Journal of Educational Measurement, 21, 215-237.

MESSIK, S. (1993). Trait equivalence as construct validity of score interpretation across multiple methods of measurement. In R.E. Bennett & W.C. Ward (Eds.), Constrction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment (pp. 60-73). Hillsdale, N.J. Lawrence Erllbaum Associates.

MICHELL, I. (1990). An introduction to the logic of psychological measurement. Hillsdale, N.J. Lawrence Erlbaum.

MISLEVY R.J. (1986). Bayes model estimation in item response models. Psychometrika 51, 177-195.

MISLEVY, R. J. (1989). Foundations of a new test theory (Research Report No. RR-89-52-ONR). Princeton, NJ: Educational Testing Service.

MISLEVY, R.J. (1993a). Foundations for a new test theory. En Frederiksen, R.J. Mislevy y I. Bejar (Eds.) Test theory for a new generation of tests, pp. 19-39. Hillsdale, N.J., Lawrence Erlbaum.

MISLEVY, R.J. (1993b). Test theory reconsidered. En Frederiksen, Mislevy, R.J. y Bejar, (1993). Test theory for a new generation of tests. Hillsdale, N.J. Lawrence Erlbaum.

MISLEVY R.J. (1994). Evidence and inference in educational assessment, Psychometrika 59, 439-483.

MILEVY y BEJAR, (1993). Test theory for a new generation of tests. Hillsdale, N.J. Lawrence Erlbaum.

MISLEVY , R. J.y VERHELST , N. (1987). Modeling item responses when different subjects employ different solution strategies (research Report RR-87-47-ONR). Princeton, NJ. Educational Testing Service.

MISLEVY , R. J.y VERHELST , N. (1990). Modeling item responses when different subjects employ different solution strategies. Psychometrica, 55, 195-215.

MISLEVY, R. J y WILSON, M. (1996). Marginal maximun likelihood estimation for a psychometric model of discontinous development. Psychometrika 61, 41-71.

MOLENAAR, I.W. (1997). Nonparametric models for polytomous responses. In W.J. van der Linden & R.K. Hambleton (Eds.). Handbook of modern item response theory (pp. 369-380). New York. Springer.

MUÑIZ J., (1990). Teoría de Respuesta a los Items. Ediciones Pirámide, España.

MUÑIZ J. (1992). Teoría Clásica de los Tests, Ediciones Pirámide, España.

MURAKI, E. (1990). Fitting polytomous item response model to Likert- type data. Applied Psychological Measurement 14, 59-71.

MURAKI, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement, 16, 159-176.

MUTHÉN, B. (1988). Some uses of structural equation modelling in validity studies: Extending IRT to external variables. In H. Hainer y H. Braun (Eds.), Test validity (pp. 213-238). Hillsdale, NJ: Lawrence Erlbaum Associates.

MUTHÉN, B. (1989). Using item specific instructional information in achievement modelling. Psychometrika 54, 385-396.

NEUGEBAUER, O. (1969). The Exact Science in Antiquity, Dover Publications, New York.

NICHOLS, P. (1994). A framework for developing cognitively diagnostic assessments. Review of Educational Research, 64, 575-603.

NOSS R. (1985). Creating a mathematical environment through programming: A study of young children learning LOGO. Ph.D. Thesis, Chelsea College, University of London.

OLIVEROS, L. (1994). Dossier para el curso Elaboración de Instrumentos de Observación.

PARRA CABRERA L. Y WALLS MEDINA J. (1970). Matemáticas, Primer Curso. México. Kapelusz Mexicana.

PAULSON, J. (1985). Latent cass representation of systematic patternsin test responses (ONR Technical Report) Portland, OR: Portland State University.

PAULSON, J.A. (1986). Latent class representation of systematic patterns in test responses (Technical Report ONR-1). Portland, OR. Portland State University, Psychology Department.

PIAGET, J. (1960). Les structures maethématiques et les structures operatoires de l'intelligence. En L' enseignement des mathématiques. Neuchatel, Paris: Delacheaux et Niestlé.

PIAGET, J. (1964) Six Etudes de Psychologie. Gonthier. Ginebra.

PIAGET, J. (1967). Biologie et conaissance. Paris. Gallimard.

PIAGET, J. (1970). L' épistemologie génetique. Paris. Presses Universitaires Francaises.

PIAGET, J. (1973a). Introduction a l' épistemologie génetique (Vol I y II). Paris. Presses Universitaires Francaises.

PIAGET, J. (1973b). Remarques sur l'éducation mathématique. Math-école, 58, 1-7.

PIAGET, J. (1975). L'équilibration des structures cognitives. Paris. Presses Universitaires de France.

PIAGET, J. (1976). The Grasp of Consciousness. Harvard University Press.

PIAGET, J. (1977). Recherches sur l'abstraction réflechissante. Paris. Presses Universitaires de France.

PIAGET, J. (1978). Success and Understanding. Harvard University Press. Cambridge.

PIAGET, J. (1980). Adaptation and Intelligence. University of Chicago Press. Chicago.

PIAGET, J. y GARCÍA R. (1982). Psicogénesis e Historia de la Ciencia. Siglo XXI Editores.

QUINTERO, R.; REYES, A.; TRIGUEROS, M. y URSINI, S. (1994). Misconceptions of variable, Proceedings of the XVI Annual Meeting of the PME-NA, p. I-185.

QUINTERO, R.; URSINI, S., REYES, A. y TRIGUEROS, M. (1995). Students Approaches to Different Uses of Variable. Proceeding of the IXX PME International Conference.

RADFORD, L. y GRENIER, M. (1996). Entre les choses, les symboles et les idées... une séquence d'enseignement d'introduction à l'algèbre, Revue des sciences de l'education, XXII/2, 253-276.

RASH G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Denmarks Paedagogiske Institut.

RASCH, G. (1961). On general laws and the meaning of measurement in psychology. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, (pp. 321-333). Berkeley, CA: University of California Press.

RECKASE, M.D. (1985). Multidimensional IRT models. En D.J. Weiss (Ed.), Proceedings of the 1982 Item Response Theory and Computer Adaptive

Conference. Minneapolis, MN: University of Minnesota, Computer Adaptive Testing Laboratory.

REGGIANI M. (1994). Generalisation as a Basis for Algebraic Thinking: Observations with 11-12 Year Old Pupils. Proceedings of the XVIII PME International Conference, pp. IV-97-104.

REISER, M. (1996). Analysis of residuals for the multinomial item response model. Psychometrika. Vol. 61 pp. 509-528.

RICHARDSON, M. W. (1936). The relationship between difficulty and the differential validity of a test. Psychometrika 1, 33-49.

RILEY, M.S., GREENO, J.G., & HELLER, J.I. (1983). Development of children's problem solving ability in arithmetic. In H.P. Ginsburg (Eds.), The development of mathematical thinking (pp. 153-196). New York. Academic Press.

ROBERTS, D.M. (1993). An empirical study on the nature of trick questions. Journal of Educational Measurement, 30, 331-344.

ROJANO, T. (1994). La matemática escolar como lenguaje. Nuevas perspectivas de investigación y enseñanza, Enseñanza de las Ciencias, 12 (1), 45-56.

ROSENBAUM, P.R. (1987). Comparing item characteristic curves. Psychometrika 52, 217-233.

ROST, J. (1990). Rasch models in latent clases: An integration of two approaches to item analysis. Applied Psychological Measurement, 14, 271-282.

ROUSE BALL, W.W. (1960). A Short Account of the History of Mathematics, Dover Publications, New York.

ROYER, J.M., CISERO, C.A., y CARLO, M.S. (1993). Techniques and procedures for assessing cognitive skills. Review of Educational Research, 63, 201-243.

RUÍZ, L. y RODRIGUEZ, J.L. (1994). La modelización funcional algebraica en la Enseñanza Secundaria: incidencia de la parametrización, Proyecto de Investigación (no publicado), Universidad de Jaén.

RULON, P.J. (1939). A simplified procedure for determining the reliability of a test by split-halves. Harvard Educational Review, 9, 99-103.

RYAN, K.E. (1993, April). A comparison of the single right answer format and the multiple answer format with one correct answer: Is there a one right answer mentality? Paper presented at the annual meeting of the National Council on Measurement in Education, Atlanta.

SABATIER, P.C. (Ed.) (1990). Inverse methods in action. Springer, Berlin.

SAMEJIMA, F. (1969). Estimation of ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 34.

SAMEJIMA, F. (1972). A general model for free- response data. Psychometric Monograph, No. 18.

SAMEJIMA, F. (1979) A new family of models for the multiple-choice item. (Office of Naval Research Report 79-4) Knoxville, TN, University of tennessee.

SAMEJIMA, F. (1981). Final Report: Efficient methods of estimating the operating characteristics of item response categories and challenge to a new model for the multiple-choice item (Office of Naval Research Final Research Report of N00014-77-C0360). Knoxville, TN: Department of Pscychology, University of Tennessee.

SAMEJIMA, F. (1983). A latent trait model fro differential strategies in cognitive processes(Tecnical Report ONR/RR-83-1). Knoxville, TN: University of Tennessee.

SAMEJIMA, F. (1988). Advancement of latent trait theory 9ONR Final Report). Knoxville, TN: University of Tennessee.Samejima, F. (1979). A new family of models for the multiple-choice item. (Offfice of Naval Research Report 79-4). Knoxville, TN: University of Tennessee.

SAMEJIMA, F. (1990). Final Report: Validity study in multidimensional latent space and efficient computerized adaptive testing (Office of Naval Research Final Research Report of N00014-87-K-0320). Knoxville, TN: Department of Psychology, University of Tennessee.

SAMEJIMA, F. (1995a). A cognitive diagnosis method using latent trait models: competency space approach and its relationship with DiBello and Stout's unified cognitive/psychometric diagnosis model. In P.D. Nichols, S.E. Chipman

& R.L. Brennan (Eds.), Cognitively diagnostic assessment. Hillsdale, NJ: Lawrence Erlbaum.

SAMEJIMA, F. (1995b). Strenghts and weaknesses of mathematica models for polytomous responses in specific psychological realities. 60th Annual Meeting of the Psychometric Society. June 15-18.

SAMEJIMA, F. (1997). Departure from normal assumptions: A promise for future Psychometrics with sustantive mathematical modeling. Psychometrika-vol. 62, pp. 471-494.

SAMEJIMA, F. (1998). Efficient Nonparametric Approaches for Estimating the Operating Characteristics of Discrete Item Responses. Psychometrika, vol 63, pp. 111-130.

SAX, G., & REITER, P.B. (sin fecha). Reliability and validity of two-option multiple-choice and comparably written true-false items. University of Washington.

SCHAIE, K.W. Y STROTHERS, C. R. (1968). A cross sequential study of age changes in cognitive behaviour. Psychological Bulletin, 70, 671-680.

SCHEIBLECHNER, H. (1972). Das lernen und losen komplexer denkaufgaben (The studying and solving of some complex conceptual problems). Zeitchrif fur Experimentelle und Angewandte Psychologie, 19, 476-506.

SCHEIBELCHNER,H. (1985) Psychometric models for speed-test construction: The linear exponential model. En S. Emberston (Ed.) Test Design: Developements in Psychology and Psychometrics (pp. 219-244). New York. Academic Press.

SCHERZER, O., ENGL, H.W., KUNISH, K. (1995). Optimal a- posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems.

SERRALDE, E., ZUÑIGA, E., ZUÑIGA, I. y ZUÑIGA, J. (1982), 'Matemáticas 1', Ediciones Pedagógicas, México.

SFARD, A. (1989), Transition from Operational to Structural Conception: The Notion of Function Revisited, Proceedings of the Thirteenth PME Conference, Paris, pp.151-158.

SFARD A. (1994). The Gains and the Pitfolls of Reification - The Case of Algebra, en Educational Studies in Mathematics, 26, pp. 191-228.

SFARD, A. y LINCHEVSKI, L. (1994): The gains and the pitfalls of reification. The case of algebra, Educational Studies in Mathematics, 26, 191-228.

SHAHABI, S., y YANG, L. (1990, April). A comparison between two variations of multiple-choice items and their effects on difficulty and discrimination values. Paper presented at the annual meeting of the National Council on Measurement in Educational, Boston.

SHEPARD, L.A. (1993). The place of testing reform in educational reform-A reply to Cizek. Educational Researcher, 22, 10-13.

SHERMAN, S.W. (1996, April). Multiple-choice test bias uncovered by the uso of an "Y don't know" alternaative. Paper presented at the annual meeting of the American Educational Reseaarch Association, Chicaago.

SIEGLER, R.S. (1981). Developmental sequences within and between concepts. Monograph of the Society for Research in Child Development, 46.

SIJTSMA K., y HEMKER, BT (1998). Nonparametric Polytomous IRT Models for Invariant Item Ordering, with Results for Parametric Models. Psychometrika, vol. 63, pp. 183-200.

SIJTSMA, K., y JUNKER, B.W. (1996). A survey of theory and methods of invariant item ordering. British Journal of Mathematical ans Statistical Psychology, 49, 79-105.

SMITH, R.M., y KRAMER, G.A. (1990, April). An investigation of components influencing the diffiiiculty of form-development items. Paper presented at the annual meeting of the National Council on Measurement in Education, Boston

SMOLENSKY, P. (1986) Information processing in dynamical systems: Foundations of harmony theory. En D.E. Rummelhart, I.C. L. McClelland y The P.D.P. Research Group (Eds.) Parallel distributed processing. Vol. I: Foundations. Pp 194-281. Cambridge MA, MIT Press.

SNOW, R.E. (1993). Construct validity and constructed-response tests. In R.E. Bennett & W.C. Ward (Eds.), Constrction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment (pp. 45-60). Hillsdale, NJ: Lawrence Erllbaum Associates.

SNOW, R.E. Y LOHMAN, D.F. (1989). Implications of cognitive psychology for educational measurement. In R.L. Linn (Ed.) Educational measurement (3rd ed. pp 262-331). New York. MacMillan.

SPADA, H. (1976). Modelle des Denkens und Lernens (Models of thinking and learning). Bern: Huber. (In German)

SPADA, H. Y MC GAW, B. (1985). The assessment of learning effects with linear logistic test models. En S. Emberston (De.) Test Design: New directions in psychology and psychometrics (pp. 1669-193). New York. Academic Press.

SPEARMAN, C. (1904). The proof and measurement of association between two things, American Journal of Psychology, 15, 72-101.

SPEARMAN, C. (1907). Demostration of formulae for true measurement of correlation, American Journal of Psychology, 18, 161-169.

SPEARMAN, C. (1913). Correlations of sums and differences, British Journal of Psychology, 5, 417-426.

SPEARMAN, C. (1927). The Abilities of Man, Nueva York, MacMillan.

STEGELMANN,W. (1983). Expanding the Rasch model to a general model having more than one dimension. Psychometrika, 48, 259-267.

STERNBERG R. J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

STEVENS, S.S. (1946). On the theory of scles in measurement. Science, 103, 677-680.

SUTHERLAND R. (1987), A Longitudinal Study of the Development of Pupils' Algebraic Thinking in a Logo Environment. PhD Thesis, University of London Institute of Education.

SWAMINATHAN, H. y GIFFORD, J.A. (1983). Estimation of parameters in the three latent trait model. En D. Weiss (Ed.), New horizons in testing. New York: Academic Press.

SWAMINATHAN, H., y ROGERS, J. (1990). Detecting differential item functioning using logistic regression procedures. Journal of Educational Measurement, 27, 361-370.

SYMPSON, J.B. (1983, August). A new item response theory model for calibrating multiple-choice items. Paper presented at the annual meeting of the Psychometric Society, Los Angeles.

SYMPSON, J.B. (1983a). A new IRT model for calibrating multiple choice items. Paper presented at the annual meeting of the Psychometric Society, Los Angeles, CA

SYMPSON, J.B. (1983b, April). Extracting information from wrong answers in computerized adaptive testing. In New Developments in Computerized Adaptive Testing. Symposium conducted at the annual meeting of the american Psychological Association, Washington, D.C.

TATSUOKA, K. K. (1983a). Rule-Space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20, 345-354.

TATSUOKA, K. K. y LIN R.L. (1983b) Indices for detecting unusual patterns: lnks between two general approaches and optential applicacton. Applied Psychological Measurement, 7, 81-96.

TATSUOKA, K.K. (1990). Toward an integraation of item response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M.G. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition (pp. (453-488). Hillsdale, NJ: Lawrence Erlbaum Associates

TATSUOKA, K.K., y TATSUOKA, M.M. (1987). Bug distribution an datatern classification. Psychometrika, 52, 193-206.

TATSUOKA, K.K., BIRENBAUM, M., y ARNOLD, J. (1989). On the stability of students' rules of operation for solving arithmetic problems. Journal of Educational Measurement, 26, 351-361.

THISSEN, D. (1993). Repealing the rules that no longer apply to psychological measurement. in N. Frederiksen, R.J. Mislevy, & Y. Bejar (Eds.), Test theory for a new generation of tests.

THISSEN, D., STEIMBERG, L. y FIZPATRIK, A.R. (1989). Multiple- choice models: The distractors are also part of the item. Journal of Educational Measurement, 26, 161-176.

THISSEN, D., y STEINBERG, L. (1984). A response model for multiple- choice items. Psychometrika, 49, 501-519.

THISSEN, D., y STEINBERG, L. (1986). A taxomomy of item response models. Pshycometrika, 51, 566-577.

THOMAS M. y TALL D. (1986). The Value of the Computer in Learning Algebra Concepts, en Proceedings of the Tenth PME Conference, London, pp. 313-318.

THOMAS M. y TALL D. (1991)., Encouraging Veratile Thinking in Algebra Using the Computer, Mathematics Education Research Center, University of Warwick, Coventry, U.K.

THOMPSON, P.W. (1982). Were lions to spead, we wouldn't understand. Journal of Mathematical Behavior, 3, 147-165.

THORNDIKE E.L., COBB M.V., ORLEANS J.S., SYMONDS P.M., WALD E. y WOODYARD E. (1923). The Psychology of Algebra. The Macmillan Company, New York. Institute of Educational Research Teachers College, Columbia University.

TIKONOV, A. y ARSENIN, V. (1977). Solutions to ill-posed problems. Wiley, New York.

TORGERSON, W.S. (1958). Theory and methods of scaling. New York. John Wiley.

TRAUB, R.E. (1993). On the equivalence of traits assessed by multiple-choice and constructed-response tests. In R.E. Bennett & W.C. Ward (Eds.).

Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment (pp. 1-27).Hillsdale, NJ. Lawrence Erlbaum Associates.

TREVISAN, M.S., SAX, G., y MICHAEL, W.B. (1977). The effects of the number of options per item and sutdent ability on test validity and reliability. Educational and Psychological Measurement, 51, 829-837.

TRIGUEROS M. y CANTORAL R. (1992). On Students Understanding of Variation. Proceedings of the PME-NA Conference Blacksburg, Virginia, pp. 176-182.

TRIGUEROS M., REYES A., URSINI S. Y QUINTERO R. (1996a). Diseño de un Cuestionario de Diagnóstico acerca del Manejo del Concepto de Variable en el Álgebra, en Enseñanza de las Ciencias, 14 (3), pp. 351-363.

TRIGUEROS, M.; URSINI, S. & REYES, A. (1996b). College Students' Conceptions of Variable. Proceedings of the XX PME International Conference, Spain, pp. IV-315

TRIGUEROS, M. y URSINI, S. (1998). Dificultades de los Estudiantes Universitarios Frente al Concepto de Variable, in Hitt, F. (Ed.), Investigaciones en Matemática Educativa II, Grupo Editorial Iberoamérica.

TUCKER, L.R. (1946). Maximum validity of a test with equivalent items. Psychometrika, 11, 1-13.

URSINI S. (1994a), Ambientes Logo como Apoyo para Trabajar las Nociones de Variación y Correspondencia, en I. Simposio Metodología de la Enseñanza de las Matemátaicas, ITAM, PP. 124-128.

URSINI, S. (1994b). Pupils' Approaches to Different Characterisations of Variable in Logo, PhD Thesis, University of London Institute of Education.

URSINI S. (1996). Creación de un Potencial para Trabajar con la Noción de Variable, en Hitt F. (Ed.), Investigaciones en Matemática Educativa, Grupo Editorial Iberoamérica, pp. 423-440.

URSINI S. y TRIGUEROS M. (1997). Understanding of different uses of variable: A study with starting college students. Proceedings of the XXI PME International Conference, Finland, pp.

USISKIN Z. (1988). Conceptions of School Algebra and Uses of Variables, en Coxford A.F. y Shulte A.P. (Eds.), The Ideas of Algebra K-12, pp. 8-19.

VAN DER LINDEN, W. J. (1985). Optimun Design in Item Response Theory: Test Assembly and Item Calibration.

VAN DER LINDEN, W. J. y EGGEN T.J.J.M. (1986) An empirical Bayes approach to item banking. Applied psychological measurement, 10, 345-354.

VAN DER LINDEN, W. J., y BOEKKOOI-TIMMINGA, E. (1989). A maximum model for test design with practical constraints. Psychometrika, 54, 237-248.

VAN EGEN H. (1953). The Formation of Concepts, the Learning of Mathematics, en Febr H.F. (Ed.) The Learning of Mathematics: It's Theory and Practice, (Twenty-first Yearbook), Washington, D.C., National Council of Teachers of Mathematics, p. 69-98.

VERHELST, N.D., y GLAS, C.A. W. (1995). The one parameter logistic model. In G.H. Fischer y I. W. Molenaar (Eds.), Rasch models: Their foundations, recent development, and applications. New York: Springer.

VIGOSTKY, L.S. (1978). Mind in siociety: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

WAGNER S. (1981). An Analytical Framework for Mathematical Variables, Proceedings of the Fifth PME Conference, Grenoble, Francia, pp. 165-170.

WAGNER S. (1983). What are These Things Called Variables Variables?, Mathematics Teacher, October 1983, pp. 474-479.

WAINER, H. (1992) Measurement problems. Journal of Educational Measurement, 30, 1-22.

WAINER, H., y KIELY, G. (1987). Item clusters and computerized adaptive testing: A case for testlets. Journal of Educational Measurement, 24, 185-202.

WAINER, H., y THISSEN, D. (1993). Combining multiple-choice and constructed response test scores: Toward a Marxist theory of test construction. Applied Measurement in Education, 6, 103-118.

WHITELY, S.E. (1980) Multicomponent latent trait models for ability tests. Psychometrika.45,479-494.

WILSON, M. (1984). A Psychometric model of hyerarchical development. Unpublished doctoral dissertation, University of chicago.

WILSON, M. (1989). Saltus: A Psychometric model of discontinuity in cognitive development. Psychological Bulletin, 105 (2), 276-289.

WILSON M, y MISLEVY R.J. (1997)Marginal Maximum Likelihood Estimation for a Psychometric Model of Discontinuous Development, Psychometrika-Vol. 61, No. 1, 41-71.

WINSBERG, S., THISSEN, D. y WAINER, H. (1983). Estimation of the form of the item characteristic curve using monotone splines. Paper presented at the annual meeting of the Psychometric Society, Los Angeles, CA.

WRIGT, B. D. (1977). Solving measurement problems with the Rasch model. Journal of Educational Measurement, 14, 97-116. .

WRIGHT, B. D., y BELL, S.R. (1984). Items banks: what, why, how, Journal of Educational Measurement, 21, 331-346.

YAMAMOTO, K.(1987). A hybrid model for item responses. Unpublished dissertation. University of Illinois.

YAMAMOTO, K.(1989). HYBIL: A computer program to estimate HYBRID model parameters. Princeton, N.J: Educational Testing Services.

YAMAMOTO, K.(1990, april). Modelling the mixture of IRT and patterned responses by a modified HYBRID model. Paper presented at the 1990 American Educational Research Association Meeting in Boston.

YEN, W. M. (1993). Scaling performance assessments: Strategies for managing local item independence. Journal of educational Measurement, 30, 187- 213.

YOUSCHKEVITCH, A.P. (1976). The Concept of Function up to the Middle of the 19th Century, in Archives of History of Exact Science, Vol. 16, 1, pp. 37-85.

ZHU, C., BYRD, H., LU, P. y NOCEDAL, J. (1994) LBFGS-B: Fortram subroutines for large-scale bound constrained optimization, Report NAMII, EECS, Department, Northwestern University, USA.