



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE PSICOLOGÍA

DIVISIÓN DE ESTUDIOS DE POSGRADO

**EL CONOCIMIENTO EXPERTO EN LA
SOLUCIÓN DE PROBLEMAS DE
ESTADÍSTICA INFERENCIAL**

T E S I S

**QUE PARA OBTENER EL GRADO DE:
MAESTRO EN PSICOLOGÍA EDUCATIVA**

PRESENTA:

LUIS DOMINGO MÁRQUEZ RAMÍREZ

Directora de Tesis: DRA. BENILDE GARCÍA CABRERO
Miembro Del Comité: DRA. PATRICIA ANDRADE PALOS
Miembro Del Comité: DRA. LIZBETH O. VEGA PÉREZ
Suplente: DRA. ASUNCIÓN LÓPEZ MANJÓN
Suplente: MTRA. ESTELA JIMÉNEZ HERNÁNDEZ



MÉXICO, DF

JUNIO DE 2006



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS

Quiero expresar mi agradecimiento:

A mi Directora de Tesis: Dra. Benilde García Cabrero, por su asesoramiento y por brindarme la oportunidad de recurrir a su capacidad y experiencia científica en un marco de confianza, afecto y amistad, fundamentales para seguir creciendo intelectualmente.

A los miembros de mi comité de tesis: Dra. Patricia Andrade Palos, Dra. Lizbeth O. Vega Pérez, Dra. Asunción López Manjón y Mtra. Estela Jiménez Hernández, quienes me orientaron generosamente y compartieron conmigo este proceso enriquecedor.

A mis padres quienes me infundieron la ética y la motivación que guían mi transitar por la vida, por enseñarme que la perseverancia y el esfuerzo son el camino para lograr objetivos.

A mis hijas, Lore y Liss, mi más preciado tesoro, quienes siempre están presentes en mi corazón y con quienes quiero compartir este logro, que es también suyo.

A Ma. Luz Vega, Mariana Núñez y Sandra Jiménez por su incondicional colaboración para que este trabajo fuera llevado a buen fin.

Un especial agradecimiento por su valiosa colaboración e incondicional disposición a compartir su sabiduría al Dr. Javier Aguilar, Dra. Sofía Rivera, Mtra. Alejandra Valencia, Lic. Marcos Verdejo y Lic. Humberto Zepeda.

A Nidia Flores, Andrés Díaz, Elvia Ortega, Rosa Ma. Mendoza, Octaviano García, Vania Pineda, Estrella Alvarado, Israel Haro, Montserrat Rebolledo, Patricia Olivera, Jazmín M. del Castillo y Orlando Mejía, por apartarse un momento de sus actividades para hacer aportaciones importantes a este estudio.

A ¡A TODOS MUCHAS GRACIAS!

NOMBRE: Luis Domingo
 Máximo Ramírez
 FECHA: 13/ junio/ 2006
 FIRMA: 

INDICE

AGRADECIMIENTOS	5
RESUMEN	6
INTRODUCCIÓN	7
MARCO TEÓRICO	13
1. Dificultades en el Aprendizaje y la Enseñanza de la Estadística	13
1.1 Enseñanza de la Estadística Descriptiva	15
1.1.1 Poblaciones y muestras	16
1.1.2 El tipo de datos	16
1.1.3 La variación, la varianza y la desviación estándar	18
1.1.4 Correlación, causa y asociación	19
1.2 Enseñanza de la Estadística Inferencial	19
1.2.1 Cálculo de los intervalos de confianza para las medias muestrales centradas en μ .	20
1.2.2 La Prueba de Hipótesis	20
1.2.3. Significancia estadística vesus importancia práctica.	24
1.3 El Razonamiento Estadístico	25
1.4 Aplicación del Conocimiento Estadístico	26
2. La Investigación sobre la Solución de Problemas y los Estudios Experto-Novato	28
2.1 La Solución de Problemas	28
2.2 Los Estudios Experto-Novato	31
2.2.1 Diferencias en el conocimiento entre expertos y novatos	31
2.2.2 Diferencias cuantitativas entre expertos y novatos	33
2.2.3 Diferencias cualitativas entre expertos y novatos	35
3. La Aproximación del Análisis Cognitivo de Tareas al Diseño Instruccional	39
3.1 Métodos de Análisis Cognitivo de Tareas	41
3.1.1 Diferencias entre el Análisis de Tareas Tradicional y el Análisis Cognitivo de Tareas.	41
3.1.2 Análisis de Protocolos	46
3.1.3 Técnicas de Entrevista Cognitiva	46
3.1.4 Modelamiento Cognitivo y de Ejecución	47
3.2 Dominios de aplicación del Análisis Cognitivo de Tareas	48
3.2.1 Análisis Cognitivo de Tareas en Técnicos de Aviación	48
3.2.2 Ejemplo de la aplicación del procedimiento automatizado de análisis cognitivo de tareas DRE (Descomponer, Red, Evaluación)	52
4 La teoría de los esquemas y la representación del conocimiento	57
4.1 La teoría de los esquemas	57
4.2 Aprendizaje por modificación y generación de esquemas	59
4.2.1 Crecimiento	59
4.2.2 Ajuste	60
4.2.3 Reestructuración	60

MÉTODO	64
Objetivo	64
Preguntas de Investigación	64
Sujetos	64
Instrumentos	65
Procedimiento	65
1ª Fase: Orientación: Identificación de las Tareas Principales	68
Elaboración del espacio del problema	68
2ª Fase: Análisis Básico: Caracterización General de la Tarea y Componentes de la Pericia	68
3ª Fase: Refinamiento del Análisis: Caracterización de la Progresión Novato-Experto	70
Extracción del Conocimiento Conceptual, Procedimental y Estratégico en cada nivel de pericia	70
4ª Fase: Análisis de Datos y Representación del Conocimiento.	71
RESULTADOS	72
6.1 Primera Fase Orientación: Identificación de las Tareas Principales	72
6.1.1 El procedimiento para la elección de la prueba estadística	72
6.1.2 Conocimiento conceptual involucrado en la elección de la prueba estadística	73
6.2 Segunda Fase: Análisis Básico. Caracterización General de la Tarea y Componentes de la Pericia	78
6.2.1 Esquema del Espacio del Problema	78
6.2.2 Aspectos clave involucrados en la selección de la prueba estadística	82
6.3 Tercera Fase: Refinamiento del Análisis. Caracterización de la Progresión Novato-Experto	83
6.3.1 Tarea de clasificación de problemas estadísticos	84
6.3.1.1 Identificación de los conceptos clave para la clasificación de problemas estadísticos	84
6.3.2 Tarea de solución de un problema estadístico	92
6.3.2.1 Secuencia de solución de los expertos	93
6.3.2.2 Secuencia de solución de los novatos	99
6.4 Cuarta Fase Análisis de Datos y Representación del Conocimiento	105
6.4.1 Conocimiento conceptual, procedimental y estratégico involucrado en la selección de la prueba estadística	105
6.4.2 Principales errores conceptuales durante el proceso de elección de la prueba estadística	105
6.4.3 Aspectos problemáticos dónde se debe concentrar la enseñanza	107
DISCUSIÓN Y CONCLUSIONES	111
REFERENCIAS	121
APÉNDICE A. Guía de Preguntas para la Entrevista del Experto	125
APÉNDICE B. Lista de problemas estadísticos	127
APÉNDICE C. Solución de un problema estadístico mediante la aplicación	

<i>de ANOVA</i>	133
<i>APÉNDICE D Características generales de los doce problemas presentados.</i>	134
<i>APÉNDICE E Taxonomía de relaciones semánticas, adaptada de Lemke, (1997).</i>	136
<i>APÉNDICE F Conocimiento declarativo, procedimental y estratégico involucrado en la selección de la prueba estadística</i>	137
<i>Apéndice G Errores conceptuales en que incurren los novatos</i>	140

RESUMEN

El objetivo del presente estudio fue analizar el conocimiento experto en la solución de problemas de estadística inferencial. Se utilizó la metodología de Análisis Cognitivo de Tareas para identificar los componentes cognitivos involucrados en el proceso de la selección de la prueba estadística que utilizan tanto los expertos como los novatos en estadística en el contexto de una investigación en el campo de la Psicología. El proceso se desarrolló con base en las tres fases propuestas en el *Modelo de Análisis de Tareas Integrado* MATI, (Ryder & Redding, 1993), al modelo original se le agregó una cuarta fase (Análisis de Datos y Representación de l Conocimiento), la cuál constituye una fase analítica que resume los hallazgos de las fases anteriores. Para la realización de la entrevistas, se utilizó la técnica de análisis de *Protocolo en Voz Alta*, a través de una entrevista semi-estructurada denominada PARI (Prerrequisito, Acción, Resultado, Interpretación),(Lajoie, Azevedo, Zhao, Kinnon, & Fleiszer, 1997). Los resultados del Análisis Cognitivo de Tareas describieron las estructuras de conocimiento tanto de expertos y novatos, analizando de qué manera hacen uso los expertos los conceptos, los procedimientos y las estrategias, así como las diferencias que se presentan en los diferentes niveles de pericia, desde alumnos que se encuentran en tercer grado de licenciatura hasta nivel doctorado. Asimismo, se presentan sugerencias para diseñar un plan instruccional enfocado en la enseñanza de la estadística a nivel universitario.

INTRODUCCIÓN

La presente investigación tiene como propósito fundamental la elaboración de un modelo de conocimiento experto de la conducta de solución de problemas en el ámbito de la estadística, particularmente la estadística inferencial. La meta es describir las características de dominio en el área de la estadística y determinar los componentes subyacentes que determinan la ejecución experta. El modelo de ejecución experta en estadística permitirá contar con un marco de referencia para el desarrollo de materiales instruccionales y el planteamiento de directrices pedagógicas para el diseño instruccional en escenarios de enseñanza de la estadística en el ámbito universitario.

La adquisición de conocimientos y habilidades en estadística ha representado un problema en todos los niveles de enseñanza; la instrucción estadística se caracteriza por ser principalmente teórica. La comprensión de los conceptos y la adquisición de habilidades necesarios para tomar decisiones ante los problemas estadísticos, está fuera del alcance de la mayoría de los estudiantes, ya que la aproximación tradicional proporciona al estudiante básicamente recetas algorítmicas que no se relacionan con el razonamiento en problemas estadísticos reales (Hawkins, Jolliffe & Glickman, 1992). De la misma forma, las prácticas de estadística contemplan el análisis descriptivo e inferencial de datos, pero acompañadas de una explicación matemática descontextualizada.

Tanto las investigaciones en psicología, como en estadística y matemáticas, han proporcionado evidencias respecto al hecho de que los estudiantes no logran las metas instruccionales que se proponen los maestros (Hawkins, Jolliffe & Glickman, 1992). Las investigaciones psicológicas, al centrarse en la observación y descripción del funcionamiento cognitivo, sus éxitos y fracasos, han revelado que la gente posee concepciones erróneas o razonamientos inapropiados sobre las ideas estadísticas. Al respecto, Ávila, Márquez, y de la Rosa, (1999), elaboraron un diagnóstico del estado actual de la problemática en la enseñanza-aprendizaje de la estadística en la Facultad de Psicología de la UNAM. En el diagnóstico realizado, se detectaron las necesidades y dificultades que reportan los profesores durante el proceso de enseñanza de la estadística, destacando la opinión generalizada acerca de la falta de una relación fuerte entre la teoría y la práctica en la impartición de la asignatura. Otro aspecto que resultó problemático fue el alto índice de deserción y reprobación que se presentó en el periodo comprendido entre 1996 y 1998 en las materias de Matemáticas I y II, Estadística Descriptiva e Inferencial. En el apartado referente a la investigación en la enseñanza y aprendizaje de la estadística se proporcionará información más amplia sobre este estudio.

Es sabido que los estudiantes no parecen beneficiarse de las estrategias instruccionales utilizadas en la clase de estadística que desvinculan la teoría de la práctica. (Lajoie, Jacobs, & Lavigne, 1995). De la misma forma, la utilización de innumerables libros de texto, que presentan los contenidos estadísticos

organizados de acuerdo a la estructura de la disciplina, no han mejorado la comprensión ni el razonamiento, cuando se trata de resolver problemas en los que es necesario aplicar los conceptos y métodos estadísticos a problemas reales de las Ciencias Sociales. La aplicación práctica de los conocimientos estadísticos involucra un proceso complejo en el cual se toman múltiples decisiones que implican la utilización de gran cantidad de conceptos, procedimientos y estrategias de una manera integral y en contextos específicos.

Las matemáticas en general se han considerado tradicionalmente como una disciplina bien estructurada; por consiguiente, se ha establecido que los estudiantes sólo pueden tener respuestas correctas o incorrectas, sin considerar lo que sucede durante el proceso (Resnick, 1989). Esta misma idea se ha aplicado a la enseñanza de la estadística. La manera de abordar la disciplina ha descuidado el proceso complejo que siguen los estudiantes al resolver los problemas estadísticos con los que se enfrentan, al intentar aplicar los conocimientos. Pocas veces se les brinda la oportunidad de explicar su razonamiento cuando resuelven un problema estadístico. Por el contrario, se les pide que memoricen hechos y procedimientos (Bransford, Hasselbring, Barron, Kulewicz, Guzdial & Palincsar, 1989), en vez de activar sus capacidades de solución de problemas. Como consecuencia, los estudiantes, no comprenden los conceptos y no son capaces de aplicar el conocimiento adquirido en sus propias disciplinas y mucho menos extender el conocimiento a otras áreas.

El aprendizaje de los conceptos abstractos en el campo de las matemáticas y la estadística es difícil (Garfield & Ahlgren, 1994), desalentador y carente de significado, a menos que los métodos instruccionales hagan los contenidos menos ambiguos y más concretos, y esto ocurre con los estudiantes en todos los niveles de enseñanza. Para superar estos problemas de enseñanza, Lajoie, Jacobs y Lavigne (1995), proponen una alternativa diferente, tratar a la estadística como una disciplina poco estructurada, en la cual existe más de una respuesta correcta. Esto podría abrir las puertas a otras formas de aprendizaje y enseñanza que enfatizan el carácter constructivo y activo en vez del carácter receptivo tradicional. Las autoras afirman que la utilización de las aproximaciones de solución de problemas puede reforzar la naturaleza activa del aprendizaje de la estadística. Lajoie et al. se refieren al concepto de "empoderar" a los estudiantes, es decir, darles la oportunidad de aprender estadística "haciéndola" realmente, y así demostrar la relevancia de sus aplicaciones en el mundo real (Garfield & Ahlgren, 1988; Lajoie, Lawless, Lavigne & Munsie, 1993; Scheaffer, 1988).

Otro problema que se suma a los ya mencionados, es que en la mayoría de los programas de enseñanza de la Estadística, la estructura del currículo se basa en la propia estructura de la disciplina; dicho de otra manera, en la planeación de la enseñanza, no se ha tomado en cuenta la forma en que los estudiantes van asimilando los conceptos complejos, ni los errores conceptuales en que incurren durante el proceso de aprendizaje.

En las disciplinas científicas consideradas como “duras”, se ha intentado abordar el problema de la enseñanza y aprendizaje de contenidos complejos mediante la caracterización de la forma como piensan y resuelven los problemas los expertos de dichas disciplinas (Chi & Rees, 1983). La física es la disciplina que ha atraído el mayor número de estudios sobre el conocimiento experto; entre los más representativos se encuentran los realizados por Chi, Feltovich y Glaser (1981).

El estudio y caracterización de la ejecución experimentada en una disciplina específica (Chi, Glaser, & Farr, 1988), se ha realizado básicamente a través de los estudios experto-novato; este enfoque ha resultado prometedor para el mejoramiento de la enseñanza y aprendizaje de disciplinas complejas (Chi, Glaser, & Rees, 1982). Es bien sabido que los expertos difieren de los novatos en dos formas básicas. Los expertos tienen un conocimiento bien definido de su campo de pericia, y aplican más eficientemente los procedimientos de solución de problemas en la ejecución de la tarea (Chi & Glaser, 1985). También se sabe que las diferencias en la ejecución se deben, en parte a las diferencias en la organización del conocimiento. Por tanto, será necesario establecer en este estudio la caracterización de la organización del conocimiento experto en estadística y compararlo con diferentes niveles de pericia.

Sin embargo, cabe señalar que existe poco interés hacia los estudios experto-novato, para indagar los procesos que siguen los expertos en estadística hasta que logran un dominio pleno del campo de conocimiento, pasando por los diferentes niveles de pericia. Al revisar la literatura, se encontraron sólo algunos estudios que se han enfocado al estudio del proceso de convertirse en experto en estadística, (e.g. Alacaci, 2004 ; Gardner y Hudson, 1999; Lavigne y Glaser, 2001; Quilici y Mayer, 1996). La falta de investigaciones que enfoquen el estudio de la estadística desde esta perspectiva es una de las razones que motivó la realización del presente estudio.

Desde la perspectiva de la Psicología Cognitiva, la complejidad del proceso de toma de decisiones en la solución de problemas estadísticos, señalado anteriormente, requiere una metodología de estudio diferente. Un campo que se ha considerado de principal importancia en la Psicología Cognitiva es el análisis de tareas complejas. El análisis cognitivo minucioso tiene implicaciones para la determinación de las metas instruccionales y de la metodología de enseñanza a seguir, en vista de que permite contribuir al mejoramiento de la comprensión de los conceptos y fomentar el razonamiento estadístico. Bajo esta consideración, el presente estudio lleva a cabo el análisis de la tarea compleja de la solución de problemas estadísticos, que los profesionales de diversas disciplinas de las Ciencias Sociales enfrentan comúnmente, quienes hacen uso de la estadística como una herramienta metodológica de investigación en su propio campo de trabajo. Por tanto, en el presente estudio, se pretende indagar el conocimiento declarativo, procedimental, simbólico y estratégico involucrado, así como la

caracterización de los modelos mentales de la tarea estadística que usualmente emplean los expertos y los novatos, cuando intentan resolver un problema de estadística en Ciencias Sociales. Para la extracción, análisis y representación del conocimiento experto involucrado en esta tarea compleja, se utilizó una metodología denominada *Análisis Cognitivo de Tareas*. Una aplicación que puede derivarse del uso de esta metodología de análisis, es el desarrollo de directrices pedagógicas encaminadas al mejoramiento de la enseñanza de la estadística.

El Análisis Cognitivo de Tareas es una metodología utilizada con éxito para el diseño instruccional de conocimientos y habilidades en tareas complejas en el contexto real de la solución de problemas (Ryder & Redding, 1993). El objetivo principal de este estudio es elaborar un modelo cognitivo de la tarea compleja de la elección de la Prueba Estadística mediante la aplicación de la metodología basada en el Análisis Cognitivo de Tareas, considerando el proceso de solución de problemas en Estadística en el contexto de una situación real de investigación en el campo de las Ciencias Sociales.

El estudio de la toma de decisiones en estadística se considera un campo complejo y rico en conocimientos. Esta toma de decisiones involucra el uso de procedimientos específicos como el manejo de conceptos interrelacionados en una forma compleja, generalmente en el contexto de la solución de problemas aplicados en un área específica de conocimiento. Por tanto, al abordar el tema de la enseñanza de la estadística nos debemos remitir invariablemente al estudio de la solución de problemas, y más específicamente a la solución de problemas en el terreno de las matemáticas. En la segunda mitad de los 70's los investigadores centraron su atención en el estudio de las áreas de problemas ricos en conocimiento (para denominar a los campos que involucran el uso de una gran cantidad de conocimiento declarativo, procedimental y estratégico). Entre las áreas de problemas ricos en conocimiento estudiados, se incluyen álgebra, física, termodinámica, ajedrez, bridge, geometría, diagnóstico médico, formación en política pública, y programación en computadora (Van Lehn, 1996). Sin embargo, no existe un campo de estudios definido, que intente abordar la estadística como solución de problemas con toda la complejidad que ello implica. Esta línea de investigación intentaría determinar las características relevantes del razonamiento y la representación de los problemas estadísticos propias de los expertos y novatos, y daría cuenta de la ruta o el progreso que deben seguir los estudiantes aprendices en su camino para convertirse en expertos solucionadores de problemas estadísticos.

En los siguientes párrafos se describen los contenidos de cada capítulo para dar al lector una visión general de la forma en que se estructuró el contenido del presente estudio, con el objeto de aproximarse a la ruta de convertirse en experto en estadística, así como un conocimiento detallado de los conceptos, procedimientos y modelos mentales implicados en este proceso.

En el primer capítulo de la tesis se presenta una revisión de la investigación realizada en el ámbito de la enseñanza y aprendizaje de la estadística. En él se contemplan los estudios realizados sobre los problemas encontrados en la enseñanza de la estadística, así como las propuestas de diversos autores para mejorar y favorecer su aprendizaje. De la misma forma se abordan algunos errores conceptuales frecuentemente encontrados en la comprensión de la estadística descriptiva e inferencial.

En el segundo capítulo, la discusión se centra en el estudio de la aproximación de la solución de problemas y particularmente en la línea de investigación en la psicología cognitiva conocida como estudios *experto-novato*. El estudio de la solución de problemas estadísticos por expertos y novatos no ha sido ampliamente atendido por este tipo de investigaciones, a pesar de la complejidad de los conceptos y procedimientos que están involucrados en la toma de decisiones ante un problema estadístico.

En el tercer capítulo, se aborda el *Análisis Cognitivo de Tareas*, una metodología que aporta la psicología cognitiva con el objeto de distinguir las diferencias cualitativas entre expertos y novatos en la solución de un problema complejo, con un énfasis particular en las habilidades cognitivas requeridas y la formación de modelos mentales. Se establecen las características principales que lo distinguen del Análisis de Tareas Tradicional. Asimismo se revisan las diferentes modalidades o métodos específicos de Análisis Cognitivo de Tareas y se presentarán algunos ejemplos de su utilización en diversos ámbitos de conocimiento.

En el cuarto capítulo se aborda el tema de la teoría de los esquemas y la representación del conocimiento, conceptos que ayudan a esclarecer la perspectiva de la naturaleza del conocimiento experto y los estudios *experto-novato*.

En el quinto capítulo se describe la metodología a seguir para la construcción de un modelo experto de la solución de problemas estadísticos utilizando la metodología de *Análisis Cognitivo de Tareas*. La metodología propuesta ha sido adaptada del *Modelo de Análisis de Tareas Integrado (MATI)* de (Ryder & Redding, 1993) y se divide en tres fases principales: 1ª fase, "Orientación: Identificación de las Tareas Principales"; 2ª Fase: "Análisis Básico: caracterización General de la Tarea y Componentes de la Pericia"; 3ª Fase: Refinamiento del Análisis: Caracterización de la Progresión Novato-Experto".

En el sexto capítulo se aborda el análisis de resultados dividido en las fases correspondientes al modelo MATI, además de la cuarta fase: "Análisis de Datos y Representación del Conocimiento", la cual ha sido agregada al modelo original de Ryder y Redding (1993).

En el séptimo capítulo, Discusión y Conclusiones, se discuten las implicaciones

y alcances de la metodología propuesta para la conformación del modelo cognitivo experto de la solución de problemas estadísticos que implican la elección de la prueba estadística, para el diseño instruccional y el mejoramiento de la enseñanza y el aprendizaje de conceptos y procedimientos en estadística. Asimismo, se discutirá acerca del empleo extensivo de la metodología del Análisis Cognitivo de Tareas en otros campos del conocimiento. Finalmente se presentan las consideraciones finales, así como las limitaciones y sugerencias que podrán apoyar las futuras investigaciones en la materia.

MARCO TEÓRICO

1. Dificultades en el Aprendizaje y la Enseñanza de la Estadística

Es posible mejorar la enseñanza de la estadística, si los profesores analizan lo que realmente desean que sus alumnos conozcan como resultado de su clase y proporcionan las actividades de aprendizaje apropiadas para desarrollar la ejecución deseada. Al revisar la investigación existente acerca de lo que los profesores consideran que los estudiantes deben saber al finalizar un curso de introducción a la estadística, encontramos que entre los principales objetivos de enseñanza están de hecho, fomentar la comprensión de los principales conceptos y el pensamiento estadístico, antes que lograr que los alumnos calculen manualmente la desviación estándar, conozcan cómo convertir los valores normales a puntajes z utilizando las tablas, etc. Garfield (1992) presenta cinco puntos centrales, en los que propone enfocar la enseñanza de la estadística en las aulas:

1. Es importante aprender algunos aspectos fundamentales de estadística para comprender mejor y evaluar la información en el mundo.
2. Es posible aprender las ideas principales de la estadística trabajando consistentemente sobre ellas, aplicando hábitos adecuados de estudio y trabajando en conjunto con otros.
3. Aprender estadística significa aprender a comunicarse utilizando un lenguaje estadístico, resolviendo problemas estadísticos, estableciendo conclusiones y argumentando con fundamentos sólidos.
4. Existen diferentes formas de resolver los problemas estadísticos y por tanto, se requiere enseñar a los alumnos a pensar sobre diferentes alternativas de solución.
5. La gente puede llegar a conclusiones diferentes basadas en los mismos datos, si tienen diferentes suposiciones y utilizan diferentes métodos de análisis, lo cual es necesario incluir en las propuestas de enseñanza de la estadística.

Otro aspecto a considerar es cómo lograr que los estudiantes aprendan los conceptos y procedimientos estadísticos que se proponen los profesores y hacer que cambien sus creencias actuales. La forma como se imparte usualmente la clase de estadística es escuchar la clase teórica, resolver los ejercicios del libro de texto y practicar en el laboratorio de cómputo (Garfield, 1992). Pero necesitamos preguntarnos si estas actividades realmente ayudan a lograr las metas planteadas con los alumnos y si éstos se encuentran adecuadamente

preparados para utilizar el razonamiento estadístico para recolectar y analizar datos, así como para comunicar escrita y verbalmente los resultados al resolver problemas estadísticos reales. Los datos derivados de las investigaciones revelan que los estudiantes no están aprendiendo lo que se plantea en las metas instruccionales (Garfield, 1992).

La investigación sobre la enseñanza de la estadística ha proporcionado algunos procedimientos instruccionales efectivos. Por ejemplo, Garfield (1992) señala algunas indicaciones para ayudar a los estudiantes a comprender la estadística:

- Los cursos que se basan en actividades diseñadas para grupos pequeños ayudan a los estudiantes a superar algunas concepciones erróneas y fomentan la comprensión de los conceptos estadísticos.
- Evaluar a los estudiantes proporcionándoles “retroalimentación correctiva” sobre sus concepciones erróneas, seguidas de actividades correctivas (alentando a los estudiantes a explicar las soluciones, suponiendo respuestas antes de efectuar los cálculos, y regresando a sus respuestas para determinar si tienen sentido).
- Es posible enseñar los conceptos probabilísticos (referentes al heurístico de representatividad), haciendo que los estudiantes hagan predicciones antes de recolectar los datos para resolver problemas de probabilidad y luego comparando los resultados experimentales con sus predicciones originales.
- El uso de las simulaciones en computadora ayuda a los estudiantes a proporcionar respuestas correctas a una diversidad de problemas de probabilidad. La utilización de software permite visualizar e interactuar con los datos y mejora la comprensión de los conceptos y su correspondiente análisis.

Otra línea de investigación aborda la evaluación del conocimiento estadístico. Al respecto, Hubbard (1997) afirma que la evaluación puede utilizarse como una herramienta poderosa para alentar a los alumnos a adoptar estrategias de aprendizaje profundas en vez de superficiales, basado en la premisa de que la evaluación influye en el aprendizaje. Hubbard señala que la mayoría de las preguntas en las evaluaciones tradicionales tienden a reforzar la memorización de los procesos y no la comprensión de conceptos, por contraparte propone algunas técnicas para construir preguntas que evalúen la comprensión enfocándose en metas específicas, por ejemplo:

- a) Pedir a los estudiantes que elaboren la pregunta.
- b) Sugerir que algunos aspectos de la situación actual pueden cambiar y preguntar a los alumnos que expliquen cómo los cambios afectan la solución.
- c) Relacionar las representaciones gráficas y simbólicas de un concepto.

Por otra parte, Verkoeijen, Imbos, Van de Wiel, Berger, y Schmidt, (2002) proponen el método de recuerdo libre para determinar la calidad de las representaciones en estadística de los estudiantes.

La mayor parte de la investigación sobre la enseñanza y el aprendizaje de la estadística se ha centrado en estudiar los errores conceptuales que presentan los estudiantes cuando se abordan los temas referentes a Estadística Descriptiva e Inferencial. A continuación se presentan algunos errores en los que se incurre comúnmente al abordar los temas representativos de estas dos grandes áreas.

1.1 Enseñanza de la Estadística Descriptiva

De acuerdo con Hawkins, Jolliffe y Glickman, (1992), los temas que se abordan en la mayoría de los programas de estadística descriptiva son los siguientes:

- la colección, organización, tabulación y representación gráfica de datos;
- las estadísticas correspondientes a las medidas de tendencia central y de dispersión;
- la búsqueda de grupos modales u otros patrones en los datos (actividades que caracterizan fuertemente la fase inicial del análisis estadístico de datos AED). Una actividad típica es el Análisis Exploratorio de Datos en donde se investigan las relaciones lineales en los datos bivariados.

Sin embargo, existen algunos conceptos que presentan problemas para su comprensión. Hawkins et al. (1992), proporcionan una lista de los errores conceptuales en que incurren los alumnos comúnmente en el salón de clases:

- a) cualquier diferencia en las medias entre dos grupos se considera significativa
- b) no se toma en cuenta la variabilidad en el "mundo real"
- c) se presenta una confianza no fundamentada en las muestras pequeñas
- d) existe poco respeto por las diferencias pequeñas en muestras aleatorias grandes
- e) el tamaño de una muestra debería estar directamente relacionada al tamaño de la población.

Shoughnessy (1992) señala además otro error conceptual:

- f) se presenta una falta de conciencia del fenómeno de regresión a la media en la vida diaria.

A continuación se describen con detalle algunos de los tópicos problemáticos para el alumno desde el punto de vista conceptual.

1.1.1 Poblaciones y muestras

Una de las primeras dificultades enfrentadas por los estudiantes principiantes es la comprensión del uso estadístico de los términos *población* y *muestra*. Para una persona común, la palabra "población" se refiere a la gente que vive en alguna región, puede ser tan grande como un continente o tan pequeña como una aldea. La palabra "muestra" tiene una interpretación general más amplia, encontrándose en tales contextos como una encuesta *muestral* (considerada como una muestra de gente), *muestras* libres de productos de consumo y *muestras* de sangre y orina en las investigaciones médicas. Población y muestra pueden significar lo mismo para los estadísticos, pero también mucho más. Una *población* es el grupo completo de personas, objetos o mediciones en los que estamos interesados. Por otra parte, una muestra es parte de esa población, significa implícitamente que es una representación de esa población.

La imagen mental del proceso de muestreo que proviene de la experiencia en la vida cotidiana, es por lo general, la de tomar una muestra de elementos *físicos*. Sin embargo, en estadística se estudian las mediciones en la muestra (o población) por lo tanto, los términos "población" y "muestra" en estadística se refieren a colecciones de *observaciones* o *mediciones*.

En los cursos de estadística, casi simultáneamente se incluye el término de *variable aleatoria*, donde las *poblaciones* y *muestras* se definen en términos de distribuciones, la *población* se describe por sus *parámetros* y la muestra por sus *estadísticos*. Muy cercano a esos conceptos está la idea de *distribuciones muestrales*, que son las distribuciones derivadas de estadísticas del muestreo. Los estudiantes regularmente confunden estos tres tipos de distribución, probablemente porque sus descriptores tienen tanto en común, por ejemplo: confunden μ con \bar{X} (media de la población con media muestral) y confunden σ^2 con s^2 (*varianza de la población* con *varianza de la muestra*). Otro concepto que representa un problema para su comprensión es el de *error estándar* para representar la *desviación estándar* de la distribución muestral.

1.1.2 El tipo de datos

Uno de los primeros tópicos que se introducen en el programa de estadística descriptiva tiene que ver con los tipos de datos. La clasificación generalmente se lleva a cabo de acuerdo a dos tipos de datos: *cualitativos* y *cuantitativos*. Básicamente con los datos *cualitativos* los valores son esencialmente palabras, que pueden indicar las categorías que las definen. Con los datos *cuantitativos* los valores son atributos numéricos. Los datos *cualitativos* se convierten comúnmente a códigos numéricos, pero cualquier aritmética que se haga con los códigos no tendrá ningún significado, de igual forma que no se puede hacer aritmética con palabras. Por ejemplo con los valores de la variable *nominal* "estado civil" (soltero, casado, divorciado, viudo) se puede asignar un valor (1, 2, 3, 4, respectivamente), pero no podemos sumar solteros con casados para

obtener una medida.

El concepto *ordinal* se usa cuando hay una continuidad subyacente tras las palabras; por ejemplo: las respuestas a las preguntas: ¿Qué tan frecuentemente tiene problemas para dormir?, pueden etiquetarse como "todas las noches", "la mayoría de las noches", "algunas veces", "raramente", y "nunca". La correspondencia entre esas palabras no está claramente definida, aunque se puedan agrupar perfectamente por su orden en una escala, no se pueden establecer las distancias entre esas etiquetas.

Hawkins et al. (1992), plantean una distinción clara entre los datos *cualitativos* y *cuantitativos*: los datos *cualitativos* pueden proporcionar conteos de frecuencia para las categorías, los datos *cuantitativos* pueden dar esa posibilidad, pero tienden a ser tratados como mediciones o puntajes. Los datos *cuantitativos* (numéricos) pueden ser discretos (por ejemplo, el número de niños en una familia) o continuos (por ejemplo, la altura de un sujeto experimental en centímetros), pero los estudiantes principiantes generalmente piensan en datos numéricos discretos y datos categóricos como si fueran de un mismo tipo. Estas concepciones erróneas también se refuerzan por el hecho de que estos dos tipos de datos pueden tener representaciones gráficas similares, por ejemplo datos representados con una gráfica de barras, donde las barras están separadas para indicar categorías diferentes o lo discreto de los valores numéricos. Para superar esta confusión y enfatizar la distinción entre los datos categóricos y discretos, debemos inclinarnos por la representación gráfica de la gráfica de barras para datos categóricos, mientras que para los datos discretos podemos seleccionar preferiblemente las gráficas lineales. Por otra parte, un diagrama de tallo-hoja puede usarse tanto para datos discretos como continuos en forma bruta.

En referencia específica a los gráficos, Hawkins et al. (1992) afirman que en los gráficos lineales, la altura y anchura son proporcionales a la frecuencia, mientras que en el histograma, el área es proporcional a la frecuencia, esto quiere decir que, la altura de los bloques de anchura igual en el histograma es proporcional a la densidad de la frecuencia. Esto ayudaría a los estudiantes a comprender las propiedades y manipulaciones de las distribuciones probabilísticas que se van a encontrar posteriormente en el curso.

En libros de texto podemos encontrar que el concepto discreto significa números enteros, mientras que continuo puede significar fracciones decimales. Sin embargo, esta definición puede conducir a errores conceptuales en los estudiantes. Tomemos, por ejemplo, el caso de los números de preguntas respondidas correctamente en una prueba de ortografía. Se dice que esto representa números enteros. Sin embargo, representar los mismos datos como la proporción de respuestas correctas puede originar una concepción errónea, ya que, es claro que 21 respuestas correctas de 30 respuestas es un valor discreto que puede representarse como una proporción de 70%. Una expresión más

apropiada es que el concepto discreto se aplica cuando todos los valores posibles están separados uno del otro por valores imposibles, mientras que continuo se aplica cuando existen valores posibles entre dos medidas, por muy finos que éstos sean.

Muchos cursos introductorios de estadística incluyen métodos para el uso de datos en el nivel de escalas de intervalo, las escalas de razón se emplean raramente. Los datos intervalares implican unidades de discriminación iguales en un rango determinado y por lo tanto poseen la propiedad de aditividad, la cual es de suma importancia para computar los datos derivados de las estadísticas, tales como la media aritmética.

Una práctica de salón de clase, recomendada por Hawkins et al. (1992), para demostrar las propiedades de las escalas de intervalo, es introducir la distinción entre la rama de la estadística conocida como *paramétrica* (basada en suposiciones concernientes a la distribución de la población, descrita por sus parámetros) y *no paramétrica* (basada en procedimientos de inferencia que están libres de muchas suposiciones acerca de los parámetros de la población) para proporcionar una comprensión intuitiva básica de los tipos de suposiciones que pueden o no hacerse acerca de las distribuciones poblacionales, y los posibles errores que pueden dar lugar cuando se violan tales restricciones.

1.1.3 La variación, la varianza y la desviación estándar

Pareciera que explicar a los estudiantes lo que significa el concepto de *variación* es una tarea fácil; sin embargo, podemos ver que dos distribuciones pueden tener la misma *media* y *simetría* de formas y no obstante, diferir en su *dispersión* acerca de la media. A los estudiantes les resulta difícil comprender la noción de *variación* y las mediciones relacionadas, la *varianza* y la *desviación estándar*.

La *variación* es uno de los conceptos fundamentales, si no es que el concepto más importante, en estadística (Hawkins et al., 1992). Una barrera para la comprensión de la noción de *variación* puede ser, la introducción temprana en el programa de estadística descriptiva de la fórmula de la *variación* de un conjunto de datos y la falla al proporcionar respuestas adecuadas para las preguntas inevitables como: ¿por qué desviaciones al cuadrado? ¿por qué dividir por $n-1$? (Garfield & Burril, 1996).

Ambas medidas, la *desviación de la media* y la *desviación estándar*, son simples intentos de encontrar un enunciado representativo sobre de la cantidad de dispersión típica de la media, mientras que la suma de desviaciones, ignorando su dirección (positiva o negativa), y la suma de desviaciones al cuadrado, son enunciados que intentan representar la cantidad total de dispersión en referencia a la media.

1.1.4 Correlación, causa y asociación

Es importante enfatizar que la noción de correlación es distinta a la de la causalidad (Hawkins et al., 1992). Existe una importante confusión, al describir la *causalidad* como *correlación*. Este error conceptual involucra la atribución de una dependencia causal entre dos variables, X y Y , es decir, como resultado de una alta correlación entre las dos. La literatura está repleta con ejemplos de correlaciones espurias, en las cuales una variable está altamente correlacionada con otra, simplemente debido a que alguna está correlacionada con un tercer factor subyacente no observado. Por ejemplo, puede haber una alta correlación entre *habilidad de lenguaje* y *estatura*, como consecuencia del hecho, de que las dos se incrementan con la edad (el tercer factor subyacente).

1.2 Enseñanza de la Estadística Inferencial

La inferencia estadística que se enseña en el currículum escolar, comprende principalmente los conceptos de estimación del intervalo de confianza y prueba de hipótesis basados en suposiciones de muestras normales o grandes. En los programas de enseñanza, la estadística inferencial es probablemente la más propensa a la aparición de concepciones erróneas (Hawkins et al., 1992).

Los conceptos de la estimación paramétrica y la prueba de hipótesis que involucran: la comprensión de variables aleatorias, las distribuciones muestrales y la probabilidad han resultado ser sumamente difíciles de comprender para los estudiantes, sin embargo son muy importantes para dominar la compleja tarea involucrada en la realización de la inferencia estadística.

Para comprender los métodos tradicionales de la estadística inferencial, el estudiante debe tener una noción adecuada de la distribución muestral de un estadístico. Esto se relaciona con el concepto de la distribución muestral de la media, X (X mayúscula). Durante las primeras semanas de una clase tradicional de estadística, el estudiante se ocupa del cálculo de un estadístico descriptivo llamado distribución de la media x (x minúscula). Sin embargo, en las siguientes semanas del curso, el estudiante se encuentra con un valor numérico específico, X -*barra* es ahora un estadístico muestral, es decir, una variable aleatoria que tiene una distribución probabilística por sí misma. Generalmente el estudiante cubre el tema de las distribuciones muestrales realizando algunos ejercicios simples reforzados con alguna simulación (en ocasiones apoyada en computadora). Sin embargo, para comprender los conceptos de variable aleatoria en general y un simple estadístico en particular, se requiere una gran cantidad de tiempo de aprendizaje adicional al que usualmente se le proporciona al estudiante. Un programa de estadística tradicional aborda los métodos descriptivos como uno de sus temas iniciales, incluyendo entre sus contenidos principales la diferencia entre el *parámetro* de la población y un *estadístico* muestral; sin embargo, estos conceptos importantísimos se tratan de una manera superficial. El primero está basado en un conjunto completo de mediciones; el segundo en un subconjunto de mediciones tomadas del conjunto

en su totalidad y desafortunadamente, por lo general, los estudiantes no pueden establecer la diferencia entre ellos.

En el contexto de la inferencia estadística, el estudiante debe adquirir la noción del valor de un estadístico muestral en referencia a una variable aleatoria. Por otra parte, el verdadero parámetro de la población, (por ejemplo la media μ) es una constante, cuyo valor es desconocido y sobre el cual se desean hacer inferencias. De esta forma, hay una diferencia intrínseca entre el parámetro de la población y un estadístico muestral. La falta de comprensión de estas diferencias esenciales conduce a los errores comunes cometidos por los principiantes en estadística. Por ejemplo, en la estimación del intervalo de confianza encontramos estudiantes haciendo aseveraciones como:

$P(75 < \mu < 85) = 0.95$ - tratando una constante μ como una variable aleatoria;

El intervalo de confianza para $\mu = x$ más/menos, etc. -en este caso el signo "igual" es erróneo.

1.2.1 Cálculo de los intervalos de confianza para las medias muestrales centradas en μ

Los estudiantes no tienen problema para entender que X es la media muestral y μ es el parámetro de la población correspondiente. Sin embargo, no comprenden inmediatamente por qué no se puede establecer una hipótesis acerca de las variables aleatorias o por qué no se pueden hacer aseveraciones probabilísticas acerca de constantes fijas (por ejemplo, μ).

En el caso de la estimación del intervalo de confianza, es difícil para el estudiante principiante tratar el intervalo muestral como una variable aleatoria. Es mucho más fácil simular 100 intervalos de confianza y mostrar que aproximadamente 100 $(1 - \alpha)$ por ciento contiene el parámetro donde $(1 - \alpha)$ es el nivel de confianza. Sin embargo, a pesar de la evidencia de los ejercicios de simulación, el concepto de variable aleatoria se mantiene poco claro para los estudiantes.

1.2.2 La Prueba de Hipótesis

La prueba de hipótesis es la mayor fuente de confusión y concepciones erróneas (probablemente aún más que la estimación). Adicionalmente, existe una controversia acerca del papel de la prueba de hipótesis en el método científico en general. El primer punto a subrayar es que, el establecimiento de hipótesis y el criterio de toma de decisión se deben presentar antes de la recolección de datos y de cualquier análisis de datos. Sin embargo, en los ejemplos usados generalmente en las clases de estadística, los profesores proporcionan los datos como parte del ejemplo, y a menudo se asigna también la hipótesis. Esto significa que los estudiantes pondrán poca atención a las advertencias acerca

del orden de los pasos en la prueba de hipótesis. Hawkins et al. (1992) plantean la necesidad de una aproximación más imaginativa a los ejemplos, dividiendo las etapas en partes separadas. Así la primera etapa podría ser una discusión de los problemas de investigación y una formulación de hipótesis sin referirse a los datos.

Después de esta etapa, se formula una hipótesis como la siguiente:

$$H_0: X = 80; H_1: X > 80$$

Donde 80 es la media muestral observada, esto se ve inmediatamente como inadmisibile. Algunas veces los estudiantes escriben hipótesis en términos de **X-barra** usando la notación incorrecta para la media μ , esto refleja que no han comprendido la noción de la prueba de hipótesis. En un nivel más profundo, la prueba de hipótesis involucra probabilidades condicionales sutiles cuya manipulación correcta es esencial para la adecuada interpretación de los resultados. La confusión relacionada con estas probabilidades condicionales conduce a mayores errores y concepciones erróneas. Para profundizar en esto, en la figura 1 se muestra un diagrama de la representación de los procesos de prueba de hipótesis, el cual es útil para mostrar a los estudiantes los siguientes aspectos:

1. El valor del parámetro de la población es desconocido. Se postulan dos hipótesis relacionadas al parámetro, pero no se tiene la seguridad de cual es la correcta.
2. El estadístico sirve para inferir datos acerca de la población. Esto sólo puede hacerse sobre la base de la información parcial representada por una muestra de observaciones tomadas de la población. El estadístico puede, por tanto ser erróneo.
3. En la figura 1 podemos observar que son posibles dos tipos de error; si se decide rechazar H_0 cuando H_0 es verdadera, éste es un error tipo 1; si se decide no rechazar H_0 cuando H_1 es verdadera, este es un error tipo 2.

		Decisión estadística basada en la muestra de observaciones	
		No rechazar H_0	Rechazar H_0
Estado verdadero de la población	H_0	Correcto	Error tipo 1
	H_1	Error tipo 2	Correcto

Figura 1 Inferencia Estadística Clásica (Tomada de Hawkins, A. ; Jolliffe, F. y Glickman, L., 1992, p. 85).

Como podemos notar, tales eventos como "error tipo 1" y "error tipo 2" son resultados posibles en el proceso de toma de decisiones y pueden tener probabilidades asociadas, expresadas de la siguiente forma:

Pr (Error tipo 1)

Esto se conoce como el "*nivel de significancia*" de la prueba de hipótesis y se denota regularmente como α . (el *nivel de significancia* se cita como 100α por ciento). Algunas veces los términos "*error tipo 1*" y α se consideran erróneamente como la misma cosa. Aún cuando los estudiantes tengan el conocimiento de que α es una probabilidad, la describen como:

Pr (Error tipo 1)

Esta afirmación no expresa claramente la naturaleza condicional de esta probabilidad. Como se puede observar en la figura 1, la forma más apropiada es la siguiente expresión:

Pr (Rechazar H_0 / H_0 es verdadera)

para α .

Es importante tener en mente que los estudiantes pueden encontrar dificultades con la aproximación clásica de la prueba de hipótesis, porque sus intuiciones los llevan a buscar instancias confirmatorias, en lugar de tratar de negar la proposición. Evans y St. (1986) sugieren una explicación alternativa para esta dificultad: "la predisposición confirmatoria de los estudiantes puede no deberse a la tendencia a buscar confirmaciones, sino a la inhabilidad de los estudiantes para formular pruebas negativas".

Quizás debido a las dificultades anteriores, en la prueba de hipótesis se tiende a aplicar una aproximación casi algorítmica de "libro de cocina". Tanto los libros de texto como los profesores se refieren frecuentemente a "una receta". Por ejemplo, Johnson (1981), sugiere que la enseñanza de la prueba de hipótesis se da en seis pasos, como se muestra en la figura 2.

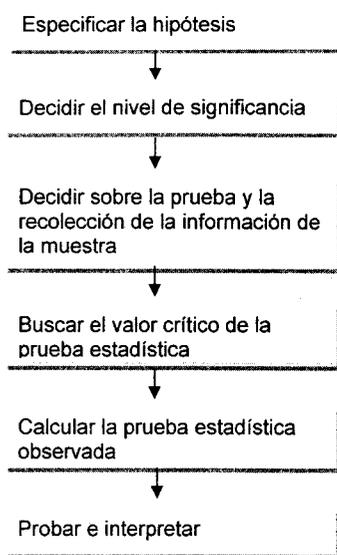


Figura 2. Un marco para la prueba de hipótesis (Tomada de Hawkins, A. ;Jolliffe, F. y Glickman, L., 1992, p. 86).

Para el estudiante que utiliza la prueba de hipótesis como parte de un método clásico de investigación científica, los seis pasos deben ser directivos en el sentido de que se debe asegurar el establecimiento de la hipótesis antes de recolectar los datos y así proporcionar una protección contra la "manipulación de datos". Sin embargo, suponen incorrectamente que el proceso es mecánico, y que el nivel de significancia escogido en el nivel 2 (a menudo uno de los valores "mágicos" de 1 y 5 por ciento) tiende a ser el único indicador de que una hipótesis puede ser retenida o rechazada. En contraparte, es preferible, citar un valor- p (la probabilidad de que un resultado igual o más extremo que el observado podría haber resultado por casualidad) y se pueden encontrar herramientas para hacerlo en muchos paquetes estadísticos. Un valor- p pequeño sugiere que la hipótesis nula debe ser rechazada. Sin embargo cuando se trata de pruebas de dos colas, los estudiantes tienen problemas para llegar a interpretar los resultados y realizar conclusiones.

El último paso de la prueba de hipótesis es la decisión de la prueba estadística a aplicar. Generalmente se piensa que este proceso de decisión para un grupo de datos determinado es puramente mecánico; sin embargo, aún en este caso

existen muchas respuestas y supuestos que probar, por ejemplo: ¿estamos tratando con una o más poblaciones? ¿podemos asumir que las poblaciones están normalmente distribuidas? ¿el tamaño de la muestra es grande?, etc.

1.2.3. Significancia estadística versus importancia práctica.

Gardner y Altman (1989), afirman que el procedimiento de prueba de hipótesis de seis etapas (ver figura 2) es inadecuado. Los autores señalan que, más importante que saber si los resultados son o no *estadísticamente significativos*, es saber si son *estadísticamente significativos* o son *prácticamente significativos*.

Según Gardner y Altman, cuando se lleva a cabo una prueba de hipótesis y se llega al paso de la decisión en la “receta” de seis pasos, se puede concluir que la prueba estadística es *significativa* y rechazar H_0 . Se hace referencia a estas pruebas como pruebas de *significancia*. En su forma de lenguaje natural el término “significativo” tiene la connotación de “importante” y los estudiantes inmediatamente asocian el resultado de una prueba significativa (usualmente una diferencia significativa) con un resultado que sugiere una diferencia importante en el contexto amplio de investigación.

En los libros de texto se advierte a los estudiantes (Daniel, 1977) que deben diferenciar cuidadosamente entre *significancia estadística* y *significancia práctica*. Al respecto, resulta útil recordar que el objetivo de la prueba de hipótesis es llegar a una conclusión acerca de una población sobre la base de la evidencia encontrada en una muestra de la población.

Daniel (1977) sugiere que se debe distinguir entre el concepto de *significancia estadística* y el de *significancia práctica* de la siguiente manera: Una prueba de hipótesis revelará *significancia estadística* la cual es aplicable a los resultados de la muestra. Ésta no puede detectar *significancia práctica*, la cual se aplica a condiciones existentes en la población. Daniel afirma que no se debe usar la palabra “significativo” cuando se establece la hipótesis nula, como en:

H_0 : No hay diferencia significativa entre ...

Los investigadores que hacen esto parecen estar pidiendo la subsecuente prueba estadística para indicar la presencia o ausencia de *significancia práctica*, y esto es algo que las pruebas no pueden hacer. Siempre se debe recordar que las hipótesis nulas son afirmaciones acerca de las poblaciones. Por tanto, éstas no deben contener referencias a las muestras.

La problemática de la adquisición de conocimientos en estadística no sólo ha sido abordada desde la perspectiva de los errores conceptuales de los estudiantes e investigadores. Existe un campo muy importante de investigación que surge de la ciencia cognitiva, éste es el estudio de la solución de problemas en diversos ámbitos, los estudios más significativos se han llevado a cabo

analizando tareas complejas, tales como ajedrez, física, química, medicina. El tema de la solución de problemas estadísticos puede ser abordado desde esta perspectiva, tomando en consideración que generalmente la estadística se aborda en el contexto de un problema no estructurado. La solución de problemas estadísticos es además una tarea compleja por la gran cantidad de cursos de acción que es posible tomar en un proceso de decisión que involucra el dominio de innumerables conceptos, procedimientos y estrategias, para llegar a diferentes metas posibles. Por tanto es conveniente abordar la estadística como una disciplina compleja que involucra la solución de problemas y la toma de múltiples decisiones. Por la importancia que reviste el estudio de la adquisición de habilidades y conocimientos en estadística desde la perspectiva de la solución de problemas, el tema se tratará en una sección independiente.

1.3 El Razonamiento Estadístico

De acuerdo a Garfield y Gal (1997), el razonamiento estadístico puede definirse como la manera en que la gente razona con ideas estadísticas y adquiere el sentido de la información estadística. Esto involucra las interpretaciones basadas en grupos de datos, representaciones gráficas y resúmenes estadísticos. Gran cantidad del razonamiento estadístico combina ideas acerca de los datos y la probabilidad, lo cual conduce a hacer inferencias y a interpretar los resultados estadísticos. Subyacente a este razonamiento está la comprensión conceptual de ideas importantes, tales como distribución central, dispersión, asociación, incertidumbre, aleatoriedad y muestreo.

Por su parte, Chervany, Collier, Fienberg, Johnson, y Neter (1977) y Chervaney, Benson, e Ilyer (1980) definieron el razonamiento estadístico como lo que un estudiante es capaz de hacer con el contenido estadístico (recordar, reconocer, y discriminar entre los conceptos estadísticos) y las habilidades que los estudiantes demuestran al utilizar los conceptos estadísticos en los pasos de un problema específico. Estos autores consideraron el razonamiento estadístico como un proceso de tres pasos:

- Comprensión (viendo un problema particular como similar a una clase de problemas)
- Planeación y ejecución (aplicando métodos apropiados para resolver un problema), y
- Evaluación e interpretación (interpretación de resultados y su relación con el problema original)

Sin embargo, Hawkins, Jolliffe y Glickman (1992) reconocen que se sabe poco acerca de los procesos de razonamiento estadístico o acerca de cómo los

estudiantes aprenden las ideas estadísticas.

Nisbett (1993) editó una colección de investigación conducida por él mismo y muchos otros colegas sobre el razonamiento estadístico y las formas de entrenar a la gente para seguir las reglas de razonamiento. Esta colección presenta una serie de generalizaciones basadas en la investigación en esta área, las cuales se resumen a continuación:

1. La gente tiene reglas intuitivas que aplica a algunos problemas estadísticos en la vida diaria.
2. Hay diferencias individuales en el grado en que la gente comprende estas reglas y las aplica a problemas concretos.
3. La instrucción en estadística cambia la manera en que la gente ve el mundo.

Nisbett afirma que la gente puede mejorar su razonamiento estadístico si aprende las reglas estadísticas y que el uso que hace cada quién de las reglas estadísticas se puede mejorar por medio de la instrucción directa, que esta instrucción puede ser abstracta, y que la gente puede ser instruida acerca de cómo "decodificar el mundo" para facilitarle la aplicación de estas reglas. El autor admite que los profesores pueden mejorar la enseñanza de las reglas estadísticas.

La investigación de este autor acerca de la evaluación del razonamiento estadístico (Garfield, 2002), reveló que los estudiantes pueden desempeñarse bien en un curso de estadística, obtener buenas calificaciones en su tarea, exámenes, y proyectos y a pesar de todo, tener calificaciones bajas en una prueba de razonamiento estadístico, tal como "Evaluación del Razonamiento Estadístico" (Garfield & Ahlgren, 1988). Estos resultados sugieren que los profesores de estadística no enseñan a los estudiantes específicamente cómo usar y aplicar los tipos de razonamiento. En su lugar, la mayoría de los profesores tienden a enseñar conceptos y procedimientos, proporcionar oportunidades a los estudiantes para trabajar con datos y software, y esperar que el razonamiento se desarrolle como resultado. Sin embargo, como afirma Garfield (2002), parece que el razonamiento no se desarrolla de esta forma. Diversos autores se encuentran trabajando actualmente en estudiar la forma de desarrollar el razonamiento estadístico (delMas, Garfield y Chance 1999), particularmente en el área de inferencia estadística.

1.4 Aplicación del Conocimiento Estadístico

De acuerdo a Gardner y Hudson (1999), los estudiantes universitarios tienen dificultades para saber cuándo aplicar los conceptos estadísticos que han aprendido en los cursos, al resolver problemas reales de estadística. Gardner y Hudson realizaron un estudio en una comunidad académica australiana para identificar las pruebas más comúnmente utilizadas entre la gente involucrada en el uso de la estadística en la educación; entre los entrevistados se encontraban

profesores de estadística, estudiantes y autoridades académicas. A los participantes se les solicitó que calificaran la importancia de diversos procedimientos estadísticos descriptivos e inferenciales, para incluirse en un curso introductorio de estadística para las ciencias de la conducta, tomando como base una lista inicial de 34 procedimientos estadísticos. Para la evaluación, se utilizó una escala de cinco valores (5=esencial, 4=muy importante, 3= importante, 2=relativamente importante e 1=innecesario). Los resultados mostraron que la estadística descriptiva obtuvo el mayor puntaje, seguida de los procedimientos más comunes en la estadística inferencial (prueba t para muestras independientes, correlación de Pearson, prueba t para muestras relacionadas, ANOVA one-way para grupos independientes, prueba de asociación Ji Cuadrada, regresión simple, ANOVA factorial, prueba t para una muestra, etcétera). Se obtuvieron bajos puntajes para estadísticas escasamente utilizadas (correlación semiparcial), análisis más complejos, como los análisis multivariados (MANOVA, path análisis, cluster análisis, análisis discriminante) y, se obtuvieron puntajes moderados para los procedimientos no paramétricos: Kruskal-Wallis, Friedman, U de Mann-Whitney, Wilcoxon, prueba de signos, contrario a lo que podría pensarse, en virtud de que se trata de métodos estadísticos que se incluyen en los cursos introductorios de estadística a nivel universitario.

En el mismo estudio se analizó la capacidad de aplicación del conocimiento adquirido en la solución de 21 problemas estadísticos escritos aplicados a estudiantes universitarios. La tarea consistió en determinar cuál prueba estadística se debería aplicar a los problemas estadísticos planteados. Los resultados confirmaron la enorme dificultad de la tarea. La probabilidad obtenida de que cada estudiante identificara el procedimiento estadístico apropiado, fue de (.25). Gardner y Hudson (1999) sugieren que se debería dar a los estudiantes más tiempo para practicar las habilidades adquiridas, lo cuál podría significar la reducción de los tópicos cubiertos en los cursos de estadística. El análisis cualitativo reveló que existen diferentes patrones de dificultades, relacionados con:

- a) Interpretación errónea de la pregunta de investigación
- b) Incapacidad para nombrar el procedimiento que estadístico
- c) Dificultades para reconocer el nivel de medición de los datos
- d) Respuesta inapropiada para identificar las claves verbales de los problemas
- e) Dificultades con la interpretación de diferentes formas de presentación de los datos

2. La Investigación sobre la Solución de Problemas y los Estudios Experto-Novato

2.1 La Solución de Problemas

La solución de problemas ha sido estudiada principalmente por psicólogos y recientemente por investigadores en inteligencia artificial (Bransford & Stein, 1993); (Simon, 1986) ; (Wilson, Fernández, & Hadaway, 1993) . El estudio de la solución de problemas en humanos se ha realizado generalmente en laboratorios, usando problemas que pueden ser resueltos en periodos relativamente cortos y frecuentemente se busca acumular una gran cantidad de datos acerca del proceso de solución, (Lajoie & Vivet, 1999). Aunque virtualmente cualquier actividad humana puede ser vista como la solución de un problema, a través de la historia del estudio de la solución de problemas, mucha de la investigación realizada, ha tenido que ver con tareas que toman minutos u horas para ejecutarse. Típicamente los sujetos llevan a cabo muchas acciones observables durante este periodo, y estas acciones se interpretan como la parte visible externamente del proceso de solución (Van Lehn, 1996). Sin embargo existen muchos procesos cognitivos que no son visibles a simple vista y requieren de procedimientos complejos para extraer los procesos implicados. Antes de entrar en la discusión sobre la solución de problemas, se requiere comprender la unidad básica de análisis de este campo de estudio, es decir el *problema*.

Desde el punto de vista del procesamiento de la información, un problema tiene tres elementos básicos: Un *estado inicial*, un *estado final* y un conjunto de *procesos* (operadores), que pueden transformar un estado en otro (Garnham & Oakhill, 1996). El *estado inicial* es el estado del "mundo" que plantea el problema, este mundo puede ser real o ficticio como en el caso de los juegos, los rompecabezas y otros juegos. El *estado final* es un estado del mundo en que el problema está solucionado. Los *procesos* son cosas que se le hacen al mundo para cambiarlo del *estado inicial* al *estado final*; en el caso del ajedrez, serían los movimientos permitidos por las reglas del juego.

En las últimas tres décadas, la mayoría de los problemas se han enfocado en problemas multipasos (Van Lehn, 1996), lo que quiere decir que para encontrar la solución no basta con ejecutar un solo operador, sino que se deben llevar a cabo un número considerable de pasos de manera correcta. Un ejemplo de tales tareas es resolver una ecuación algebraica. La solución es una secuencia de transformaciones algebraicas aplicadas correctamente. La dificultad en el problema radica en la decisión de qué transformaciones se deben aplicar, recordarlas en forma precisa y aplicarlas correctamente. De esta forma, la responsabilidad en la solución se distribuye sobre el proceso completo de solución y no solamente en el descubrimiento de uno o dos pasos claves. La elección de tareas a estudiar, originó que la investigación se enfocara principalmente en la manera en que la gente organiza el proceso de solución, en cómo deciden los pasos que se deben aplicar, en qué circunstancias se llevan

cabo, y cómo su conocimiento del área, determina su visión del problema y el descubrimiento de la solución. Finalmente, los pasos previamente separados en la ejecución de la tarea, se compilan en procedimientos cada vez más grandes. Conforme se van practicando extensivamente, pueden alcanzar un estado de *automaticidad* en el cual pueden ser ejecutadas con poco conocimiento consciente.

Las investigaciones en el laboratorio acerca del proceso de solución de problemas, se han complementado con estudios de campo, en los cuales se analiza a los profesionales durante la solución de problemas del mundo real, por ejemplo, médicos haciendo diagnósticos ó expertos en ajedrez analizando las posiciones del juego, en otros casos se han investigado ejecutivos de negocios tomando decisiones de inversión. Actualmente, se están usando registros históricos, incluyendo cuadernos de laboratorio de los científicos para estudiar procesos de solución de problemas en el descubrimiento científico (Simon, 1986).

Para aproximarse al estudio de la solución de problemas existen diversos métodos y técnicas que intentan extraer información, generalmente de expertos, solucionando problemas. La técnica más utilizada para estudiar la solución de problemas son "Los protocolos de pensamiento en voz alta", que consisten en solicitar a los sujetos que piensen en voz alta mientras trabajan. Esta técnica fue vista al principio con suspicacia por los conductistas, definiéndola como subjetiva e "introspectiva", sin embargo, recientemente ha recibido una atención metodológica especial. Hoy en día la técnica de "Pensar en voz alta" se usa formalmente para obtener datos acerca de la conducta de los sujetos en un rango amplio de escenarios. De esta manera, se les pide a los sujetos que resuelvan problemas en sus mentes, se les solicita que piensen en voz alta mientras trabajan, y el protocolo verbal resultante se interpreta como una secuencia de acciones. Estos protocolos son una aproximación a los procesos que se encuentran entre el estado inicial y el estado final del problema (Simon, 1986).

De los estudios empíricos, surge una descripción de los procesos de solución de problemas que se considera válida para un amplio rango de actividades. La solución de problemas generalmente se conduce por una búsqueda selectiva a través de un conjunto grande de posibilidades, usando atajos (heurísticos) para guiar la búsqueda. En vista de que las posibilidades en las situaciones de los problemas realistas son generalmente muchísimas, la búsqueda por ensayo y error simplemente no funciona; la búsqueda debe ser altamente selectiva.

Un procedimiento usado a menudo para guiar la búsqueda, es el denominado "*escalar la montaña*", utilizando alguna medida de aproximación a la meta para determinar posteriormente, dónde es más provechoso buscar la solución. Otro procedimiento más poderoso y común es el análisis *medios-fines*. En el análisis *medios-fines*, el solucionador del problema compara la situación actual con la

meta, detecta una diferencia entre ambos, y luego busca en su memoria las acciones que es más probable que reduzcan la diferencia. Si la diferencia es cincuenta millas hacia la meta, el solucionador del problema buscará en la memoria el conocimiento acerca de autos, motos, bicicletas y otros medios de transporte, las acciones de caminar y volar serán probablemente descartadas como inapropiados por la distancia.

Otra cuestión que se ha aprendido acerca de la solución de problemas, especialmente cuando el solucionador es un experto, es que se requiere el manejo de grandes cantidades de información que están almacenadas en la memoria y que son recuperables cuando el solucionador reconoce claves señalando su relevancia. De esta manera, el conocimiento experto de un médico se evoca por los síntomas presentados por el paciente; este conocimiento conduce a la recolección de información adicional que se necesita para discriminar entre enfermedades alternativas y, finalmente para establecer el diagnóstico.

En pocos casos ha sido posible estimar cuántos patrones debe reconocer un experto para acceder al conocimiento relevante almacenado en la memoria (Simon, 1986). De las actividades más estudiadas es el juego de ajedrez. Un master en ajedrez debe ser capaz de reconocer alrededor de 50,000 configuraciones diferentes de piezas de ajedrez que ocurren frecuentemente en el curso de los juegos de ajedrez. Un médico debe ser capaz de reconocer decenas de miles de configuraciones de síntomas; un botánico o zoólogo especialista en taxonomía, cientos de miles de características de especímenes que definen sus especies. Por comparación, los graduados del colegio típicamente tienen un vocabulario en sus lenguas nativas de 50,000 a 200,000 palabras; (sin embargo, estos números son muy pequeños en comparación con las situaciones del mundo real que los expertos enfrentan: tal vez hay un 10 a la 120 potencia de ramas en el árbol del juego de ajedrez, jugado solamente con seis tipos de piezas en un tablero de 8 por 8).

Uno de los logros de la teoría contemporánea de la solución de problemas ha sido proporcionar una explicación de los fenómenos de la *intuición* y el *juicio*, observados frecuentemente en la conducta de los expertos (Norman, 1985). El almacenamiento del conocimiento experto "indexado" por las claves de reconocimiento que lo hacen accesible y combinado con algunas capacidades inferenciales básicas (tal vez en la forma de análisis *medios-fines*) da cuenta de la habilidad de los expertos para encontrar soluciones satisfactorias para los problemas difíciles, y algunas veces para encontrarlas casi instantáneamente. La *intuición* y el *juicio* de los expertos se derivan de esta capacidad de reconocimiento rápido, vinculada a un gran almacenamiento de conocimiento. Cuando la intuición inmediata falla en producir una solución del problema o cuando una solución probable necesita ser evaluada, el experto retrocede a los procesos inferiores de análisis e inferencia, que son característicos de los sujetos novatos. A continuación se describirá con mayor amplitud en qué

consisten los estudios experto-novato.

2.2 Los Estudios Experto-Novato

En las últimas dos décadas se han realizado numerosas investigaciones sobre la diferencias entre sujetos expertos y novatos en la forma de resolver problemas y ejecutar tareas específicas (Chi et al., 1988). En la compleja sociedad actual estamos siempre en manos de expertos, personas críticas en un área o un conjunto de tareas especializadas. Desde pilotos de avión al abogado laborista, pasando por el mecánico de coches, el oftalmólogo, el psicólogo clínico, muchas de las profesiones que se ejercen en nuestra sociedad requieren una larga y costosa formación especializada, miles de obras de práctica distribuidas a lo largo de unos cuantos años (Norman, 1985). Al respecto, (Pozo, 1993), se plantea los siguientes cuestionamientos: ¿Qué diferencias hay entre esa persona experta y todos los que somos novatos en esa materia?, ¿En qué consiste ser experto? ¿Cuáles son los procesos psicológicos responsables de la transición al conocimiento experto? Preguntas que han sido materia de interés de diversos investigadores de la ciencia cognitiva (Norman, 1985).

La extensa literatura en psicología cognitiva que contrasta la ejecución experta con la novata en una amplia variedad de dominios (Chi et al., 1988) tiene implicaciones para el análisis de las tareas de una actividad. La investigación demuestra claramente la importancia del conocimiento en un campo específico en la ejecución experta, esto tiene implicaciones en el diseño del entrenamiento efectivo. La aproximación experto-novato va en contra de la práctica prevaleciente en el desarrollo de los cursos de entrenamiento, la cual asume, que el diseñador instruccional puede aplicar habilidades de diseño generalizadas independientemente del contenido y se ignora la importancia de la instrucción en un contexto situado, en el que usualmente se desarrolla la tarea. De acuerdo con Chi et al. (1988), si realmente queremos entrenar gente hacia la pericia, debemos comprender que se requiere una gran cantidad de conocimiento acerca del área que se trate. Esto sugiere que debemos involucrarnos en el diseño del currículo instruccional y pasar una cantidad importante de tiempo aprendiendo sobre el área. De la misma forma, el análisis de tareas en estos campos deberá ser un esfuerzo colaborativo y minucioso, en el cual los expertos en el campo constituyen un elemento esencial.

2.2.1 Diferencias en el conocimiento entre expertos y novatos

El surgimiento de esta área de estudio para la psicología se ha debido no sólo las necesidades teóricas de la psicología cognitiva imperante, a partir de la recuperación del aprendizaje por el procesamiento de información, sino muy especialmente a las necesidades tecnológicas derivadas del diseño de *sistemas informáticos expertos* en la solución de problemas específicos (Pozo, 1993). Un *sistema experto* es un programa que tiene una amplia base de conocimientos en

un dominio restringido y usa un razonamiento diferencial complejo para ejecutar tareas que haría una persona experta. Los sistemas expertos se basan en la emulación de ciertos rasgos del conocimiento experto. Ello ha impulsado sin duda el interés por conocer más sobre la naturaleza psicológica de la pericia¹. Y en fechas recientes ha impulsado también el estudio de los procesos mediante los que evolucionan los conocimientos del experto. Dado lo complejo que resulta construir un sistema experto, es muy rentable conocer cómo podría aprender, ya que eso posibilita el diseño de sistemas expertos más flexibles, capaces de adaptarse a nuevas situaciones.

Dado que el conocimiento en los expertos se caracteriza básicamente por tener "bases de conocimiento específicas", que dependen del área temática de que se trate, resulta muy difícil realizar una comparación entre los diversos campos de dominio. Entre los trabajos de investigación más citados y numerosos en la literatura científica, se encuentran los que se refieren a la mecánica Newtoniana, estos estudios ofrecen un cuadro más completo de las diferencias cuantitativas y cualitativas entre expertos y novatos.

De acuerdo a Pozo (1993), todos los estudios comparativos entre expertos y novatos parten de unos presupuestos comunes, independientemente de su especificidad temática, estos supuestos pueden ser explícitos o implícitos:

- a) La diferencia experto/novato es básicamente una diferencia de conocimientos y no de procesos cognitivos básicos o capacidades generales de procesamiento.
- b) Esa diferencia de conocimiento es tanto cuantitativa como cualitativa; esto es, los expertos no sólo saben más que los novatos, sino que sobre todo tienen organizados sus conocimientos de una forma distinta.
- c) La pericia es un efecto de la práctica acumulada, esto es, un efecto del aprendizaje, desdeñándose, por tanto, los factores innatos y las posibles diferencias individuales.
- d) La pericia está circunscrita a áreas específicas de conocimiento, de forma que se es experto o no con respecto a algo. Un mismo sujeto puede tener grados diversos de pericia para problemas conexos de una misma área.

En cuanto a la metodología empleada en estos trabajos, existen también numerosas variantes. No obstante, suelen basarse en la solución de un problema por un grupo de expertos y otro de novatos. Los problemas planteados pueden ser de naturaleza más cualitativa (toma de decisiones en espacios de problemas más o menos restringidos, organización jerárquica de los conceptos relevantes para la solución del problema, etc.) o simplemente cuantitativa - solución de ecuaciones o cálculos matemáticos. Asimismo, la información se

¹ En inglés: Expertise

extrae mediante diversas técnicas o métodos (protocolos de pensamiento en voz alta, categorización de problemas, recuerdo libre, Análisis Cognitivo de Tareas, etc.). Los estudios experto-novato se han enfocado principalmente en encontrar tanto diferencias cuantitativas como cualitativas entre los expertos y los novatos en diversos ámbitos del conocimiento (Jonassen, Primavera, 2003).

2.2.2 Diferencias cuantitativas entre expertos y novatos

Herbert Simon (1978), fue uno de los pioneros de la psicología cognitiva del procesamiento de información. El método empleado por él y muchos otros investigadores es el siguiente: se presenta a un grupo de sujetos expertos y otro de novatos (cuyo número oscila entre uno y once sujetos) una serie de problemas de cinemática, consistentes en la resolución de varias ecuaciones sobre movimientos rectilíneos. Por lo general se presentan problemas fáciles, que pueden resolverse en uno o dos minutos. A los sujetos se les pide que "piensen en voz alta" mientras resuelven el problema. Finalmente, con los datos recogidos se realiza un sistema de producción para cada sujeto, que posteriormente se traduce a un programa de computadora, que simula las ejecuciones de expertos y novatos.

En el proceso de solución de esas ecuaciones por los expertos y los novatos surgen diferencias de naturaleza cuantitativa. Así, los novatos cometen más errores y tardan, en término medio, cuatro veces más que los expertos en resolver los problemas. Estas diferencias son consecuencia de las distintas estrategias usadas por uno y otro tipo de sujetos. Pozo (1993) describe estas diferencias de la siguiente manera:

- a) Los expertos, antes de empezar a aplicar ecuaciones, se detienen en una fase de representación, en la que realizan un análisis cualitativo, basado en una intuición física. En cambio, el novato inicia inmediatamente las ecuaciones.
- b) Ambos grupos difieren en la estrategia seguida para la solución del problema. Los novatos hacen un análisis hacia atrás, esto es, parten de la meta final (solución desconocida del problema) y luego buscan los datos disponibles que pueden ser útiles para alcanzar. Por el contrario, los expertos trabajan hacia delante, utilizando los datos conocidos para buscar la solución desconocida. Esta diferencia desaparece significativamente cuando el problema se complica, ya que entonces también los expertos trabajan "hacia atrás". Esta estrategia consiste en un análisis medios-fines, similar al que usan los sujetos en la solución de problemas carentes de carga semántica, como la *Torre de Hanoi* y, en los que no pueden aplicar conocimientos específicos previos.
- c) Al aplicar las ecuaciones, el novato las hace de una en una, mientras el experto las realiza agrupadas, es decir calcula varias ecuaciones juntas,

después hace una pausa y a continuación realiza otro grupo ecuaciones. Se considera que estos grupos de ecuaciones corresponden a "Chunks" de conocimientos, análogos a los empleados por los jugadores expertos de ajedrez en el recuerdo de las posiciones, y se revelarían la existencia de "racimos de conocimiento" organizados en la mente del experto, frente conocimiento más aislado del novato.

- d) Por último, los novatos realizan muchos más meta-enunciados con respecto al propio proceso de solución. Esto es, dedican más tiempo a pensar sobre la estrategia que deben seguir y a analizar y planificar los pasos siguientes. Por el contrario los expertos apenas "piensen en voz alta" sobre las operaciones que deben hacer para solucionar las ecuaciones. Se diría que los expertos no necesitan tomar decisiones sobre la forma en que deben resolver el problema, mientras que los novatos deben dedicar una buena parte de su tiempo precisamente a hallar la vía de solución adecuada.

Todas las diferencias anteriores, en especial las tres últimas, han sido interpretadas como una muestra de que la actuación de un experto consiste básicamente en un proceso de automatización de sus conocimientos, de tal forma que no necesita ir tomando decisiones a medida que resuelve problemas, por lo que parte directamente de los datos, sin hacer, como novato, acercamientos progresivos a la meta final. Obviamente la solución del experto resulta así no sólo eficaz, sino sobre todo más eficiente. Consume menos tiempo, se equivoca menos y realiza "racimos de ecuaciones" que ha automatizado en una misma secuencia. Todas estas ventajas se deberían a un proceso de automatización, como el postulado por Anderson (1983) dentro de su teoría ACT. El experto se caracterizaría por haber compilado o agrupado su conocimiento en secuencias de acción automáticas que no precisan de esfuerzo atencional para su ejecución. Ello no sólo hace más rápida y eficiente la ejecución, sino que además libera espacio de procesamiento para dedicarlo a los rasgos novedosos de la situación o controlar el propio procesamiento. El proceso de convertirse en experto -y por lo tanto, la naturaleza psicológica de la propia pericia, sería similar a lo que sucede cuando una persona está aprendiendo a conducir (Anderson, 1982). Las acciones que inicialmente, ejecuta el novato, de una en una y con un costo atencional alto, ocurren como consecuencia de la práctica repetida y acompañada generalmente de retroalimentación, a ejecutarse en el experto en una misma secuencia automatizada, compilada y procedimentalizada.

Pero esa diferencias, existentes sin duda entre los expertos y los novatos en la vida cotidiana (piénsese en el llamado "ojo clínico" o diagnóstico casi automático de enfermedades por parte de los médicos expertos en comparación con el más concienzudo pero titubeante diagnóstico de un médico novato), parecen depender también de la naturaleza del problema presentado. Aunque es cierto que los expertos resuelven por procesos automáticos problemas familiares

para ellos –que en cambio requieren procesos controlados en los novatos-, esta diferencia puede diluirse si enfrentamos a los expertos a problemas nuevos. De hecho, podemos preguntarnos si la solución de ecuaciones de mecánica -como en el diagnóstico de enfermedades cotidianas- los expertos están resolviendo realmente un problema o simplemente realizando un "ejercicio". La simple aplicación de estrategias automatizadas o compiladas no se ajusta a lo que normalmente se entiende por solucionar un problema.

Chi, Glaser y Rees (1982) replicaron los trabajos anteriores, pero con problemas aparentemente más difíciles. Sus resultados son diametralmente opuestos: no existía ninguna diferencia en el tiempo de solución (si acaso los novatos tardaban menos), ni tampoco había diferencias en las pausas entre ecuaciones. En cuanto al análisis cualitativo previo que según Larkin (1980) es característico de los expertos, también se producía en los novatos. La única diferencia importante entre expertos y novatos consistía en el mayor número de errores que estos cometían, como consecuencia tanto de errores de cálculo como, especialmente de la aplicación al problema de conocimientos inadecuados o insuficientes.

Por tanto, al complicar la tarea desaparecen los efectos de la automatización del conocimiento experto, pero empiezan a aparecer diferencias cualitativas más profundas en la forma en que expertos y novatos organizan su conocimiento. El tipo de problemas y el método utilizado en las investigaciones reseñadas hasta ahora, no permite profundizar en el análisis de esa diferencia de conocimientos entre expertos y novatos. Según Chi, Glaser y Rees (1982), el análisis de las verbalizaciones de los sujetos durante la solución del problema no basta para desentrañar la estructura de sus conocimientos. Además, los problemas cuantitativos pueden resolverse en muchos casos sin aplicar análisis conceptuales exhaustivos. Por ello es necesario recurrir a una nueva metodología para identificar las diferencias cualitativas entre el conocimiento novato y experto en un área concreta.

2.2.3 Diferencias cualitativas entre expertos y novatos

En la misma investigación, Chi, Glaser y Rees (1982), presentan una extensa investigación en la que analizan las diferencias en la estructura cualitativa de los conocimientos de expertos y novatos en la misma área de mecánica, utilizando diversas técnicas de toma de datos, tales como categorización y clasificación jerárquica de problemas, resúmenes de textos e identificación de los rasgos relevantes del problema. En cada tarea comparan a un grupo de expertos y otro de novatos (que varían de dos a ocho sujetos), a los que en algunos casos incorporan un tercer grupo de sujetos intermedios. Cada tarea consiste en realizar una determinada actividad cualitativa, como las anteriormente señaladas. Con un número variable de problemas generalmente extraídos de un manual introductorio de Física. El análisis de los datos no se basa en protocolos verbales, sino en una disección cualitativa de las respuestas de los sujetos a

partir del marco conceptual proporcionado por la mecánica newtoniana.

Una de las tareas que Chi y colaboradores planteaban a sus sujetos, consistía en clasificar una serie de problemas sobre mecánica. Los expertos y novatos diferían en los criterios o atributos en que basaban su clasificación. Los expertos categorizaban los problemas según su estructura conceptual profunda, determinada por las leyes y conceptos físicos relevantes para su solución, mientras que los novatos guiaban sus clasificaciones por la estructura superficial, basada en objetos reales y términos explícitos en el enunciado de los problemas. Los conceptos activados por expertos y novatos -los esquemas desde los que asimilan la realidad- son en general diferentes.

Estas diferencias resultaron aún más patentes en otro de los estudios de la misma investigación. Los autores diseñaron una serie de 20 problemas, algunos de los cuales tenían una considerable semejanza superficial, ya que implicaban los mismos materiales reales (por ejemplo, poleas o rampas) pero diferían en los conceptos implicados en su solución. En cambio otros problemas, utilizando distintos materiales, compartían principios explicativos comunes. Como cabía esperar, los novatos tendieron a clasificar juntos los problemas superficialmente similares pero estructuralmente distintos, mientras que los expertos hacían justo lo contrario. Se puede expresar que entre sujetos expertos y novatos: mientras que los expertos se guían por abstracciones conceptuales, las representaciones de los novatos consisten en objetos reales organizados según parámetros de tiempo real. Esta tendencia de los novatos a comprender los problemas a partir de sus rasgos superficiales puede llevarles a cometer serios errores conceptuales.

Esta misma investigación sirve para ilustrar otra de las diferencias más relevantes observadas por Chi y colaboradores entre los expertos y los novatos. Ambos difieren en sus conocimientos declarativos básicos, esto es, en el contenido de los esquemas o conceptos que activan. Aún cuando utilicen las mismas etiquetas verbales, lo que no es muy frecuente, el significado de las mismas varía. Si se descomponen las leyes de Newton en sus elementos conceptuales básicos, los esquemas o los modelos mentales expertos contienen más conceptos físicos básicos que los novatos. Esto es algo que no nos debe sorprender, sin embargo, lo que llama la atención es la escasa comprensión que tienen los novatos de los conceptos físicos básicos. Los sujetos novatos cometían numerosos errores conceptuales, a pesar de haber sido instruidos en física durante varios años, no llegaban a comprender conceptos como la *inercia* o el *movimiento parabólico* en la caída libre de la bola. Estos errores conceptuales se deberían a la existencia en los novatos de fuertes *concepciones espontáneas* o *teorías implícitas* sobre el movimiento de los objetos, que han dado lugar a una gran cantidad de investigaciones, un verdadero género literario dedicado a la comprensión intuitiva de la ciencia por parte de los novatos. Según esos estudios, el paso de novato a experto no implicaría únicamente un cambio en el contenido declarativo de los esquemas o conceptos, sino que requeriría un

verdadero cambio conceptual, una auténtica revolución conceptual similar a las ocurridas en el proceso del conocimiento científico a lo largo de la historia (Bransford, Brown, & Cocking, 2000) .

Una de las consecuencias de entender así el paso de novato a experto es que debe implicar cambios estructurales y no sólo de contenidos. La investigación de Chi, Glaser y Rees (1982) avala la existencia de diferencias estructurales entre conocimiento experto y novato. Además de diferir en su contenido, los esquemas o conceptos activados por unos y otros difieren también en su organización. Así, tras una primera clasificación de un conjunto de 40 problemas se pidió a ambos grupos de sujetos que hicieran nuevas categorizaciones. Se encontraron notables e imprevistas diferencias individuales en ambos grupos que, por lo demás, no son exhaustivas de esta investigación. No obstante, se observaron ciertos rasgos característicos de las organizaciones conceptuales respectivas de novatos expertos.

En primer lugar, puede observarse que los novatos tienen un mayor número de categorías básicas. Además, muchas de esas categorías incluyen un solo problema por lo que no son susceptibles de nuevas discriminaciones. Por último, el contenido de las categorías de los novatos suele basarse en rasgos superficiales u objetos reales de la situación. Estas clasificaciones novatas contrastan con las que hacen los expertos: hay menos categorías básicas, pero más discriminaciones e interacciones; cada categoría es más comprensiva y está relacionada con otras categorías. De hecho, las clasificaciones jerárquicas de los expertos se corresponden con las características que tienen las estructuras conceptuales de los expertos: "en la red conceptual almacenada por el experto hay múltiples rutas de cada uno de los conceptos a los otros conceptos; podemos decir que cada concepto tiene en el diccionario mental del experto múltiples referencias cruzadas. Esta mayor densidad de conexiones entre conceptos en la red conceptual del experto significa a su vez que la probabilidad de que cualquier concepto dado evoque otros conceptos relacionados es mayor". Esta mayor interconexión en los conceptos del experto se refleja también explicaciones causales: los novatos tienden a dar explicaciones basadas en una sola causa o en varias yuxtapuestas, mientras que los expertos integran diversas causas en un sistema común, analizando las interacciones entre ellas.

Por tanto, los novatos y los expertos difieren no sólo en sus conocimientos específicos en un área sino también en la forma en que organizan esos conocimientos (Bransford et al., 2000). De hecho, ambos aspectos están íntimamente relacionados. La estructura de una teoría o modelo explicativo está muy relacionada con su contenido, de tal forma que puede afirmarse que el progreso en el conocimiento se debe tanto un incremento cuantitativo "del saber cuánto" a una reorganización cualitativa de las ideas que componen la teoría. De hecho, el contenido y la estructura de las teorías mantenidas por expertos y novatos en tareas de mecánica, correlacionan positivamente de modo significativo. La relación entre el proceso estructural y el avance de los

conocimientos específicos se refleja en otro de los estudios realizados por Chi, Feltovich y Glaser (1981). Cuando analizaron la representación jerárquica que tenían sus sujetos novatos y expertos de un problema de plano inclinado, se encontraron con diferencias muy informativas. Aunque muchos conceptos coinciden en ambas representaciones, su lugar dentro de la estructura es diferente. En la parte superior de la jerarquía del novato se encuentran objetos reales o categorías superficiales, a los cuales se hallan subordinados algunos principios físicos. En cambio, la representación del experto está presidida por leyes y principios físicos, de forma contraria a los novatos, los conceptos están subordinados a los principios. Según los autores, los expertos poseen conceptos de nivel superior que los que presentan los novatos, que integran los objetos del mundo real que constituyen las categorías básicas del novato. La principal carencia conceptual de los novatos reside precisamente en el poco uso que hacen de los principios generales de una disciplina específica para resolver los problemas.

Además de esta organización jerárquica, por la que los conocimientos de los novatos quedan subsumidos y, por tanto, modificados, en esquemas de nivel superior, los esquemas de los expertos tienen también mayor contenido procedimental, ya que son en sí mismos, procedimientos idóneos para la solución efectiva de problemas, cosa que no sucede en el caso de los novatos. En otros términos, los modelos conceptuales expertos pueden concebirse como sistema de producción compuestos por unidades condición-acción. En cambio, los modelos conceptuales de los novatos revelan que carecen de información sobre cuándo deben activar su esquema. La representación del experto específica las condiciones de aplicación de las leyes y principios generales, mientras que la del novato no contiene esas especificaciones.

Una última diferencia digna de mención entre los expertos y los novatos reside en el proceso selectivo que realizan ante problemas mecánicos. Cuando se les pide que seleccionen los rasgos más importantes para la solución del problema, los novatos a su vez, se refieren a objetos reales o términos mencionados en el problema. Muy raramente mencionan por sí mismos rasgos de segundo orden, esto es, no perceptible directamente. En cambio, los rasgos destacados por los expertos pueden caracterizarse como descripciones de los estados y condiciones de la situación física descrita por el problema. *En otras palabras, enfrentados al mismo problema, un experto y un novato no ven el mismo problema, ya que sus esquemas de asimilación son diferentes, una prueba más de la naturaleza constructiva del conocimiento.*

En suma, el paso de novato a experto, implica no sólo cambios cuantitativos en la cantidad de conocimientos como consecuencia de la práctica, sino también una verdadera reorganización de esos conocimientos. Cuando una persona se convierte en un experto, no sólo automatiza sus destrezas sino que además reestructura sus conocimientos. Sin embargo, quedan algunas preguntas sin contestar (Pozo, 1993), por ejemplo: ¿Cómo se produce esa transición?,

¿Cuáles son las condiciones que favorecen la reestructuración de conocimientos que llamamos pericia?, ¿Qué conexión genética hay entre automatizar y reestructurar? Desgraciadamente, no hay todavía respuestas firmes a algunas de estas preguntas. Hasta la fecha, los estudios comparativos entre novatos y expertos han tendido a ser demasiado estáticos, descriptivos, tal vez por las dificultades metodológicas que implica estudiar la transición de novato a experto. No obstante, la comparación experto-novato está siendo propuesta por algunos autores como un modelo adecuado para analizar el desarrollo cognitivo. Hay incluso quien afirma que los niños no son sino novatos universales y que lo que otros autores consideran diferentes estadios de desarrollo, con sus correspondientes estructuras intelectuales distintas, no son sino niveles diferentes de pericia en dominios específicos, debido a que los niños tienen menos conocimientos y están menos organizados. Esta posibilidad incrementa el interés por analizar los procesos de reestructuración en la conversión de una persona novata en un experto.

Otro hallazgo de la literatura experto-novato es la importancia de la manera en que el problema se representa mentalmente. Para analizar las representaciones mentales sobre la solución de problemas, se ha utilizado una metodología para extraer el conocimiento experto y novato, surgida de la Psicología Cognitiva conocida como *Análisis Cognitivo de Tareas*, la cual agrupa una serie de técnicas específicas que varían de acuerdo a las características propias del campo a estudiar. En el siguiente apartado se abordan detalladamente las técnicas mayormente utilizadas.

3. La Aproximación del Análisis Cognitivo de Tareas al Diseño Instruccional

El diseño de un ambiente educativo basado en la computadora para el entrenamiento del conocimiento experto en una tarea compleja como es la aplicación e interpretación de análisis estadísticos, requiere la aplicación de técnicas utilizadas en Inteligencia Artificial (IA) para el diseño de Sistemas Tutoriales Inteligentes (STI), tales como la representación del conocimiento y su inferencia en la educación basada en computadora (Lajoie & Vivet, 1999).

La fundamentación para construir ambientes instruccionales complejos apoyados en el uso de la computadora tales como los STI, se basa en la afirmación de que es posible modelar, trazar y corregir los procesos cognitivos en el contexto de la solución de problemas (Lesagold, Lajoie, Logan, & Eggan, 1990). El desarrollo de un STI se sustenta en la aplicación del Análisis Cognitivo de Tareas.

El propósito central del Análisis Cognitivo de Tareas es evaluar las diferencias cuantitativas y cualitativas de las descripciones, conceptos, principios y procedimientos, mientras los sujetos (generalmente expertos) se involucran en la ejecución de una tarea específica.

El Análisis Cognitivo de Tareas que se puede emplear para diseñar un Tutor Inteligente, el cual es un proceso de múltiples fases, en el que los investigadores deben comprender, en primera instancia, los aspectos que representan mayor dificultad del campo en cuestión, antes de construir tareas que establezcan la diferenciación entre una ejecución experta y otra menos experta en dicho dominio (Lajoie et al., 1997). En el proceso se extrae y sistematiza la siguiente información.

1) *Tareas que constituyen la meta de la ejecución de la actividad:*

Las tareas que representan los segmentos más importantes y de mayor dificultad.

2) *Habilidades involucradas en la ejecución de las tareas meta:*

La manera en que se resuelven los problemas, cuando son llevados a cabo adecuadamente, y qué tipo de habilidades procedimentales y conocimientos del sistema están involucrados.

3) *Estrategias heurísticas y modelos mentales utilizados por los expertos:*

Cuáles son las estrategias generales que utilizan los expertos a lo largo de diferentes tipos de problemas, y cuál es la naturaleza y el contenido de los planes para la solución de problemas utilizados por los expertos.

4) *Las dificultades encontradas por los novatos:*

Qué es lo que hacen los novatos ante este tipo de problemas. De qué tipo de conocimiento, habilidades y estrategias carecen, si muestran algún tipo de concepciones erróneas, y si son capaces de emplear los procedimientos en una variedad de circunstancias.

5) *Dónde concentrar la enseñanza:*

Cuáles habilidades, conocimientos o estrategias empleadas por los expertos se adquieren regularmente a través de entrenamiento. Donde se localizan los problemas que no logra resolver la instrucción tradicional o que son resistentes a la instrucción.

El Análisis Cognitivo de Tareas que se plantea emplear en este estudio pretende seguir los pasos anteriores como base para *describir y analizar los componentes cognitivos que subyacen al proceso de aplicación adecuada de la prueba de hipótesis en la solución de problemas estadísticos en Ciencias Sociales, situados en contextos realistas de enseñanza. El producto de dicho análisis dará como resultado la elaboración de un modelo cognitivo de la pericia en*

estadística, el cual servirá como base para diseñar en una fase posterior, un ambiente educativo basado en computadora. Para analizar los componentes de una tarea, tradicionalmente se ha utilizado el "Análisis de Tareas"; sin embargo, las características puramente conductuales de esta metodología, no permiten extraer los componentes cognitivos de las tareas complejas, por lo que ha surgido una metodología que retoma muchas herramientas del Análisis de Tareas tradicional, pero incorpora el componente cognitivo. Para comprender con claridad las diferencias de cada metodología de análisis se plantean de manera específica las características que las distinguen.

3.1 Métodos de Análisis Cognitivo de Tareas

3.1.1 Diferencias entre el Análisis de Tareas Tradicional y el Análisis Cognitivo de Tareas.

El Análisis de Tareas, como se conoce tradicionalmente, se basa en una aproximación de sistemas para el entrenamiento (generalmente llamada "diseño instruccional de sistemas", o simplemente "diseño instruccional") (Zachary, Ryder, & Hicinbothom, 2000). Este tipo de análisis (Zachary et al., 2000) domina el entrenamiento para el trabajo que se ofrece en el ejército y en la mayor parte de los escenarios corporativos (Means, 1993). Esta aproximación se conceptualiza como una secuencia sistemática de actividades, empezando con la identificación de las tareas en términos de conductas que se requieren entrenar y el análisis de ellas, lo cual proporciona el fundamento para definir los objetivos instruccionales, diseñar pruebas y seleccionar las estrategias instruccionales (ver figura 3).

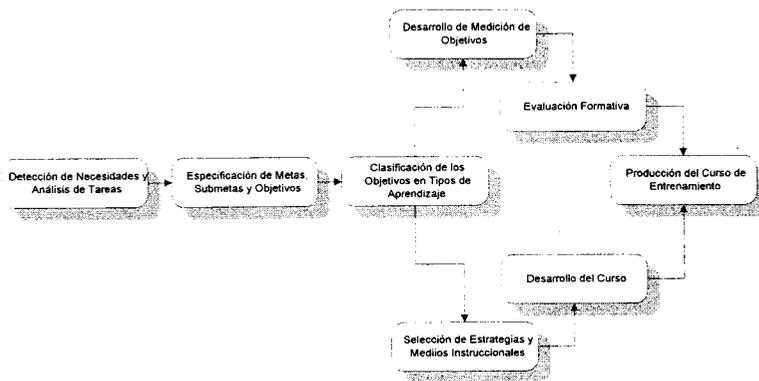


Figura. 3 Etapas en el desarrollo de sistemas instruccionales (Tomado de Means, B., 1993, pág. 98)

En la aproximación conductual al análisis de tareas, la actividad se descompone en una serie de pasos procedimentales. Se considera la tarea ejecutada por expertos, y se enfatizan las acciones concretamente observables. Normalmente, el diseñador instruccional puede especificar los pasos por sí mismo, basándose en la documentación de la tarea disponible, o bien entrevistando al responsable de una tarea. Las preguntas que se suelen utilizar son: ¿Qué es lo que realiza usted primero?, ¿Qué es lo que hace después?, ¿Qué es lo que viene después? De esta forma se especifica un algoritmo para la ejecución de la tarea. En la figura 4 se presenta un ejemplo de la realización de las preguntas. El analista procede a identificar los prerrequisitos para cada paso en la ejecución de la tarea. Puede haber sub-habilidades incorporadas o información acerca de cuándo se debe ejecutar algún paso. El diseñador instruccional identifica los prerrequisitos preguntando: ¿Qué debe saber hacer o conocer el aprendiz cuando se enfrenta a la tarea de aprender esta nueva habilidad? (véase Figura 4).

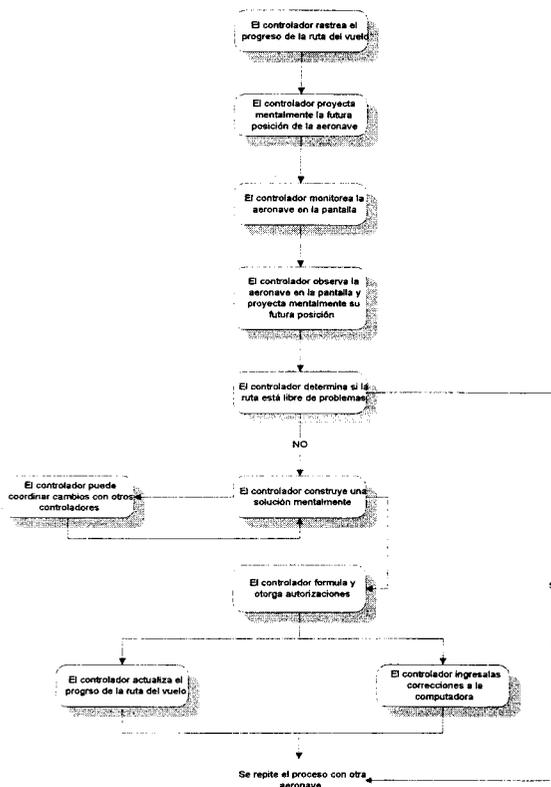


Figura 4. Análisis de Tareas Tradicional del control de tráfico aéreo (Tomado de Means, B., 1993, pag. 100).

La estrategia que utilizan los diseñadores instruccionales para el análisis de tareas tradicional, consiste en entrevistar a los responsables de una actividad o a los supervisores acerca del conocimiento, habilidades y actitudes requeridos para ejecutar diversas tareas (Means, 1993). Dada una lista de tareas y un conjunto de habilidades, elementos de conocimiento y actitudes, los entrevistados indican cuáles son las tareas que son realmente ejecutadas en la actividad y los conocimientos, habilidades y actitudes asociadas a esas tareas. Este tipo de análisis proporciona el fundamento para decidir qué habilidades, conocimientos y actitudes deberían ser entrenados. Sin embargo, esta aproximación tiene una desventaja, los datos recabados no le dicen nada al diseñador instruccional acerca de la secuencia de los pasos en la ejecución de la tarea, por ejemplo ¿cómo se utilizan las habilidades y conocimiento para realizar una tarea en particular?

La aproximación tradicional conductual para el análisis de tareas puede ser bastante eficiente para aquellas tareas que son lo suficientemente determinísticas para que puedan ser especificadas por un algoritmo, especialmente cuando los pasos críticos en la ejecución de una tarea son conductas observables, más que operaciones cognitivas, tales como formulación de hipótesis, juicios o solución de problemas. Para las tareas que no cumplen con estos criterios la aproximación tradicional conductual no es muy apropiada, por ejemplo las tareas que están involucradas en el control de tráfico aéreo. Este es un ejemplo de una tarea con demandas principalmente de tipo cognitivo, la cual involucra identificar: cuándo es necesario intervenir (cuando se debe indicarle a la aeronave que debe modificar su altitud o dirección con objeto de alejarse de otra aeronave), decidir cuál acción es la más apropiada y priorizar los múltiples problemas de este tipo que se deben atender rápidamente, mientras se continúa monitoreando otras situaciones.

En el tipo de análisis tradicional no se describen realmente los procesos críticos, tales como determinar la posición futura de una aeronave, detectar posibles conflictos, generar y evaluar posibles soluciones. Este análisis, por sí mismo no le dice nada al estudiante acerca de cómo ejecutar esas partes vitales de la tarea, y no le dice nada al instructor acerca de cómo enseñarlas. Es precisamente en áreas en que las tareas involucran la solución de problemas y la toma de decisiones en las cuales se encuentran los retos más difíciles para el entrenamiento y es aquí donde la aproximación cognitiva es más apropiada.

Por otra parte, el *Análisis Cognitivo de Tareas* (ACT) tiene sus fundamentos en la extensa literatura en psicología cognitiva, la cual ha influenciado el análisis de las tareas que conforman una actividad determinada al contrastar la ejecución de expertos y novatos en una gran variedad de disciplinas (Means, 1993). La investigación demuestra claramente la importancia del conocimiento en un área específica en la ejecución experta y su implicación en el diseño de entrenamiento efectivo. Si deseamos entrenar gente realmente para lograr la pericia, se requiere un vasto conocimiento acerca del área que estamos

estudiando. Esto implica que las personas involucradas en el diseño instruccional, necesitan dedicar un tiempo considerable aprendiendo sobre el área determinada, y que el análisis de tareas debería constituir un esfuerzo colaborativo en el cual los expertos tienen una participación determinante.

Otra aportación de las investigaciones experto-novato en el *Análisis Cognitivo de Tareas* es la importancia de la forma en que se representan mentalmente los problemas. Los expertos agrupan los problemas que muestran los mismos principios subyacentes, mientras que los novatos los agrupan considerando características superficiales. Cuando se les pide a los novatos que resuelvan problemas, éstos recurren inmediatamente a las ecuaciones, en tanto que los expertos dedican mucho más tiempo tratando de entender cualitativamente el problema, a menudo dibujando diagramas antes de comenzar a escribir una ecuación, lo cual constituye su representación mental del problema que determina las decisiones que realizan los expertos.

La investigación en ciencia cognitiva describe los procesos en los cuales, los pasos separados previamente en la ejecución de la tarea se compilan en procedimientos más grandes. Conforme se va practicando esto extensivamente, se va convirtiendo en procesos automáticos, los cuales pueden ser ejecutados con poca o nula atención consciente. Esto tiene ventajas para la ejecución de la tarea, porque estos procedimientos pueden ser ejecutados sin agotar los recursos mentales; sin embargo, representa una dificultad cuando los expertos son entrevistados para identificar los pasos en la ejecución de la tarea. Frecuentemente tienen dificultad para describir cada paso en la ejecución de la tarea, debido a muchos de ellos se han fusionado.

Una implicación de estas investigaciones es que los expertos no serían capaces de proporcionar a los encargados del entrenamiento todo el conocimiento o los procedimientos que los novatos necesitan aprender. Por lo tanto, se requiere una interacción amplia entre el experto y el diseñador instruccional, en la cual el diseñador requiere del experto la profundización en los detalles cada vez más finos de la tarea (ver Figura 3). A continuación se comparan en detalle las características tanto del *Análisis de Tareas Tradicional* como del *Análisis Cognitivo de Tareas* a fin de visualizar las ventajas de la utilización de éste último en tareas que implican una demanda cognitiva considerable.

En el *Análisis de Tareas Tradicional*, las tareas se evalúan en términos de las respuestas conductuales dadas a cada estímulo durante el proceso de completamiento de una tarea. En contraste, la meta del *Análisis Cognitivo de Tareas* (ACT) es delinear los procesos mentales y habilidades necesarios para ejecutar una tarea a niveles de alta capacidad, así como los cambios en la estructura de conocimiento y procesamiento, conforme las habilidades se desarrollan con el tiempo (Ryder & Redding, 1993).

En tanto que el análisis tradicional se enfoca sobre la ejecución deseada, el ACT

se basa en la pericia: la estructura de conocimiento y las estrategias de procesamiento de la información involucradas en la ejecución. Los métodos tradicionales enfatizan la identificación del conocimiento requerido para cada elemento de la tarea individual. Por el contrario, la aproximación cognitiva enfatiza la base de conocimiento para la tarea completa, su organización e interrelación entre conceptos. Esta aproximación proporciona información útil para estructurar el entrenamiento y facilita tanto el aprendizaje inicial como la progresión hacia la organización del conocimiento utilizado por los expertos o bien por gente que ejecuta la tarea adecuadamente. De la misma forma, desde la aproximación cognitiva, se identifican las habilidades para el trabajo como un todo, no para cada tarea por separado. Otra diferencia consiste en que el análisis cognitivo de tareas considera la especificación de los modelos mentales utilizados en la ejecución de la tarea, lo cual no se considera en los métodos tradicionales (véase Tabla 1).

Tabla 1
Comparación del análisis de tareas tradicional vs el análisis cognitivo de tareas

Análisis de Tareas Tradicional	Análisis Cognitivo de Tareas
Enfatiza la conducta	Enfatiza la cognición
Analiza la ejecución objetivo	Analiza la pericia
Analiza el conocimiento para cada tarea separadamente	Evalúa el conocimiento para la tarea en su conjunto
Segmentación de la tarea basada conductualmente	Segmentación de la tarea basada en habilidades
No se consideran los modelos mentales	Se consideran los modelos mentales

Nota. Tomada de Ryder y Redding, (1993).

La aproximación tradicional no contempla la caracterización de la variabilidad en la ejecución, tanto entre como dentro de los individuos, debido a que un individuo puede ejecutar una tarea en diferentes formas, utilizando diferentes secuencias y métodos.

Por el contrario, los métodos de ACT intentan identificar las estrategias de solución de problemas que pueden manifestarse en secuencias de acción variables, dependiendo de la dinámica de desarrollo de cada tarea, así como las diferencias individuales relevantes.

El *Análisis Cognitivo de Tareas* se fundamenta en investigaciones de campo y de laboratorio en ciencia cognitiva para obtener métodos y técnicas para extraer y analizar los requerimientos de conocimiento y ejecución para determinadas tareas que involucran habilidades cognitivas complejas (Ryder & Redding, 1993).

3.1.2 Análisis de Protocolos

El análisis de protocolos consiste en que los sujetos piensen en voz alta mientras ejecutan o describen una tarea, para después utilizar la verbalizaciones para inferir el procesamiento cognitivo de los participantes (Ryder & Redding, 1993).

La precisión de los reportes verbales depende de los procedimientos para extraerlos (por ej. las instrucciones dadas, si los reportes son retrospectivos o concurrentes). De la misma forma, en el caso de los expertos las habilidades de la tarea son más automáticas, y por lo tanto los procesos mentales intermedios se vuelven inaccesibles para los reportes verbales. Por lo tanto, aunque el análisis de protocolos es una valiosa herramienta analítica, debe ser utilizada, en algunos casos, junto con otros métodos.

El análisis de protocolos es utilizado principalmente para derivar información acerca de los procesos de razonamiento, las estrategias de toma de decisión, y los modelos mentales (Ryder & Redding, 1993). Los resultados posibles incluyen, por ejemplo, conjuntos de reglas de producción, árboles de decisión, estrategias heurísticas, algoritmos, etc. Se pueden realizar comparaciones específicas (por ej. hipótesis generadas, errores cometidos, procedimientos utilizados) entre los protocolos obtenidos de los individuos de diferentes niveles de pericia, lo cual es útil para determinar las progresiones que se tienen que dar en el entrenamiento.

3.1.3 Técnicas de Entrevista Cognitiva

El análisis cognitivo de tareas no necesariamente requiere del uso de técnicas complicadas que consumen mucho tiempo, en cambio, se puede obtener información valiosa a través de entrevistas. Las técnicas de entrevista cognitiva involucran procesos estructurados o semiestructurados, en los cuales la selección de la pregunta se basa en una teoría cognitiva de la pericia y en un marco de referencia deseado para representar o usar los resultados, tal como una técnica de modelamiento específico o un marco de referencia de diseño instruccional. La característica que hace "cognitiva" a una entrevista es el enfoque en el conocimiento y los procesos cognitivos que subyacen la ejecución: los conceptos y sus interrelaciones, las estrategias de toma de decisión, etc. Por ejemplo, solicitar dibujos de relaciones físicas o funcionales entre los componentes de un sistema puede proporcionar información acerca de los modelos mentales que rigen el pensamiento del sujeto.

Muchas técnicas de entrevista incluyen los protocolos retrospectivos, es decir, discusiones de la ejecución de tareas realizadas. El método de decisión crítica es una técnica que utiliza la recolección de los incidentes no rutinarios que requiere el juicio del experto. Una vez que el incidente se ha descrito, se utilizan preguntas para extraer información acerca de las metas que guiaron la ejecución, las opciones evaluadas, las claves utilizadas, y los factores situacionales relevantes.

3.1.4 Modelamiento Cognitivo y de Ejecución

Uno de los desarrollos recientes en ACT involucra la construcción de modelos cognitivos y/o de ejecución para dominios complejos. El modelamiento cognitivo intenta capturar los procesos cognitivos que describen con precisión o predicen la ejecución humana. Los modelos, que son ya sea computacionales (simulación por computadora) o conceptuales (flujo de procesos), normalmente incluyen características del ambiente físico, el conocimiento del operador, y los métodos y estrategias del operador para la ejecución de la tarea. Los modelos de ejecución usualmente no tienen componentes que representen los operadores del conocimiento; tienen solamente secuencias de los elementos de la tarea, aunque pueden incluir tanto elementos cognitivos como conductuales. Los modelos pueden ser ejercitados variando las entradas, las características del operador, o las características de la tarea a fin de evaluar su efecto en la ejecución simulada. Este proceso es adecuado para la investigación sistemática de la variabilidad en la conducta, la ejecución y los procesos cognitivos conforme las condiciones cambian.

Al grado de que un modelo describe con precisión la ejecución, la representación del conocimiento, las estrategias de razonamiento, o las descomposiciones de la tarea pueden asumirse para modelar estructuras y procesos cognitivos humanos. Así, tales modelos proporcionan una especificación de los conocimientos y habilidades requeridos para el entrenamiento.

Un ejemplo de modelamiento cognitivo es GOMS (Goals, Operators, Methods, and Selection rules), una metodología desarrollada por Card, Moran y Newell (Ryder & Redding, 1993) que ha probado su efectividad. Los modelos de GOMS representan las tareas como jerarquías de medio-fin que consisten en:

- **Metas:** deseadas y estados.
- **Operadores:** acciones perceptuales, cognitivas y motoras que el usuario debe utilizar.
- **Métodos:** secuencias de acciones que constituyen procedimientos para cumplir con las metas.
- **Reglas de selección:** criterios para seleccionar entre los diferentes métodos para hacer una tarea y en consecuencia, definir las reglas de uso posibles.

Las metas de la tarea de alto nivel se descomponen en submetas de bajo nivel, y luego en "métodos" y "operadores". Tales métodos pueden usarse para representar habilidades procedimentales para el diseño del entrenamiento.

Otro método que está basado en GOMS es el denominado COGNET (COGNitive NETwork of tasks) desarrollado por Zachary, Ryder, & Hicinthom, (En prensa). Este método agrega el modelamiento de tareas que involucran la multitarea en tiempo real, esto es, tareas en las cuales los operadores tienen muchas tareas

que completar y demandan atención en ellas en base a los cambios situacionales momento a momento. Adicionalmente COGNET incluye una representación del conocimiento global que es utilizada en la ejecución de todas las tareas, e incluye también un mecanismo para los cambios en el contexto del problema para agregarse a la representación del problema. Debido a que COGNET modela el papel del conocimiento, experiencia, y los cambios situacionales en la ejecución de la tarea, proporciona elementos adicionales a la metodología de modelamiento de GOMS basada solamente en las metas. Cada uno de los componentes de COGNET, la representación del problema y los modelos de la tarea (los cuales incluyen submetas y condiciones de iniciación) se pueden usar para entrenar explícitamente los aspectos cognitivos del dominio.

3.2 Dominios de aplicación del Análisis Cognitivo de Tareas

3.2.1 Análisis Cognitivo de Tareas en Técnicos de Aviación

Means (1993) realizó una investigación acerca de las habilidades y conocimiento requerido por los técnicos en electrónica trabajando en tareas de la Fuerza Aérea que requieren mantenimiento de equipo de alta tecnología. En este estudio se ilustraron tres características del análisis cognitivo de tareas: a) la importancia de conocimiento extensivo del área y del papel de expertos del área, b) el enfoque sobre cómo se representan los problemas, y c) la necesidad de trabajar con instancias concretas para ayudar a los expertos en el área para acceder al conocimiento que se ha convertido en tácito y las habilidades que se han vuelto automáticas.

El aspecto central en el desarrollo de la metodología de análisis de tareas para el estudio sobre los técnicos de aviación de la Fuerza Aérea, fue la utilización de problemas realistas específicos que son representativos de las tareas críticas de la actividad. En la investigación de Means (1993), los técnicos de mantenimiento de aviación participaron en una actividad que inicia al recibir las "cajas negras" de las aeronaves que se han diagnosticado con fallas por la línea aérea. El trabajo consiste en localizar la falla y arreglarla (si es posible hacerlo reemplazando un microcircuito) o enviar la parte al depósito para una posterior reparación. Para evaluar las cajas negras, normalmente se utilizan grandes estaciones de prueba controladas por computadora. La estación de prueba simula las funciones de la aeronave, enviando señales a la caja negra y luego recibiendo señales de la caja comparándola con las especificaciones. Se requería contar con un conjunto de problemas representativos de las tareas más difíciles del trabajo, de manera que sirvieran para el entrenamiento de los técnicos. Means y colaboradores (1993), se percataron de la conveniencia de presentar los problemas en forma verbal, en lugar de utilizar el equipo real que era voluminoso, caro y de baja demanda operacional. Esto significa que cada vez que se le presentaba un problema a un sujeto y éste indicaba la acción (por ejemplo, medir el voltaje en un punto particular) el equipo de investigación

debería ser capaz de decirle cuál sería el resultado posible. Tanto el desarrollo de los problemas como la retroalimentación que se les proporcionaba a los sujetos que se encontraban solucionando el problema, requerían un cierto grado de conocimiento del campo de reparación de fallas en el equipo electrónico de aviación

En respuesta a esta situación, Means (1993) desarrolló un método que incluía expertos en el área, no sólo como sujetos sino también como colaboradores de investigación. Para cada actividad particular, se identificaron los sujetos con mayores habilidades (cuatro a seis individuos). Se pidió a los sujetos definir los tipos de problemas de reparación de fallas que, de acuerdo a su experiencia, fueran los cuellos de botella para los técnicos, aún después de completar su programa normal de entrenamiento y después de varios años de experiencia en el trabajo. Los tipos de problemas se referían a cosas que están mal (p. ej. clavija doblada 18 en la tarjeta A8). Después de que cada experto generó una lista de tipos de problemas que ocurren frecuentemente y que dificultan la reparación de fallas, los expertos se reunieron en un grupo para clasificar los tipos de problemas en categorías más generales (p. ej. problemas de control de datos). Las categorías de problemas resultantes representaron los requerimientos de dificultad de reparación de fallas para el trabajo.

Para cada categoría de problema, se le dio la responsabilidad a un experto para desarrollar un problema representativo que sería apropiado para usarse como una herramienta de diagnóstico o como un contexto de entrenamiento. El experto trabajó con un investigador para trazar los síntomas que aparecían primero, los pasos alternativos que un técnico podría tomar, y cual sería la falla del sistema hipotético en respuesta a cada acción. Estos procesos incluyeron la descripción del experto de como él mismo repararía la falla y su representación gráfica de las partes del sistema relevantes para el problema.

Una vez que se trazó el problema y fue discutido minuciosamente por el experto en el área y el investigador, les fue administrado a otros técnicos. Se inició por darle el problema a otro técnico experto. El experto que desarrolló el problema describió el escenario en el cual se observaron los síntomas iniciales. Mientras el experto se encontraba resolviendo el problema, se le pidió que describiera qué haría una vez que encontrara esos síntomas y por qué razones. Se le permitió al experto consultar diagramas u otros documentos técnicos (tal como sucede en el trabajo real). En el caso de que indicara que tomaría cierta medición o correría una prueba de diagnóstico automatizada, el experto actuó simulando el sistema, expresando cuál sería el resultado de tal prueba o medición. La tarea del investigador fue extraer el pensamiento subyacente del experto sobre su elección de acciones de reparación de fallas y la interpretación de la información que recibió de cada acción. En este trabajo, los problemas de reparación de fallas más difíciles son los que surgen cuando la estación de prueba en sí misma tiene una falla. La estación de prueba consiste en tres bancos con más de una docena de cajones cada uno, con cada cajón lleno de circuitos impresos. La

documentación técnica del sistema consta de 60 volúmenes.

El siguiente problema, bastante simple, ilustra el tipo de información obtenida: En este tipo de problema, el técnico se encuentra en el taller de mantenimiento e intenta encender la estación automática de prueba. Cuando el disco está instalado y se enciende el interruptor, el disco gira y se enciende la luz indicando que está listo, pero el sistema no inicia.

Se elaboró un rastro de solución del problema para cada sujeto evaluado, al intentar resolverlo. El rastro consistió en unidades llamadas *secuencias*. Las secuencias se estructuraron alrededor de las acciones del técnico, las cuales son concretas y fáciles de explicar por parte de los entrevistados. Se identificaron dos clases de acciones:

- a) actividades de recolección de información, tal como leer los esquemas, y
- b) acciones en el sistema, tales como correr una prueba de diagnóstico.

Después de obtener la secuencia de acciones (registro conductual), se identificó la representación del problema que motivó cada acción. Se examinó al técnico en cuanto al razonamiento, el planteamiento de hipótesis y el establecimiento de metas subyacentes para cada acción. También se registró el resultado de cada acción (la información obtenida en el caso de las actividades de recolección de información o la conducta del sistema, en el caso de las acciones sobre el sistema) y la interpretación de esa información realizada por el técnico.

La tabla 2 muestra el rastro de solución del problema de un experto, a partir de su protocolo verbal.

Tabla 2. Ejemplo de registro de la solución de un problema de descompostura realizado por técnicos de aviación

Secuencia	Rastro de Solución	Conocimiento Declarativo	Conocimiento Procedimental
S. 1	<p>RP: La computadora no es parte del problema porque se reinicia por sí misma. El problema debe estar en el disco duro o en la interfaz entre el disco y la computadora. Si la luz del de los datos del disco parpadea, los datos de la computadora van hacia el disco duro.</p> <p>A: Reiniciar y observar la luz de los datos en el disco duro.</p> <p>R: La luz de los datos no enciende.</p> <p>I: El problema no está el disco duro. Se sospecha de la interfaz entre la computadora y el disco duro. (por ejemplo la tarjeta controladora del disco, la controladora del disco, o los cables).</p>	<ul style="list-style-type: none"> • Conocer los componentes que involucran el reinicio. • Conocer la función de la luz de datos del disco duro. 	<ul style="list-style-type: none"> • Modelo del Recurso: Subsistema de interfaz de la computadora. • Modelo del recurso: Disco duro.
S. 2	<p>RP: Se asume que los cables están bien porque nadie estuvo trabajando en la estación.</p> <p>A: Mirar los indicadores de falla de energía dentro del controlador del disco.</p> <p>R: Las luces no están encendidas</p> <p>I: El problema no parece ser la energía en el controlador del disco, pero podría no ser una falla en la luz porque tiene 0 energía.</p>	<ul style="list-style-type: none"> • Conocer el rango de falla de del controlador del disco. • Conocer cómo funcionan las luces indicadoras de falla de energía. • Conocer dónde están los indicadores de falla de energía dentro del controlador del disco. 	<ul style="list-style-type: none"> • Conocimiento de las probabilidades de falla (controlador del disco). • Modelo del recurso: Indicadores de falla. • Modelo del recurso: Controlador del disco.
S. 3	<p>PR: Si no hay corriente hacia el controlador del disco, las luces indicadoras de falla no pueden encender. Se puede revisar la corriente que llega al controlador del disco en el fusible.</p> <p>A: Se quita el fusible y se verifican los ohms utilizando el DMM.</p> <p>R: Cerca de 0 ohms.</p> <p>I: El fusible del controlador del disco está bien.</p>	<ul style="list-style-type: none"> • Conocer donde está el fusible dentro del controlador del disco. • Conocer cómo usar el Multímetro Digital (DMM). 	<ul style="list-style-type: none"> • Modelo del recurso: Controlador del disco. • Utilización del DMM para verificaciones de continuidad y resistencia.
S. 4	<p>PR: Todavía podríamos tener una falla de corriente si está mal el regulador P/S. Los reguladores de corriente fallan frecuentemente.</p> <p>A: Medir el voltaje en los puntos de prueba del regulador de corriente.</p> <p>R: + - 5 y + - 20 volts.</p> <p>I: La energía que sale del regulador de corriente está bien, así que el controlador de disco debe estar bien.</p>	<ul style="list-style-type: none"> • Conocer como se relaciona el regulador P/S con las luces indicadoras de falla de energía. • Conocer que el regulador P/S del controlador del disco falla a menudo. • Conocer cuales puntos de prueba en el controlador del disco se deben medir. • Conocer como medir voltajes con el DMM. 	<ul style="list-style-type: none"> • Modelo del recurso: Regulador de suministro de corriente. • Conocimiento de probabilidades de falla (Regulador de suministro de corriente). • Modelo del recurso: controlador del disco. • Utilización del DMM para probar voltajes.
S. 5	<p>PR: Se sospecha del cableado, porque se pueden oír las cabezas tratando de acceder, el cable de control de los datos y el cable entre la computadora y el controlador del disco están probablemente bien.</p> <p>A: Verificar el cable de lectura/escritura entre el controlador del disco y el disco duro.</p> <p>R: El cable está desconectado en la entrada del controlador del disco.</p> <p>I: Falta identificada.</p>	<ul style="list-style-type: none"> • Conocer que los cables a menudo se desconectan del controlador del disco. • Conocer que si la cabeza del disco duro gira, se está transfiriendo información de la computadora al disco duro a través del cable de control. 	<ul style="list-style-type: none"> • Conocimiento de probabilidades de falla (cables). • Modelos de los recursos: Disco duro y cable de control.

NOTA. Claves: PR.= Representación del Problema², A.= Acción, R.= resultado, I.= Interpretación. (Tomada de Means, 1993, p. 106-107).

² Por sus siglas en inglés: PR: Problem Representation, se conservó este orden en la nomenclatura para preservar el nombre original de la técnica.

Una vez grabada la solución del problema como se muestra en la primera columna de la tabla 2, se procedió a trabajar con los expertos en el área para analizarlo en términos del conocimiento, habilidades, y estrategias empleadas. Éstas se muestran en la segunda columna de la tabla. Mientras que la columna 2 describe las habilidades y conocimiento en términos específicos para este problema (por ejemplo “el conocimiento de que el suministro de energía del disco controlador falla en la mayoría de los casos”). La columna 3 contiene una descripción más genérica (por ejemplo, “el conocimiento de la probabilidad de fallas”), escrito en una forma que es aplicable de manera general a los problemas. Ésta última habilidad y categorías de conocimiento, las cuales son más amplias y más genéricas que las habilidades en la columna 2, son útiles para obtener el sentido de la importancia de diversas habilidades, conocimiento, y componentes de la estrategia dentro de un trabajo.

Cómo podemos ver, el ejemplo anterior ilustra de manera detallada la realización de la extracción del conocimiento experto en un dominio complejo, como es la detección de fallas electrónicas en equipos de aviación. Esta investigación que parece sencilla, por la forma en que se describe, implica el involucramiento en una disciplina, que consiste en indagar los aspectos claves del conocimiento de la disciplina y la inmersión en el propio ambiente y las actividades típicas del personal que realiza el trabajo en escenarios naturales, esto tiene implicaciones en el consumo de tiempo de análisis y recolección de información.

3.2.2 Ejemplo de la aplicación del procedimiento automatizado de análisis cognitivo de tareas DRE (Descomponer, Red, Evaluación)

Generalmente, las técnicas del Análisis Cognitivo de Tareas son extenuantes y requieren una cantidad considerable de tiempo de análisis, además requieren de la participación cercana del grupo de expertos en la materia en cuestión (lo cual no siempre es posible) trabajando largas jornadas en coordinación con los experimentadores. Valerie Shute y su grupo de colaboradores (Shute, Surgue, & Willis, 1997), han trabajado en los años recientes en desarrollar una metodología automatizada para la extracción del conocimiento experto, que pueda ser aplicable a un amplio rango de disciplinas. El resultado es un software construido con los elementos de un sistema experto que genera una secuencia de preguntas automatizadas, conocido como “Descomponer, Red, Evaluación” DRE (en inglés se conoce como DNA: Decompose, Network, Assess), que se van actualizando con las respuestas del experto usuario, las respuestas del experto se almacenan en una base de datos, para formular un modelo cognitivo de la disciplina compleja. A continuación se presenta un ejemplo proporcionado por la misma investigadora, sobre la aplicación del prototipo de DRE en el juego de *Blackjack*. La secuencia de preguntas que se describe a continuación, incluye ejemplos de preguntas de DRE, las cuales analizan el dominio del juego *Blackjack* (Shute et al., 1997). Las preguntas generales generadas por DRE se asocian con las variables que proporcionan las respuestas de los expertos a

pruebas previas. Las preguntas generales están redactadas de tal manera que puedan ser usadas a través de una variedad de dominios, mientras que las variables son usadas para hacer preguntas específicas del dominio a analizar. (Nota: en los siguientes ejemplos, las variables están indicadas por letras *itálicas* dentro de la pregunta).

Un EM (experto en la materia) responde a las diversas preguntas escogiendo palabras y frases de diferentes cuadros de diálogo que contienen las palabras designadas por las respuestas de los EM, y también tecleando el texto. Los cuadros de diálogo se usan con la intención de superar el problema del lenguaje natural. Las formas para teclear texto libre permiten a los EM verbalizar efectivamente el conocimiento y las habilidades requeridas. En los ejemplos proporcionados, las opciones seleccionadas de los cuadros de diálogo se despliegan en texto subrayado, mientras que las respuestas de texto libre se despliegan en fuente normal.

Secuencia

El DRE comienza haciendo preguntas amplias, seguidas de unas preguntas para extraer información asociada con un nodo especificado. El experto tiene la opción de contestar cada pregunta con, ya sea "Yo se como" hacer algo (procedimental), o "Yo entiendo" algo (conceptual).

Ejemplo 1:

DRE: ¿Qué es lo que hace que domine el *Blackjack*?

EM: Yo sé cómo contar cartas.

La ruta "Yo sé cómo" está diseñada principalmente para extraer las HP (habilidades procedimentales). Sin embargo se toman las provisiones para extraer el CS (conocimiento simbólico) y CC (conocimiento conceptual) asociado con los elementos de las HP (habilidades procedimentales) descritos (por ej., "contar cartas"). La ruta "Yo se como" se describe con mayor detalle más adelante. Con base en la respuesta actual, el programa pregunta al experto lo siguiente.

Ejemplo 2:

DRE: ¿Qué es lo que hace que domine el *Blackjack*?

EM: Yo comprendo las estrategias de contar cartas.

La ruta "Yo comprendo" está diseñada para obtener más CC (conocimiento conceptual) así como CS (conocimiento simbólico) y HP (habilidades procedimentales) asociadas. La ruta "Yo comprendo" se describe con mayor detalle más adelante.

La ruta "Yo sé cómo" es un conjunto de preguntas organizadas linealmente diseñadas principalmente extraer el CP (conocimiento procedimental). Sin embargo al incluir la pregunta "por qué es importante este paso" se extrae CC adicional asociado a ese paso. Al incluir la pregunta "cuándo ejecuta usted este paso" se extraen elementos de CC adicionales, lo mismo sucede con la pregunta "bajo que condiciones se ejecuta este paso". Los expertos tienen la opción de definir los términos y conceptos relacionados durante el desenvolvimiento de cada paso. Esta opción extrae el CS asociado con el paso especificado. La ruta "Yo sé cómo" está organizada de manera lineal debido a que se considera que los procedimientos se perciben inherentemente de forma lineal, (e.g. paso por paso).

Ejemplo 3:

DRE: ¿Cuál es el primer paso que se sigue generalmente cuando usted *cuenta cartas*?

EM: Inicializo mi contador mental de cartas.

Ejemplo 4:

DRE: Asígneme un nombre

EM: Inicializo el contador mental

A continuación se presenta al EM una pantalla que presenta una pregunta de "por qué" diseñada para extraer CC asociado a ese paso.

Ejemplo 5:

DRE: ¿Por qué es importante el paso *inicializar el contador mental*?

EM: Debido a que todas las cartas están en el zapato, las probabilidades son todas iguales.

También se encuentra en la pantalla un botón disponible para definir cualquier término que pueda ser usado en la descripción del EM. Cuando se selecciona este botón, se presenta al EM una pantalla que le permite ingresar el término y la definición asociada. También se le solicita al EM que incluya un ejemplo del término en la definición.

Ejemplo 6:

DRE: Término

EM: Zapato

DRE: Definición:

EM: Es una cosa en el Blackjack donde se guardan muchas barajas.

Luego se le presenta al EM una pantalla la cual presenta una pregunta "cuándo" diseñada para extraer CC adicional asociado con ese paso.

Ejemplo 7:

DRE: ¿Cuándo realiza usted el paso: *inicializar el contador mental*?

EM: Cuando se ha barajado el zapato.

Después de ingresar la información asociada con un paso, se le permite al EM ingresar pasos adicionales, editar los pasos existentes, o borrar los pasos ingresados anteriormente. Cuando se han completado todas las revisiones de los pasos, se le pide al EM colocar todos los pasos en el orden en que se ejecutan típicamente. Una vez que se han ordenado los pasos, el DNA regresa al EM a la pregunta "Qué es lo que hace que domine _____" la cual permite al EM definir una meta supraordenada.

La ruta "yo comprendo" está organizada horizontalmente porque el conocimiento no necesariamente está en orden lineal. El EM no tiene que contestar todas las preguntas (como las enlistadas anteriormente) ya que algunas pueden ser irrelevantes. La opción "ninguna de las anteriores se aplica, pero lo que quiero decir" se incluye en el caso en que el EM no encuentre ninguna de las preguntas relevantes. Al seleccionar la opción "Ninguna de las anteriores" el EM pasa a una pantalla que acepta teclear texto libre sin ninguna limitación. Nota: al seleccionar la pregunta "cómo" se conduce al EM a la rama "yo sé cómo" para obtener información acerca de los pasos. Después de completar la rama "yo sé cómo" se regresa al EM a seleccionar preguntas adicionales.

Ejemplo 1:

DRE: ¿Por qué es importante para usted comprender *las estrategias de contar cartas*?

EM: Porque esto me permite estimar las posibilidades de las 10 que se pueden repartir.

Ejemplo 2:

DRE: ¿Cuál es la definición de *las estrategias de contar cartas*?

EM: Una estrategia en el Blackjack que involucra mantener un conteo del número de las 10 que se han jugado en relación con el número total de cartas jugada en un momento determinado.

Ejemplo 3:

DRE: ¿Cuáles son los conceptos y definiciones relacionados con *las estrategias de contar cartas*?

DRE: Concepto:

EM: La situación que tenga la posibilidad con menor promedio de las 10 cartas repartidas.

DRE: Definición:

EM: Una gran cantidad de los juegos de 10 ya han sido jugados, así que pocos juegos de 10 permanecen en la baraja. Consecuentemente, hay una posibilidad reducida de los juegos de 10 que ya han sido repartidos.

Este proceso interactivo se repite para definiciones adicionales.

En las secciones anteriores se han mencionado conceptos como, conocimiento, habilidades, conceptos, aprendizaje, representación, etc.; todos ellos pertenecientes al campo de estudio de la Psicología Cognitiva e íntimamente relacionados con el estudio del conocimiento experto en el contexto de la solución de problemas en dominios complejos, tal como la adquisición de habilidades y conocimientos en estadística. Como se ha podido percibir en la lectura de los diversos autores que se han abocado a la descripción y caracterización del conocimiento experto en diferentes disciplinas, un aspecto central es la representación del conocimiento. La teoría de los esquemas ocupa un papel preponderante, junto con otras teorías surgidas de la psicología cognitiva para comprender el camino complejo que sigue un novato para convertirse en experto en un dominio específico. En la siguiente sección, se describirá brevemente la teoría de los esquemas de Rumelhart y Norman (1975), la cual ha sido sumamente influyente para la comprensión y estudio de las representaciones.

4 La teoría de los esquemas y la representación del conocimiento

4.1 La teoría de los esquemas

La teoría de los esquemas pertenece a un grupo de teorías sobre el aprendizaje conocidas como computacionales, correspondiente al paradigma del procesamiento humano de la información. Entre las teorías del aprendizaje computacional referentes a la adquisición del conocimiento, es posible mencionar tres ejemplos típicos: la teoría ACT³ (Control Adaptativo del Pensamiento) de Anderson (1983), la teoría de los esquemas de Rumelhart y Norman (1978) y la teoría más reciente de la inducción pragmática de Holland Holyoak y Thagard (1986). En los últimos años, el enfoque del procesamiento de la información ha permitido generar teorías del aprendizaje basadas en los supuestos computacionales. Entre los primeros modelos producidos por el enfoque del procesamiento de la información, se encuentra el *Solucionador General de Problemas* propuesto por (Newell & Simon, 1972), la característica de estos modelos es que se diseñaban a partir de sistemas de procesamiento con una gran capacidad sintáctica general, que les permitía enfrentarse a cualquier tarea sin necesidad de conocimientos específicos. Sin embargo, esto no fue suficiente para explicar los procesos de solución de problemas. Los investigadores se dieron cuenta entonces que, además de una capacidad sintáctica, los sistemas de procesamiento, ya sean artificiales o humanos, requerían de una cierta cantidad de conocimientos específicos para enfrentarse a un problema complejo y además rico semánticamente. Entonces apareció el diseño de sistemas expertos en áreas temáticas específicas, dotados de grandes conocimientos en esa área que les permitiría resolver problemas enormemente complejos, con una gran eficacia, siempre y cuando pertenecieran al mismo dominio de conocimientos. Entonces surgió una disciplina en constante evolución, que se conoce como *Inteligencia Artificial*. Buena parte de los esfuerzos de esta disciplina se enfocan en la actualidad al diseño de *sistemas expertos*.

De esta manera, surge la necesidad de construir sistemas que cuenten con una amplia base de conocimientos semánticamente ricos, y además surge otra necesidad tanto teórica como técnica, de proporcionarles una cierta capacidad para adquirir esos conocimientos. Entonces, el aprendizaje representa un problema relevante para la ciencia cognitiva. Desde un punto de vista teórico es necesario postular los mecanismos que expliquen cómo adquieren los conocimientos los sistemas humanos de procesamiento. Desde un punto de vista técnico, es necesario dotar a las computadoras de la capacidad para adquirir por sí mismas los conocimientos complejos, incluidos los conceptos. Estos esfuerzos han desembocado en una serie de teorías computacionales del aprendizaje que difieren entre sí en aspectos importantes, sin embargo, comparten un núcleo conceptual. En esencia, todas ellas intentan reducir la semántica a reglas sintácticas. Por tanto, la adquisición de conceptos deberá

³ En inglés: Adaptive Control of Thought

también explicarse sintácticamente. A continuación revisaremos sólo la teoría de los esquemas por ser de relevancia para el presente estudio, ya que se encuentra íntimamente relacionada con la representación del conocimiento y la caracterización del conocimiento experto en dominios complejos, como la solución de problemas de estadística.

Desde la perspectiva del procesamiento humano de información, el concepto *esquema* tiene un origen computacional. Más concretamente, el término de *esquema* ha sido recuperado en el marco de los estudios sobre inteligencia artificial. Diversos autores consideran que elaborar un programa computacional significa que debe ser capaz de realizar las tareas cognitivas que generalmente hacen los humanos, dicho de otra manera, se trata de crear cierta inteligencia artificial: la cual debe tener no sólo una considerable potencia sintáctica sino también un componente semántico. Por ejemplo, si se quiere diseñar un programa que comprenda textos, no basta con proporcionarle un conocimiento sintáctico y un léxico general. El programa debe tener una cierta cantidad de conocimientos específicos a la temática del texto presentado. Estos conocimientos constituyen los esquemas del sistema de procesamiento (Pozo, 1993).

De acuerdo a Rumelhart (1982), los esquemas son los conceptos de que dispone el sistema de procesamiento "una esquema es una estructura de datos para representar conceptos genéricos almacenados en la memoria". Por este motivo la teoría de los esquemas puede considerarse como una teoría de la representación y utilización de los conceptos almacenados en la memoria. Según Rumelhart, todo procesamiento de información depende de la activación de esquemas.

La teoría de los esquemas tiene que ver con un modelo general de cómo se representa el conocimiento y cómo se usa el conocimiento almacenado (Rumelhart & Ortony, 1982). La unidad básica de procesamiento son los esquemas, que consisten en *paquetes información* sobre conceptos genéricos. Los esquemas son unidades molares relativamente complejas, no son atómicos. Un esquema contiene, una red de interrelaciones que existe entre los constituyentes del concepto en cuestión. Dicho de otro modo, se puede describir un esquema como un concepto de un objeto, persona o situación, cuyos rasgos característicos son correspondientes con los atributos del concepto. Más específicamente, una teoría del esquema implica una teoría *prototípica del significado*. Rumelhart y Orthony (1982) comparan un esquema con el libreto de una obra de teatro. Cada representación concreta de la obra de teatro diferirá en algunos aspectos del texto original, sin embargo, en todas las representaciones se reconocerá fácilmente el texto original. Lo mismo sucede con los esquemas. Cada vez que se actualiza un esquema, existen ciertos rasgos diferenciales que no interfieren en la aplicación del mismo. Rumelhart considera que los esquemas son como el guión de un juego, como teorías o como procedimientos prototípicos.

Asimismo, los esquemas se caracterizan por su distinción entre conocimiento declarativo y procedimental. Aunque las características del conocimiento procedimental parecen asemejarse más al funcionamiento cognitivo humano, las representaciones exclusivamente procedimentales resultan difícilmente auto modificables. Por tanto surge la necesidad de considerar los procesos de aprendizaje de los esquemas lo que obliga a dotarles también de un carácter declarativo. El *saber que* y el *saber cómo*, no serían sino dos usos distintos de una misma representación de conocimiento en forma de esquemas, constituidos en redes preposicionales. La naturaleza flexible de los esquemas les permite ser utilizados tanto de modo declarativo como procedimental. Esto es posible porque los esquemas son, de acuerdo a Rumelhart, "paquetes de conocimiento en los que, además del propio conocimiento hay información sobre cómo debe usarse ese conocimiento".

Norman et al. (1975) proponen que todo el conocimiento de los esquemas puede representarse en forma de redes proporcionales. Se pueden considerar cuatro características fundamentales de los esquemas:

- 1) Los esquemas tienen variables
- 2) Los esquemas pueden encajarse unos en otros
- 3) Los esquemas representan conceptos genéricos que varían en sus niveles de abstracción
- 4) Los esquemas representan conocimientos más que definiciones

4.2 Aprendizaje por modificación y generación de esquemas

Una diferencia de la teoría del aprendizaje de esquemas de Rumelhart y Norman (1978) con la teoría ACT de Anderson es que, ya que se origina en una fundamentación empírica psicológica, como consecuencia de ello, los mecanismos del aprendizaje de esquemas van a carecer de la precisión y coherencia teórica que tenía el aprendizaje en el ACT: sin embargo, una de sus virtudes principales, es que intenta superar algunos de los límites del cambio de producciones en el ACT. Podríamos decir que Anderson se limita a proponer aquellos mecanismos que son computacionalmente posibles, mientras que Rumelhart y Norman formulan los mecanismos que un sistema cognitivo debería detener para adquirir esquemas semánticamente complejos; sin embargo, esos mecanismos no necesariamente se encuentran en un programa de cómputo. Según Rumelhart y Norman (1978), desde un punto de vista lógico pueden distinguirse tres tipos de aprendizaje: el *crecimiento*, la *reestructuración* y el *ajuste*.

4.2.1 Crecimiento

Mediante el *crecimiento* se acumula nueva información en los esquemas existentes. Las leyes que rigen este crecimiento son básicamente asociativas: siempre que se encuentra nueva información, se supone que se guarda en la

memoria alguna huella del proceso de comprensión. Esta huella de memoria es la base del recuerdo. Generalmente, se supone que esas son copias parciales del esquema original actualizado. El *crecimiento* es el mecanismo básico por el que el sistema adquiere las bases de datos con las que se rellenan las variables de los esquemas. Es un aprendizaje de hechos que no difiere en lo esencial de los procesos clásicos de aprendizaje por asociación. Se trata del mecanismo más frecuentemente usado y el más estudiado la literatura sobre aprendizaje. Pero también se trata del mecanismo que menos cambios produce en el sistema de esquemas. De hecho, el crecimiento no modifica la estructura interna de los esquemas ni genera por si mismo esquemas nuevos. La nueva información no modifica la estructura de los conocimiento y existentes. Los *esquemas* necesarios para interpretar la información tienen que estar presentes ya antes de que tenga lugar el aprendizaje. Si la única forma de aprender fuera mediante crecimiento de los esquemas, nunca se formarían nuevos conceptos. Para que esto suceda, es lógicamente necesaria la acción de los otros dos mecanismos: el *ajuste* y la *reestructuración*.

4.2.2 Ajuste

Cuando los *esquemas* disponibles no bastan para comprender o interpretar una situación es necesario generar nuevos esquemas o modificar los ya existentes. La modificación o evolución de los esquemas disponibles tiene lugar mediante un proceso de *ajuste*. Se utiliza cuando, para formar el nuevo esquema, basta con realizar modificaciones en las variables y constante de un esquema sin que sea necesario cambiar la estructura interna del mismo. El *ajuste fino* de los esquemas puede producirse de tres formas.

- a) En primer lugar, un esquema se puede ajustar, mediante *modificación de los valores por defecto* de un esquema, en función de la experiencia en la aplicación del mismo.
- b) Un segundo tipo de ajuste consiste en una *generalización del concepto* al sustituir un valor constante del esquema por una variable.
- c) Por último, el ajuste por *especialización de concepto* se produce cuando se sustituye una variable del esquema por una constante.

El ajuste es producto de la práctica y tiene como consecuencia la modificación del campo de aplicación de un esquema. En realidad no produce esquemas auténticamente nuevos, ya que no afecta la estructura de los esquemas.

4.2.3 Reestructuración

La generación o creación de nuevos esquemas tiene lugar mediante el proceso de reestructuración, que consiste en la formación de nuevas estructuras conceptuales o nuevas formas de concebir las cosas. Según Norman, para que se produzca la reestructuración, sería necesario un *insight* o *comprensión súbita* de una nueva estructura en el área temática correspondiente. Mediante una

reestructuración surgen estructuras conceptuales o interpretativas que anteriormente no estaban presentes en el sistema. El nuevo esquema se origina en otros esquemas ya disponibles. Para formar los esquemas a partir de los precedentes, puede ser de dos formas:

- a) mediante una *generación pautada* del nuevo esquema, que consiste en una copia con modificaciones de un esquema anterior. Esa copia se forma mediante un proceso analógico. Así, por ejemplo un niño que ignora lo que es un rombo podemos presentarle el concepto mediante una analogía: un rombo es a un cuadrado, lo que un paralelogramo es a un rectángulo. Igualmente, los niños, que tienen serias dificultades para entender muchas nociones del tiempo histórico, pueden comprender mejor el tiempo histórico a partir de una analogía con las representaciones espaciales que ya han adquirido. Pero la formación analógica de esquemas implica generar una copia modificada del esquema. Esta modificación puede ser de tres tipos distintos:
 1. *Generalización* (cambio de una constante por una variable)
 2. *Especialización* (introducción de una constante en lugar de una variable) o
 3. *Sustitución de variables o constantes* del esquema original por nuevas variables o constantes. La generación pautada bajo formación analógica de esquemas sería el mecanismo habitual para la formación de conceptos.
- b) El segundo mecanismo de reestructuración, la *inducción de esquemas*, es poco frecuente. La inducción de un nuevo esquema se produciría cuando se detectase que cierta configuración esquemas ocurre sistemáticamente según ciertas pautas espacio-temporales. Se trataría de un modo de aprendizaje por *contigüidad*. Pero esa detección de contingencias entre esquemas requiere mecanismos de los que carece el sistema de esquemas, por lo que la inducción es un mecanismo poco frecuente de generación de esquemas. La dificultad de la inducción reside en el descubrimiento de regularidades, en la aplicación de esquemas próximos en el espacio o en el tiempo.

Los tres tipos de aprendizaje mencionados, no sólo difieren entre sí en sus mecanismos subyacentes, sino también en una serie de características. Así, cada tipo de aprendizaje es sensible a diferentes procedimientos de instrucción, y por lo tanto, debe medirse de modo distinto, tiene una capacidad de transferencia diferente y está sujeto una interferencia con otros aprendizajes anteriores de distinta intensidad. Estas diferencias en los resultados producidos por cada uno los procesos de aprendizaje hace que sean complementarios, más que independientes. Según Rumelhart y Norman (1978), *el crecimiento, la reestructuración y el ajuste* coexisten e interactúan durante todo el ciclo de aprendizaje de un sistema jerarquizado de conceptos. De hecho, existiría una conexión necesaria entre los tres tipos aprendizaje: de tal forma que no hay reestructuración sin crecimiento previo, ni ajuste sin reestructuración anterior.

Pero, aunque los tres mecanismos coexisten en todo ciclo de aprendizaje, su importancia relativa varía según una pauta temporal característica. Al comienzo del aprendizaje en un área conceptual predominaría el *crecimiento*. La acumulación de conocimiento acabaría produciendo una *reestructuración* de los esquemas. Una tercera fase se caracterizaría por un *nuevo crecimiento* de los esquemas generados, que finalmente desembocarían en un *ajuste progresivo* de los mismos. Pero, tanto los mecanismos postulados como la importancia relativa de los mismos en cada una de las fases del aprendizaje se basan más en criterios lógicos que en un apoyo empírico. Frente a la precisión y exactitud del ACT (Anderson, 1983), que permite su falsación y la continua evolución de la teoría, los esquemas resultan en muchos casos demasiado vagos.

En los capítulos previos hemos revisado diversas maneras de abordar el conocimiento en estadística inferencial. La primera aproximación surge precisamente desde la misma disciplina, en la que han participado numerosos profesores, quienes han reunido múltiples experiencias en sus clases de estadística y en el intercambio con otros colegas, ésta forma de aproximarse se apoya básicamente en el análisis de las concepciones erróneas y las dificultades en el aprendizaje y la enseñanza de la estadística. La otra aproximación surge de la ciencia cognitiva y se basa esencialmente en los estudios experto-novato y la concepción de la estadística como solución de problemas, y no como una simple adquisición de conceptos y procedimientos descontextualizados de su uso real. Esta postura supone que la estadística es una disciplina compleja y que representa una dificultad para el aprendizaje en la mayoría de los niveles estudiantiles, y por tanto, es posible abordarse con los métodos de la ciencia cognitiva, tal como el análisis cognitivo de tareas, el cuál es un método evolucionado a partir del método tradicional de análisis de tareas adicionando el componente cognitivo. Muy relacionado con los aspectos anteriores se encuentra el tema de la representación del conocimiento, en virtud de que para solucionar un problema lo primero que realizan tanto expertos es representarse un problema, independientemente de la disciplina que se trate. En los estudios experto-novato se han encontrado diferencias importantes en la forma en que los expertos y los menos expertos se representan los problemas y en consecuencia, toman diferentes decisiones para llegar a la meta. La mayoría de ejemplos que podemos encontrar se sitúan en disciplinas reconocidas por su complejidad, como el ajedrez, las matemáticas y la física. En el caso de la estadística no es la excepción (aunque no existen suficientes estudios sobre el conocimiento experto en estadística), ya que los sujetos expertos (usualmente también son profesores) poseen un cuerpo de conocimiento más estructurado sobre la disciplina, lo que les permite identificar el tipo de patrones característicos de los diferentes problemas a los que se enfrentan cotidianamente. Por su parte los sujetos menos expertos, tienen serias dificultades para representarse los problemas, mayor tiempo para resolver los problemas además de identificación de claves erróneas. En las siguientes secciones se muestra una propuesta metodológica para abordar el estudio de la estadística desde la perspectiva de la solución de problemas, en la cuál se pueden identificar ciertos patrones de problemas

estadísticos, factibles de entrenamiento, lo cual posibilitaría acortar el tiempo y los errores en el proceso instruccional, siempre y cuando se tenga un conocimiento más preciso de las dificultades específicas que se enfrentan en cada nivel de pericia, en este campo de conocimiento.

MÉTODO

Objetivo

Analizar el conocimiento experto involucrado en la solución de problemas de estadística inferencial y las diferencias que se presentan en cada nivel de pericia, en el contexto del proceso de la selección de la prueba estadística mediante la utilización de la metodología de "Análisis Cognitivo de Tareas".

Preguntas de Investigación

Preguntas principales:

¿Cómo se caracteriza el conocimiento experto en la solución de problemas que involucran la selección de la prueba estadística?

¿Cómo se diferencia la ejecución de los expertos, con respecto a la de los novatos en la elección de la prueba estadística?

Preguntas específicas

¿Cuáles son las tareas clave involucradas en la selección de la prueba estadística?

¿Cuáles son las habilidades involucradas en la ejecución de las tareas clave en la selección de la prueba estadística?

¿Cuáles son las estrategias heurísticas y modelos mentales utilizados por los expertos en la selección de la prueba estadística?

¿Cuáles son las dificultades encontradas por los novatos en la selección de la prueba estadística?

Sujetos

GRUPO DE EXPERTOS

Se seleccionaron 5 expertos en estadística (profesores o investigadores del área de Ciencias Sociales). Los expertos debían cubrir al menos dos requisitos de la siguiente lista:

- 1) Experiencia de al menos 5 años realizando investigación utilizando metodología cuantitativa.
- 2) Experiencia de al menos 5 años impartiendo asignaturas de estadística en la Licenciatura en Psicología.
- 3) Experiencia de al menos 3 años como asesores de tesis de licenciatura o posgrado.
- 4) Ser calificado como "Buen Profesor" en evaluaciones previas realizadas por los alumnos de la Facultad de Psicología.

- 5) Ser ampliamente reconocidos por los profesores de la Facultad como "Expertos en Estadística"

GRUPOS DE NOVATOS

Se seleccionaron 10 estudiantes de diferentes grados de la Facultad de Psicología de la UNAM, considerados como novatos, de acuerdo con tres niveles de pericia.

Nivel 1: Cuatro estudiantes de posgrado en Psicología (2 de maestría y 2 de doctorado).

Nivel 2: Tres alumnos en proceso de desarrollo de tesis de licenciatura en Psicología, (que implique la utilización de un procedimiento estadístico).

Nivel 3: Tres estudiantes entre 3º y 5º semestre de la licenciatura en Psicología.

Instrumentos

1.- Guía de entrevista para aplicar en la 1ª fase a los cinco expertos en estadística (Apéndice A). El instrumento servirá para identificar los aspectos generales involucrados en la toma de decisiones respecto de la prueba estadística más apropiada cuando intentan resolver un problema.

2.- Serie de 12 problemas de estadística, correspondientes a las siguientes pruebas estadísticas: 3 de Prueba t, 3 de Ji cuadrada, 3 de Correlación, y 3 de ANOVA One Way) presentados en forma escrita en tarjetas, (Tarea de clasificación, Apéndice B). Se seleccionaron estas pruebas estadísticas ya que en estudios anteriores, se señalan como las más frecuentemente utilizadas en investigaciones de Ciencias Sociales (Lavigne & Glaser, 2001, 2003; Quillici & Mayer, 1996). La tarea de clasificación de los 12 problemas tiene por objetivo identificar las formas de representar los problemas estadísticos.

3.- Problema de estadística relacionado con la aplicación de ANOVA One Way (Apéndice C). La aplicación de la tarea de solución de un problema, sirve para analizar de manera específica, cómo los expertos y novatos toman decisiones en el contexto particular de un problema estadístico.

Procedimiento

El proceso se desarrolló con base en cuatro fases o etapas, de las cuales las tres primeras etapas planteadas corresponden al *Modelo de Análisis de Tareas Integrado* MATI, (Ryder & Redding, 1993) y la cuarta etapa agregada al modelo

original consiste propiamente en el análisis y la representación de los datos recabados. Las diferentes etapas constituyen ciclos progresivos de recolección de datos, análisis y toma de decisiones, divididos en tres componentes del conocimiento experto: conocimiento, habilidades y modelos mentales. La primera etapa es *Orientación*, se encarga de crear un amplio cuerpo de conocimientos de la actividad y la determinación de los métodos de análisis de cada tarea que se llevaron a cabo en las etapas sucesivas. La siguiente etapa, *Análisis Básico*, se encarga del análisis de la ejecución competente. La tercera etapa, *Adquisición de Habilidades y Refinamiento del Análisis* se ocupa del análisis de la progresión de la adquisición de habilidades de novato a experto, considerando una fase intermedia (véase Figura 5). Finalmente, la cuarta y última etapa adicionada al modelo MATI original denominada *Análisis de datos y Representación del Conocimiento*, consiste en organizar y representar la información obtenida en las tres fases anteriores, lo que permitió elaborar el modelo cognitivo de la tarea de solución de problemas estadísticos. A continuación se describen en detalle las características particulares de cada fase del modelo.

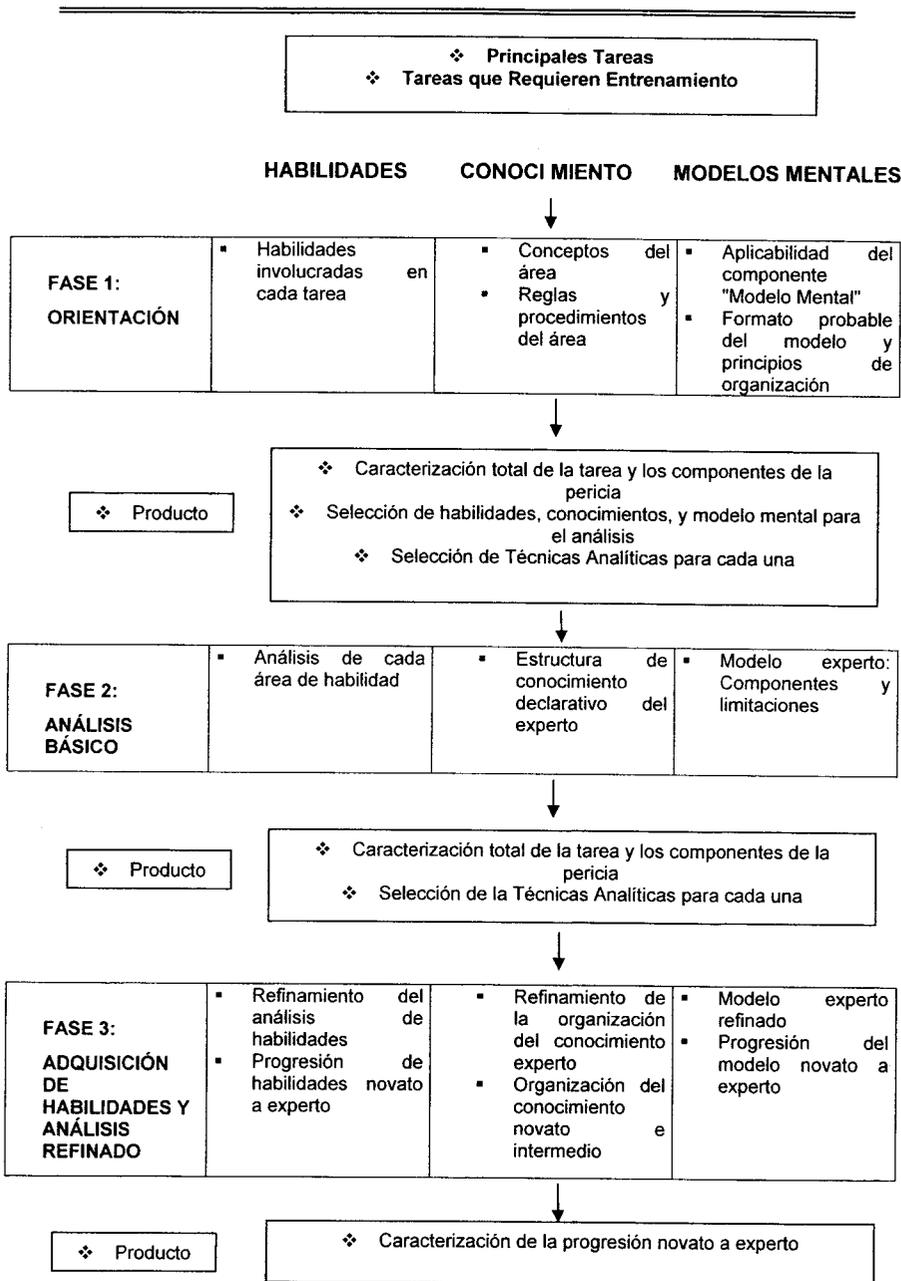


Figura 5. Esquema del Modelo de Análisis de Tareas Integrado (MATI). (Tomado de Ryder y Redding, 1993, Pág. 87).

1ª Fase: Orientación: Identificación de las Tareas Principales

El objetivo es identificar, de manera general, las tareas y los principales componentes involucrados en la pericia para la solución de un problema de estadística. Las actividades que comprenden esta etapa son las siguientes:

- Se revisó la literatura existente con respecto a los materiales instruccionales y programas para la enseñanza de la estadística, además de investigaciones realizadas sobre el tema.
- Se realizó una entrevista a profundidad a dos de los cinco *expertos* en estadística del área de Ciencias Sociales (sólo fue posible entrevistar a dos expertos, debido a la escasa disponibilidad de tiempo de los demás expertos para la entrevista). Las sesiones fueron video grabadas y transcritas para su análisis.

Para la realización de la entrevista a profundidad (ver Apéndice A) se utilizó una adaptación del formato de preguntas del *Procedimiento Automatizado de Análisis Cognitivo de Tareas* denominado DNA⁴ (Descomponer, Red, Evaluar), diseñado por Shute et al. (1997). El procedimiento DNA fue elaborado originalmente para automatizar las preguntas realizadas a los expertos durante la entrevista (que serían integradas a un software para automatizar el método de entrevista del Análisis Cognitivo de Tareas). Las preguntas realizadas intentan extraer el conocimiento conceptual y procedimental involucrado en la actividad. Para el caso de este estudio se utilizaron las preguntas diseñadas por Shute et al. (1997) en la entrevista a los expertos.

Elaboración del espacio del problema

Con los datos obtenidos a partir de la extracción del conocimiento experto, se elaboró un esquema de representación del *Espacio del Problema* del proceso de toma de decisiones para la elección de la prueba estadística, que consiste en recabar información sobre los principales pasos utilizados por los expertos para solucionar problemas estadísticos que involucran la elección de la prueba estadística.

2ª Fase: Análisis Básico: Caracterización General de la Tarea y Componentes de la Pericia

Esta fase consiste en la caracterización de la ejecución competente de la solución de problemas estadísticos, de manera más específica que en la primera fase; esto incluye: identificación de las habilidades, conocimiento conceptual y estrategias involucrados en la tarea de solución de problemas estadísticos relacionados con la elección de la prueba estadística. El objetivo fue llevar a cabo una evaluación de las diferencias cualitativas en el conocimiento que

⁴ En inglés se denomina Decompose, Network, Assess.

utilizan tanto los expertos y los novatos cuando toman decisiones, durante el procedimiento de *Elección de la Prueba Estadística*.

Para ello se llevó a cabo una sesión individual con cada uno de los sujetos novatos (10 estudiantes) y con los expertos (cinco expertos), con una duración aproximada de 1 hora. Inicialmente se les dieron las instrucciones para realizar dos tareas, 1) clasificación de problemas estadísticos (representación del problema) y 2) solución de un problema estadístico (registro del proceso de solución de un problema). Las sesiones fueron video-grabadas y se tomaron notas para su análisis posterior.

- 1) **Tarea de clasificación de problemas estadísticos.** Se proporcionó un paquete de 12 tarjetas a los sujetos, cada una de las 12 tarjetas contenía uno de los doce problemas estadísticos escritos que se seleccionaron, tomando en consideración las pruebas estadísticas más comúnmente utilizadas en la investigación en Psicología (Gardner & Hudson, 1999). Los problemas fueron adaptados de diversos libros y artículos de revistas (ver Apéndice B). El investigador solicitó a los entrevistados que clasificaran los problemas de acuerdo con las características que tienen en común, sin intentar resolverlos. Se les solicitó que pensarán en voz alta mientras realizaban la actividad en cuestión. En seguida se solicitó la explicación de las razones para realizar tal clasificación, alentando a los sujetos a que elaboraran argumentos sobre sus clasificaciones. En seguida se formularon las siguientes preguntas específicas: ¿Por qué son similares los problemas de cada grupo?, ¿En qué se diferencia cada grupo formado?, ¿Podrías hacer subclasificaciones a partir de los grupos formados?, ¿Habrías una clasificación más general con los grupos formados?, ¿Podrías elaborar otra clasificación diferente a la anterior, reorganizando los problemas?, ¿Cómo la harías?, ¿Qué criterios utilizaste para reorganizar los problemas?

- 2) **Tarea de solución de un problema estadístico.** Se presentó a los sujetos un problema estadístico (ANOVA One Way) escrito en una tarjeta (ver Apéndice C). Se proporcionaron las instrucciones consistentes en leer completamente el problema de investigación. Los entrevistados debían *pensar en voz alta* mientras intentaban resolverlo, proporcionando la solución más apropiada; asimismo se les proporcionaron hojas blancas a los entrevistados para realizar anotaciones en caso necesario. Durante la ejecución de la tarea se solicitó a los sujetos subrayar en el texto los conceptos o frases relevantes para la solución del problema. Asimismo se les indicó que realizaran la interpretación y análisis de los resultados de la aplicación del análisis estadístico con el programa SPSS, siguiendo la misma técnica de *Pensar en Voz Alta*. El investigador intervino ocasionalmente para pedir una elaboración o una clarificación adicional. Para obtener información adicional, durante el proceso de solución, se utilizó la metodología **PARI** (Representación del Problema, Acción,

Resultados e Interpretación). Esta metodología es una técnica de entrevista semiestructurada, en la cual se proporciona a los sujetos un problema de investigación y se les formulan una serie de preguntas, hasta que el sujeto resuelve el problema. En cada acción que desee emprender el sujeto, el investigador pregunta por qué realizaría tal acción. Una vez que se lleva a cabo la acción, el investigador plantea un resultado posible y el entrevistado debe formular una interpretación de los resultados de la acción. Las entrevistas de cada sujeto fueron grabadas en audio y video durante el proceso de solución del problema estadístico.

3ª Fase: Refinamiento del Análisis: Caracterización de la Progresión Novato-Experto

El objetivo principal de la presente fase es determinar la manera en que las habilidades, conocimientos y modelos mentales difieren entre los diferentes niveles de pericia. Esto tiene que ver con la comprensión de la ejecución de las tareas que componen la actividad de los sujetos expertos y los menos expertos (novatos y niveles intermedios); y las estrategias heurísticas, conocimientos y habilidades finas que los diferencian.

La tercera fase se enfocó básicamente en los aspectos que diferencian a los expertos de los menos experimentados. Adicionalmente esto permite refinar la comprensión general de la pericia en la solución de problemas estadísticos que se obtuvo en la 1ª y 2ª Fase. El procedimiento consiste en transcribir, segmentar y categorizar los datos de ambas tareas: 1) clasificación de problemas estadísticos y 2) solución del problema estadístico, con el objeto de identificar la estructura de representación de los problemas estadísticos por parte de los expertos y los novatos y los pasos en el proceso de solución de un problema estadístico.

Extracción del Conocimiento Conceptual, Procedimental y Estratégico en cada nivel de pericia

El procedimiento de la 3ª fase se construye sobre los datos obtenidos en la fase 1 y 2 para refinar la comprensión de cada habilidad y determinar en que forma difiere la ejecución de una habilidad, en un continuo progresivo de novato a experto. Conforme se desarrolla la pericia, los componentes específicos de la habilidad se vuelven más automáticos, los procedimientos de la tarea se integran a componentes más grandes y la ejecución se hace más rápida y sincronizada. El análisis de la progresión en la habilidad es básico para proporcionar las condiciones para el desarrollo de la misma durante el entrenamiento.

En esta etapa se refina y evalúa la organización del conocimiento desarrollada durante las fases 1 y 2. Adicionalmente se determina la organización del conocimiento de los novatos y de los que poseen un nivel intermedio. Las diferencias entre novatos y expertos proporcionan la base para determinar la

secuenciación instruccional e indican las concepciones erróneas potenciales que ocurren comúnmente en los novatos, cuando intentan resolver un problema estadístico. El conocimiento de los errores de los novatos permite extender la instrucción en algunos conceptos o relaciones conceptuales o pasar por alto otros que son innecesarios.

4ª Fase: Análisis de Datos y Representación del Conocimiento.

Esta fase, agregada al modelo original de análisis de tareas (MATI), constituye un trabajo analítico y sintético que resume los hallazgos de las fases anteriores del modelo MATI. Los productos de esta fase son los siguientes:

- Conocimiento conceptual experto para la elección de la prueba estadística
- Conocimiento procedimental experto para la elección de la prueba estadística.
- Conocimiento estratégico experto para la elección de la prueba estadística.
- Principales errores conceptuales en que incurren los novatos
- Aspectos problemáticos donde se sugiere concentrar la enseñanza de la elección de la prueba estadística.

RESULTADOS

6.1 Primera Fase Orientación: Identificación de las Tareas Principales

La entrevista a expertos en la primera fase, dio cuenta de los aspectos generales involucrados en la toma de decisiones respecto de la prueba estadística más apropiada cuando se enfrentan a un problema específico. A continuación se presenta el análisis de la fase de Orientación, en la que se planteó como objetivo la identificación de las tareas principales que componen la elección de una prueba estadística. Las descripciones se acompañan de algunos extractos de la entrevista que se aplicó a dos expertos.

6.1.1 El procedimiento para la elección de la prueba estadística

Ante la pregunta de qué pasos se siguen para tomar la decisión acerca de qué prueba estadística aplicar, los expertos refiere inmediatamente como primer paso o aspecto a considerar, el nivel de medición de las variables. Sin embargo más adelante afirman que son varios los pasos y que no tiene un orden específico. Aunque ninguno de los dos expertos acepta que exista un orden determinado, enumeran los pasos en el siguiente orden: objetivo de investigación (de comparación, de relación), el nivel de medición, las hipótesis (de relación, de diferencias, de predicción), el instrumento, el muestreo (probabilístico, no probabilístico), el diseño (de una sola muestra, dos muestras, tres muestras, cuatro muestras).

I= Investigador

E2= Experto 2

E5=Experto 5

I: ¿qué pasos sigues para tomar la decisión de qué prueba estadística aplicar?

E5: el nivel de medición de la variable, ¿quieres que te lo anote?

I: sí, por favor.

E5: básicos para mí son varios, y no los tomo como tú dices, cuál primero y cuál después.

I: ¿no tienen un orden?

E5: el conjunto, pero efectivamente, *primero el objetivo de la investigación*. Si *mi objetivo es comparar, comparar si hay relación, el objetivo es básico, ¿sí? Después tomaría el nivel de medición* y obviamente *ligado al objetivo están las hipótesis*. Y es lo mismo, *es de relación, es de diferencias, son de predicción*, son de... ¿sí?, y luego *el nivel de medición de la variable, y realmente para mí está vinculado el instrumento, después el muestreo, si es probabilístico, no probabilístico*, etcétera. Después *el diseño, si es de una sola muestra, dos muestras, tres muestras, cuatro muestras*, etcétera, y básicamente para mí es lo más importante.

I: ¿qué pasos sigues para tomar la decisión de qué prueba estadística aplicar?

E2: Entonces, dependiendo de lo que yo tenga es como empiezo a decidir qué prueba resulta pertinente, con todos estos elementos desde *¿cuáles son mis variables? ¿de qué tipo son? básicamente definiéndolas como independiente/dependiente. ¿en qué escala de medida están? ¿de cuántas muestras estoy hablando?* Y ahora si *¿qué condiciones estadísticas satisfacen mis datos para definir, ya finalmente, cuál es la prueba estadística que más resulta conveniente emplear de acuerdo con los datos que elegí? ¿sí?* Y de esa forma es cómo decido.

Te lo presento así, en esta secuencia, porque es la forma que yo manejo, pero igual puede variar. *Lo que si no puede variar es que identifiques tus variables.*

I: ¿Las variables?

E2: Sí, *lo primero que tienes que identificar son las variables*, de qué tipo de relación estamos hablando y ya de ahí, si quieres, puedes “chechar” *cuántas muestras tienes, de qué tipo son, en qué escala de medida están, qué supuestos se están cumpliendo*, pero tienes que partir primero...

6.1.2 Conocimiento conceptual involucrado en la elección de la prueba estadística

Entre los conocimientos que poseen los expertos vinculados con la elección de la prueba estadística, se mencionan aspectos referentes al método experimental de investigación y al dominio de conceptos sobre probabilidad y medición, relacionados con la estadística. El experto va más lejos aún, al considerar que los conceptos básicos de aritmética representan un conocimiento básico que es necesario dominar adecuadamente como un prerrequisito para el aprendizaje de los conceptos estadísticos.

I: ¿qué conocimientos debes poseer, es decir, para ti qué es lo más importante para decidir qué prueba estadística aplicar?

E5: *los conocimientos se reducen al método, saber metodología de la investigación.*

I: ¿y pensando en la estadística?

E5: *es curva, eso es curva.* Para mí, todo lo anterior que es “*Matemáticas*” es *básico*, pero se supone que todo alumno ya lo sabe, ya lo llevo desde hace muchos años, obviamente si tú ves aquí los puntos importantes que son antecedentes a la estadística, pero no a la investigación para tomar decisiones en la estadística. Sí, en la estadística, obviamente es “*Probabilidad*” y “*Medición*”. Son dos rubros grandes de las matemáticas y obviamente, la *probabilidad está vinculada a la estadística*, pero también está vinculada a las matemáticas, porque empiezas a ver aquí diagramas de “Venn” y todo esto. Y *medición obviamente tiene que ver con la curva, teorema del Límite Central*, todas las pruebas, pero para mí esto, si es un antecedente pero si no lo traes nos vamos a ir hasta las *pruebas aritméticas*. ¿estás de acuerdo? Porque si tú me dices que estos antecedentes, pues yo ya me voy a ir a la *aritmética. Tienen que saber, sumar, dividir, multiplicar, restar, sacar raíz cuadrada*, si no, no puedes llegar a hacer esto para la probabilidad necesitas aritmética. y entonces así me voy a ir hacia atrás. A estas alturas del partido para mí ya no... obviamente quien no sabe esto, difícilmente va a saber metodología y difícilmente va a saber estadística.

I: ¿se puede desglosar esto en algo más fino? Me mencionaste algo referente a la metodología.

E5: sí, es que lo más fino sería esto, bueno lo más fino es ya yo poner toda la clasificación de todo esto que es muchísimo, **objetivo, ver los tipos de objetivos** que existen, los niveles de los objetivos, la taxonomía de Bloom, la nueva, ver **los tipos de hipótesis que existen**, porque **plantear una hipótesis estadística, una de trabajo, una alterna, ver una nula, los cuatro niveles de medición, ¿por qué las variables están asociadas al instrumento?** y tendría yo que hablar de **definición operacional de variables, los instrumentos**, los tipos de instrumentos que hay, si se parecen en términos de **cómo la defines (la variable), los tipos de muestreo**, y cada uno por qué se elige y a qué está asociado, y **el tipo de diseño**, que es muchísimo.

6.1.3 Dificultades para la comprensión de la estadística

Uno de los expertos plantea que la identificación de variables es un aspecto problemático en la comprensión de la estadística, debido a que confunde las variables con las características de los sujetos que permanecen constantes (por ejemplo la edad). Otro de los expertos, refiere la dificultad en la comprensión de la curva (refiriéndose a la distribución normal) y la probabilidad. De acuerdo al experto, éstos son conceptos que los alumnos no logran integrar a un esquema de conocimiento sobre la estadística, debido a que cuando se remiten a los conocimientos en semestres posteriores, los alumnos no pueden vincular estos términos vistos en los semestres anteriores. Sin embargo, uno de los mayores problemas para los alumnos parece ser de tipo actitudinal, se trata de una fuerte aversión a la estadística, por tratarse de una parte de las matemáticas, adquirida desde edades tempranas, en los primeros años de aprendizaje escolar (primaria y secundaria). El experto identifica el concepto de probabilidad como el más difícil de comprender para los alumnos, incluyendo los conceptos asociados, tales como la curva normal y las distribuciones "F" y "T". Por otra parte el contraste de hipótesis no representa un problema para la comprensión de los alumnos. La principal sugerencia para hacer más entendibles los conceptos asociados a probabilidad y distribuciones a los alumnos es utilizar un lenguaje familiar a los alumnos y ejemplos de la vida cotidiana para ejemplificar, además de darles la oportunidad de escoger la temática de las investigaciones que los alumnos van a llevar a cabo, como parte de las prácticas de la materia de estadística.

I: y en tu opinión, ¿cuáles serían los problemas, en estadística que son más complejos, que son más difíciles de comprender?

E2: Qué es una **variable**.

I: ¿Por qué?

E2: Porque a veces resulta que una característica la manejas como una variable. Me ha tocado, ya en mi experiencia, que me dicen mis alumnos: "Quiero determinar si los anuncios televisivos influyen en la autoestima de las mujeres que van de dieciocho a veinte años". ¿Sí? y les dices, cuáles son tus **variables**, ah, pues uno es el anuncio. ¿y el otro cuál es?, no pues la autoestima, ¿y el otro cuál es?, no pues que es la edad. Pues sí, tienes de dieciocho a veinte, pero si la edad es una variable necesita tener un rango más amplio para

saber que tanto le pegan los anuncios a las edades. Entonces, esto más bien es una característica, no una variable.

E5: como que ya lo hago en automático, dependiendo de lo que quieres, cómo quieres, cómo lo mides, cómo lo seleccionas y qué es lo que vas a hacer, yo ya sé que prueba...

I: ¿y en cuanto al aprendizaje?

E5: ¿para los muchachos?... todo

I: ¿qué es lo más difícil?

E: siempre que les preguntas sobre lo que vieron en primer y segundo semestre, no les queda claro, nada de la curva, ni nada de la probabilidad, tienes que volver a empezar.

I: ¿cuáles crees que son los cuellos de botella en cuanto a la enseñanza?, o sea ¿dónde crees que está el problema de que los alumnos no comprendan la estadística?

E5: Bueno, me voy a ver muy mal, porque no es que yo sea muy buena, pero ya cuando están conmigo, como que le entienden, a lo mejor les queda más claro, pero ya como que la manejan, no perfecto, porque además **no la quieren, no quieren la materia**, esa es la otra.

I: ¿hay una aversión digamos?

E5: pero bueno, **es la aversión a la estadística, es la aversión a las matemáticas, a los números**

I: ¿vienen arrastrando una aversión?

E5: pero bueno, **el problema de la educación de las matemáticas, obviamente que está en la primaria, yo creo es la didáctica que se sigue para enseñar.**

I: y en tu opinión **¿cuáles son los conceptos específicos que son más difíciles de comprender?**

E5: **probabilidad, es el más, es terrible para ellos. No les queda claro, hay que hacérselos demasiado simple. Probabilidad y curva, todo lo que tiene que ver con Teorema del Límite Central, la distribución Ji cuadrada, la distribución F, la distribución t, no como van la distribuciones**, entender como están la distribuciones y la probabilidad asociada ahí, porque el contraste de hipótesis para ellos es regalado, ¡y es probabilidad! Sin embargo hay que decirles: "pero tiene esta relación y es por esto, y esto". Hay algunos que me dicen que luego yo lo hago demasiado digerido, y no permite que ellos piensen... pero si se los hago digerido, al rato ya piensan y ya lo tienen.

I: Es decir, ¿primero se los pones de una manera muy simple?

E5: **muy simple, pero aplicada a una investigación, muy, muy simple.**

I: ya después se va complicando más

E5: claro, y así les digo: primero les voy a hablar en español y luego con términos técnicos, de hecho así les digo y les da risa, pero la probabilidad... **la ven demasiada**

abstracta.

I: o sea como que no tiene conexión con lo que ellos están haciendo ahora y tampoco con lo que hacen en su vida ¿no?

E5: y siendo que la probabilidad es diaria, las matemáticas es diaria.

I: o sea de hecho, nosotros de alguna manera aplicamos probabilidad, pero de manera intuitiva, ¿no?

E5: correcto, **cuando dices ah yo creo que va a llover, eso es probabilidad y de hecho así se los pongo yo, a ver, cuando dicen, ¡ay! a lo mejor mi novia no llega hoy, ¡eso es probabilidad!**

I: ¿en qué se están basando para decirlo?

E5: y tu les das los elementos: "¡ya llegó tarde otra vez!", de que siempre hay justificación, de que el metro se retrasó, de que hubo trafico, de que..., si todos esos elementos, **entonces, es la probabilidad de que llegue o no llegue, es lo mismo en la investigación, pero es que hay que hacérselos así de simple, ¿no? entonces así es regalado, muchas veces me han dicho ¡ah! ¿Así de fácil? ¡Pues, sí! ¡Así de fácil!**

6.1.4 Propuestas de los expertos para solucionar los problemas en la enseñanza de la estadística

De acuerdo a la opinión del experto, la enseñanza de la estadística debería ser más aplicada y relacionada con el campo de estudio en que se utiliza. De hecho el experto ha modificado su enfoque de enseñanza de la probabilidad, cambiando los ejemplos clásicos de los libros de texto sobre probabilidad en que se utilizan lanzamientos de dados y monedas, por ejemplos relacionados con la investigación en psicología, que los alumnos van desarrollando, de acuerdo a sus propios intereses personales relacionados con la disciplina. En suma los alumnos tienen problemas para relacionar los conceptos aprendidos acerca de la probabilidad con la estadística. El experto sugiere abordar la probabilidad de una manera más simple, con términos coloquiales, tales como: "la probabilidad sirve para saber si lo que están haciendo es confiable".

I: la forma en que se enseña, **¿cómo crees que se debe enseñar?**

E5: no sé... **más simple, más aplicada, y aplicada a la carrera**, porque eso es un problema también, yo me acuerdo, y ahorita te puedo decir, cuando recibo los grupos en tercero me dicen: ¿pero no nos vas a poner dados o monedas o cartas verdad? en probabilidad.

I: ¿en probabilidad?

E5: correcto, monedas, dados y cartas. **Y les digo: ¡no!, les voy a poner problemas de investigación** que he hecho durante un tiempo, eso les preocupa sobremanera porque no saben dónde aplicarla, pues si lanzo una moneda y claro va a salir águila o sol, ¿pero lanzo a una persona? ¿Y qué me va a salir? pierna, cabeza, brazo o... así me dicen, no lo están viendo aplicado.

I: ¿se tiene que relacionar con lo que están haciendo?

E5: claro, y no lo relacionan

I: ¿es una herramienta para nosotros, no?

E5: y eso no ha sucedido, o no lo sienten así, yo que te puedo decir, también como profesor sé que muchas de las cosas que te digan, siempre el que tiene la culpa es el profesor del semestre pasado, cuando revisas al profesor que les dio el semestre pasado, y lo dio muy bien. Entonces, **yo creo que es una conjunción de varias cosas: la aversión que hay a la materia, la didáctica y que no lo siente aplicado.**

I: ¿si fuera más contextualizado, cómo sería la enseñanza?

E5: **como aplicado a su carrera**, a lo que van a hacer, y como que **darles la oportunidad de que ellos elijan lo que quieran investigar**, para que así vean que la herramienta que están usando sobre lo que ellos decidieron investigar tiene aplicación, y pueden tomar la decisión sin mayor problema.

I: ¿y cómo lo aplicas en clase?

E2: de hecho así doy yo la clase, **a los alumnos le digo: ¿ustedes que quieren investigar?** Pues, yo quiero ¡el movimiento de la rata!, ya ves que desde tercer semestre, según ellos ya van a ser fisiólogos, experimentales, organizacionales, no pues yo quiero ver la selección de personal, **lo que ellos quieran, que se enamoren de su tema y vean que la herramienta es funcional**, pero ellos sienten que no, siempre dicen ¡ay! ¿monedas?, ¿dados?... no, a mi no me interesa.

I: ¿eso se ve en preparatoria o secundaria?

E5: pues es que se los dan en primer y segundo semestre

I: entonces parece como que la probabilidad está desconectada de la estadística

E5: y eso les causa... **ellos no ven asociada la probabilidad y la estadística, y les digo, ¿saben que es la probabilidad, en términos sencillos?**, ellos me dicen: no ¿qué es? **pues a ver si lo que están haciendo realmente es confiable...** ¡así de sencillo! ¿Y cómo lo trabajas? pues eso que aventaban, el dado o la moneda, tiene que ver con esa toma de decisiones, este contraste que hacen, de hipótesis. Entonces, ¿así era de sencillo?, pues... ¡así era de sencillo!

La Tabla 3 muestra una síntesis de los productos de la Fase 1.

Tabla 3.
Productos de la fase 1 Orientación.

	HABILIDADES	CONOCIMIENTOS	MODELO MENTAL
FASE 1: ORIENTACIÓN	<ul style="list-style-type: none"> ▪ Identificar el objetivo de investigación (comparación/relación) ▪ Identificar el nivel de medición de las variables ▪ Identificar el tipo de instrumento utilizado ▪ Identificar el tipo de muestreo ▪ Identificar el Diseño (una o dos o más muestras) 	<ul style="list-style-type: none"> ▪ Probabilidad ▪ Curva Normal ▪ Aritmética ▪ Teorema de Límite Central ▪ Tipos de Instrumentos ▪ Tipos de Diseño de Muestreo 	<ul style="list-style-type: none"> ▪ La elección de la estadística pondera al menos cinco elementos clave en un problema estadístico: 1) el objetivo de investigación 2) el nivel de medición de las variables, 3) el tipo de instrumento utilizado, 4) el tipo de muestreo, y 5) el tipo de diseño

6.2 Segunda Fase: Análisis Básico. Caracterización General de la Tarea y Componentes de la Pericia

Los resultados de esta fase se concentraron en recabar datos de los componentes básicos de la pericia en la solución de los problemas estadísticos. Para recabar la información se aplicaron dos tareas, tanto a los expertos como a los novatos: La tarea de clasificación de problemas estadísticos y la tarea de solución de un problema estadístico. A partir del análisis de la primera y segunda fase se obtuvo un esquema de representación del *Espacio del Problema*, consistente en los conceptos y procedimientos generales involucrados en la selección de la prueba estadística.

6.2.1 Esquema del Espacio del Problema

Con el objeto de obtener un esquema general del espacio del problema, correspondiente a la selección de la prueba estadística se tomaron los aspectos más característicos de la ejecución experta, obtenidos tanto de la entrevista, como de la aplicación de las dos tareas a los expertos. El análisis reveló la siguiente estructura en la selección de la prueba estadística (véase Tabla 4), en la que aparecen cinco pasos o procedimientos, los cuales no tienen un orden específico, debido a que es difícil, de acuerdo al juicio de los expertos, establecer un nivel de importancia: 1) identificación de variables, 2) propósito del estudio/hipótesis (comparación o asociación), 3) nivel de medición de las variables, 3) tamaño de la muestra, 4) cumplimiento de supuestos, y propiamente 5) la decisión estadística, paso que integra la consideración de todos los aspectos en su conjunto, para decidir entre la aplicación de una prueba de tipo paramétrica o no paramétrica.

Tabla 4.
Esquema de aspectos clave considerados para la solución de un problema estadístico por los expertos

Solución de un Problema Estadístico	
I.	Identificación de variables
A.	Variable Independiente
1.	Num. de Grupos o Condiciones
a.	Muestras Independientes o Relacionadas
B.	Variable Dependiente
1.	Tipo de Instrumento
II.	Propósito del Estudio/Hipótesis
A.	Asociación
B.	Comparación
III.	Nivel de Medición de las Variables
IV.	Tamaño de la Muestra
V.	Cumplimiento de Supuestos
A.	Independencia
B.	Homogeneidad de Varianzas
C.	Normalidad
D.	Procedimientos Estadísticos
a.	Normalidad: Shapiro-Wilk
b.	Homogeneidad de varianza: Análisis de Varianza (Levene)
VI	Toma de Decisión
A:	Estadística Paramétrica
▪	Diferencias significativas
•	Nivel de Significancia
•	Valor del Estadístico
▪	¿En qué par están las diferencias?
•	Comparaciones múltiples
B.	Estadística No Paramétrica (Nota: el problema analizado no incluye este tipo de análisis)

En el esquema de la figura 6, se puede observar con más detalle en forma gráfica, cómo se relacionan entre sí los diferentes aspectos involucrados en la decisión de la prueba estadística durante la solución de un problema estadístico. En este esquema no se representan las diferencias entre los expertos, sino las coincidencias en los procedimientos de solución. Es decir, el esquema de la figura 6 representa el modelo experto del espacio del problema, para el caso de ANOVA.

Las diferencias que se observan entre cada nivel de pericia se describen en la tercera fase. Para elaborar el esquema del espacio del problema y los sucesivos esquemas de representación de los expertos y novatos se utilizó la taxonomía de relaciones semánticas elaborada por Lemke, (1997), para mayor detalle ver Apéndice D.

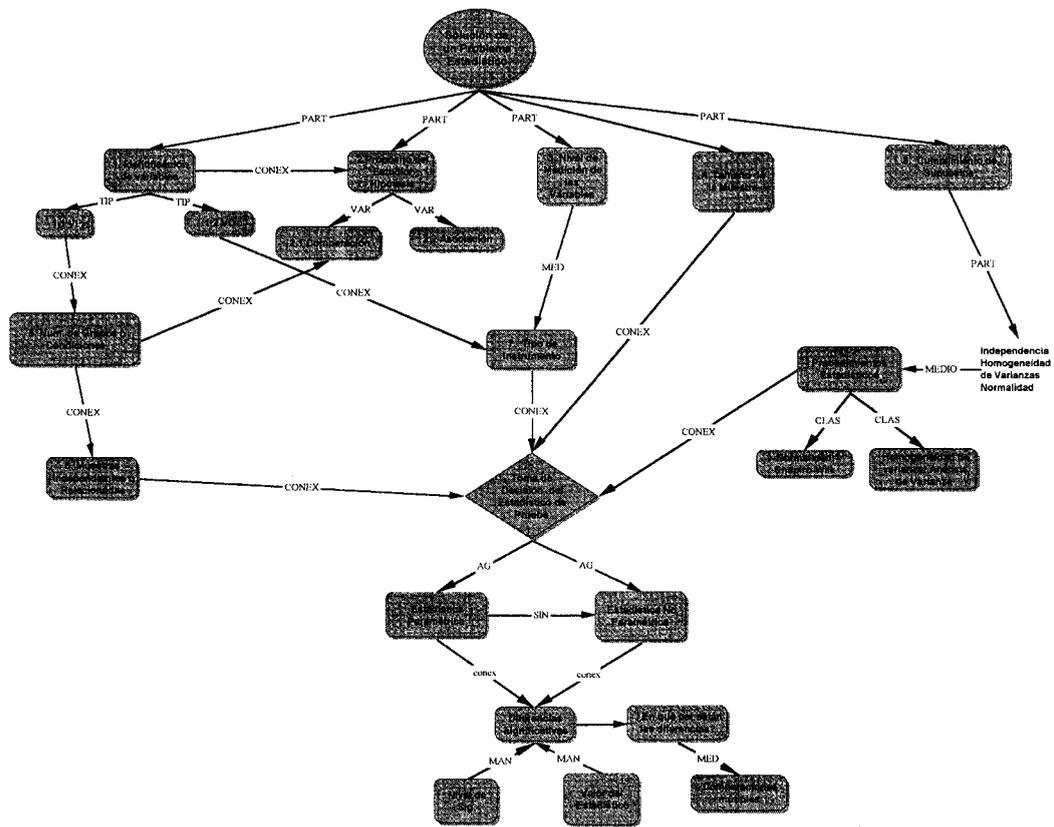


Figura 6. Esquema que representa el modelo experto del espacio del problema

6.2.2 Aspectos clave involucrados en la selección de la prueba estadística

Se analizaron los aspectos clave involucrados en la selección de la prueba estadística y el orden en que fueron mencionados por los expertos. Como se aprecia en la tabla 5, el aspecto clave mencionado en primer lugar fue la Identificación de las variables independiente y dependiente, el segundo lugar en orden de importancia para los expertos lo ocupa el propósito del estudio, el nivel de medición de las variables ocupa el tercer lugar, aunque el experto tres no menciona ninguno de los dos últimos aspectos en su explicación, mientras que el tamaño de la muestra se podría ubicar en cuarto sitio. Más alejado en orden de prioridad que los aspectos anteriores, se ubican el número de grupos o condiciones, el cumplimiento de supuestos para la pruebas paramétricas, el tipo de instrumento utilizado, la decisión de aplicar estadística paramétrica o no paramétrica, y el tipo de estadístico de prueba a utilizar. Llama la atención que sólo dos expertos mencionaron el tipo de muestras independientes o relacionadas cómo aspecto importante, a pesar de ser un aspecto esencial en la selección de la prueba estadística. Por otra parte sólo un experto mencionó el tipo de muestreo como un aspecto clave a considerar, sin embargo para los demás expertos no fue relevante en su decisión. En resumen, se puede afirmar que, aunque la mayoría de los expertos coinciden en los principales aspectos que ellos consideran relevantes para la toma de decisión de la prueba estadística, existe una importante variabilidad entre los expertos, con respecto al énfasis o grado de importancia que ellos le asignan a cada uno de los aspectos clave para lograr una decisión exitosa. Sin embargo, a pesar de la variabilidad de criterios, es posible apreciar que los cinco aspectos claves, de acuerdo a los expertos, son la identificación de variables, el propósito u objetivo del estudio, el nivel de medición de las variables, el tamaño de la muestra y el número de grupos o condiciones, respectivamente.

Tabla 5.

Orden de importancia, asignado por los expertos, a los aspectos clave involucrados en la selección de la prueba estadística

Aspecto Clave	Experto 1	Experto 2	Experto 3	Experto 4	Experto 5
Identificación de Variables (VI, VD)	2°	1°	1°	1°	2°
Propósito del Estudio	1°	3°	No mencionado	3°	1°
Nivel de Medición de las Variables	3°	2°	No mencionado	4°	3°
Tamaño de la Muestra	4°	7°	4°	2°	4°
Número de Grupos o Condiciones	5°	4°	2°	5°	5°
Cumplimiento de Supuestos	6°	8°	5°	7°	No mencionado
Tipo de Instrumento Utilizado	No mencionado	5°	3°		7°

Paramétrica o No Paramétrica	7°	No mencionado	6°	6°	8°
Tipo de Estadístico a Utilizar	8°	8°	No mencionado	7°	9°
Muestras Independientes o Relacionadas	No mencionado	6°	No mencionado	No mencionado	6°
Tipo de Muestreo (Probabilístico/No Probabilístico)	No mencionado	No mencionado	No mencionado	No mencionado	5°

En la Tabla 6 se muestra una síntesis de los productos obtenidos de la Fase 2.

Tabla 6.

Productos de la Fase 2 Análisis Básico

	HABILIDADES	CONOCIMIENTOS	MODELO MENTAL
FASE 2: ANÁLISIS BÁSICO	<ul style="list-style-type: none"> ▪ Identificar el objetivo de investigación (comparación/relación) ▪ Identificar el nivel de medición de las variables ▪ Identificar el tipo de instrumento utilizado ▪ Identificar el tipo de muestreo ▪ Identificar el Diseño (una o dos o más muestras) ▪ Aplicar procedimientos estadísticos para identificar normalidad, homogeneidad de varianzas e independencia. ▪ Aplicar el procedimiento para tomar la decisión estadística (aceptar o rechazar la hipótesis nula). 	<ul style="list-style-type: none"> ▪ Tipos de Instrumentos ▪ Tipos de Diseño ▪ Tipos de Muestreo ▪ Tamaño de la Muestra ▪ Supuestos que deben cubrir las diferentes pruebas estadísticas ▪ Pasos para tomar la decisión estadística 	<ul style="list-style-type: none"> ▪ La elección de la estadística se lleva a cabo considerando: cinco pasos o procedimientos: 1) identificación de variables, 2) propósito del estudio/hipótesis (comparación o asociación), 3) nivel de medición de las variables, 3) tamaño de la muestra, 4) cumplimiento de supuestos, y propiamente 5) la decisión estadística. ▪ El modelo experto se representa en el esquema de la figura 6.

6.3 Tercera Fase: Refinamiento del Análisis. Caracterización de la Progresión Novato-Experto

En esta fase el análisis reveló las diferencias más finas que se observan entre los diferentes niveles de pericia. Se describen a continuación en forma detallada los resultados de las tareas de clasificación de problemas y solución de problemas estadísticos.

6.3.1 Tarea de clasificación de problemas estadísticos

Se elaboró una tabla que contiene las características de los doce problemas planteados a los sujetos entrevistados con el objeto de tener una visión general del tipo de problemas presentados (véase Apéndice D), para más detalles sobre el formato de los problemas consultar el Apéndice B. Con el objeto de hacer más sintética la información presentada, se muestra únicamente un ejemplo representativo en cada nivel de la categoría de novatos.

6.3.1.1 Identificación de los conceptos clave para la clasificación de problemas estadísticos

La tarea de identificación de conceptos clave se aplicó a 4 expertos y 10 sujetos novatos (en sus tres niveles) y se realizaron dos tipos de análisis: A) la categorización de los conceptos clave subrayados en cada problema y B) la estructura de la clasificación de los doce problemas presentados.

A) Análisis de la categorización de los conceptos clave

Para el análisis de los conceptos clave presentados, los conceptos subrayados por los sujetos se clasificaron en tres categorías: basados en principios estadísticos (P), moderadamente basados en principios estadísticos (MP) y basados en conceptos superficiales (S). Asimismo se tomaron en cuenta las anotaciones que hicieron los sujetos sobre el texto en el momento de clasificarlos. En la tabla 7 se muestran los criterios de categorización y los tipos de respuesta posibles.

Tabla 7.

Categorías de conceptos clave subrayados, criterios de categorización y los tipos de respuesta posibles

Categoría	Criterios	Tipo de respuesta
(P) Basada en Principios	<ul style="list-style-type: none"> Combinación de dos o más características 	<ul style="list-style-type: none"> Propósito del estudio (diferencias vs. Relaciones) Identificación de variables (VI y VD) Nivel de medición de las variables (nominal, ordinal, intervalar) Número de condiciones o grupos (uno, dos, tres ó más)
(MP) Moderadamente basada en principios	<ul style="list-style-type: none"> Sólo una de las características 	<ul style="list-style-type: none"> Los mismos tipos de respuesta de la categoría basada en principios.
(S) Superficial	<ul style="list-style-type: none"> Características 	<ul style="list-style-type: none"> Organización de los datos

	estadísticas directamente vinculadas con la elección de la prueba estadística	no	<ul style="list-style-type: none"> • Recolección de los datos • Tipo de muestreo • Asignación al azar • Muestreo no probabilístico
	<ul style="list-style-type: none"> • Características estadísticas 	no	<ul style="list-style-type: none"> • Temática del problema (alcohol, psicoterapia, etc.) • Disciplina de estudio (psicología, medicina, política, etc.)

Los resultados mostraron que, los expertos calificaron más alto en las categorías basadas en principios ($P=88.5\%$) y moderadamente basadas en principios ($MP=11.5\%$), véase Tabla 8 y Figura 7; en tanto que en la categoría superficial (S) obtuvieron un puntaje de 0%. Lo cual significa que los expertos identificaron en la mayoría de los problemas presentados los conceptos clave para tomar la decisión de clasificar los problemas apoyándose en la identificación de las variables (VI y VD), el propósito de la prueba (identificar diferencias entre grupos o establecer relaciones entre variables), el nivel de medición de las variables y el número de grupos o condiciones.

En el caso de los novatos, destaca la variabilidad en los datos. Resulta lógico pensar, que los alumnos con un nivel mayor de escolaridad y por tanto, mayor experiencia con los contenidos de estadística (por ejemplo el grupo Nivel 1 Doctorado) tendrían un puntaje mayor en las categorías $P=70.9\%$ y $MP=20.8\%$, y un puntaje menor en la categoría $S=8.3\%$, sin embargo, como se observa en la Tabla 8 y Figura 7, los sujetos del nivel 1 Maestría, identificaron en un porcentaje ligeramente mayor los conceptos basados en principios ($P=87.5\%$) que los alumnos del nivel 1 Doctorado. De la misma forma, no se encuentra una diferencia amplia entre los sujetos del nivel 2 (sujetos en proceso de titulación) y los del nivel 1 (alumnos de los 3° a 5° semestre de licenciatura). Cabe destacar que los novatos Nivel 1 Maestría obtuvieron el mayor puntaje en la categoría de clasificación (P) entre los grupos de novatos, con un porcentaje muy cercano a los expertos (87.5%), lo cual se podría deber a la experiencia reciente que tienen en la elaboración de su proyecto de tesis, la preparación para la presentación del examen de ingreso a la maestría, además de los cursos de metodología de investigación que cursan actualmente en la residencia de maestría en psicología (ver Tabla 8 y Figura 7).

Tabla 8

Porcentaje medio de la clasificación de los conceptos clave en tres categorías (basada en principios, moderadamente basada en principios y superficial) en cada nivel de pericia

Grupo de Pericia	Basada en Principios (P) %	Moderadamente Basada en Principios (MP) %	Superficial (S) %
Expertos	88.5	11.5	0
Novatos N1 (Doctorado)	70.9	20.8	8.3
Novatos N1 (Maestría)	87.5	12.5	0
Novatos N2 (Tesis)	61.1	33.4	5.5
Novatos N3 (3°-5° sem. De Lic.)	62.4	23.2	14.4

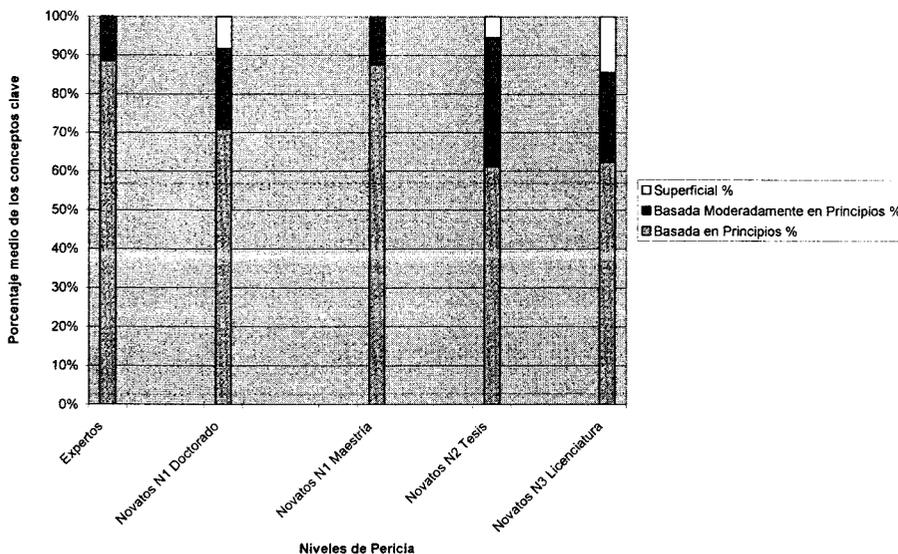


Figura 7. Porcentaje promedio de los conceptos clave identificados por los expertos y novatos para clasificar los problemas estadísticos, en tres categorías (basada en principios, moderadamente basada en principios y superficial), para cada nivel de pericia.

B) Análisis de la estructura de la clasificación de los problemas

Se elaboraron esquemas representativos de la clasificación de los doce problemas estadísticos por cada sujeto entrevistado. Esta clasificación reveló la estructura de la representación de los problemas estadísticos. A continuación se presentan ejemplos representativos de los expertos y los novatos por cada nivel.

B1) Estructura de la representación de los problemas estadísticos por parte de los Expertos

Para dar cuenta de la clasificación utilizada por los expertos, se elaboró un esquema tipo árbol de decisión. El análisis de estas estructuras reveló diferencias importantes entre los diferentes expertos con respecto a la clasificación de los problemas, lo cual se presume que tiene relación con diferentes formas en que se representan los problemas estadísticos, los expertos en la primera fase de la resolución del problema. Por ejemplo, el Experto 1 elaboró dos clasificaciones, en su primera aproximación (véase Figura 8), utilizó un criterio basado en la metodología de investigación, identificando dos grupos principales: a) problemas agrupados de acuerdo a los diseños experimentales que utilizaron, y b) problemas agrupados de acuerdo con el propósito de la investigación y el tipo de estudio (en este caso particular se consideraron los problemas cuyo propósito era buscar relaciones y además son diseños de tipo ex-post-facto). Posteriormente establece subcategorías de acuerdo a las subclasificaciones de los diseños experimentales (diseños cuasi-experimentales y experimentales puros).

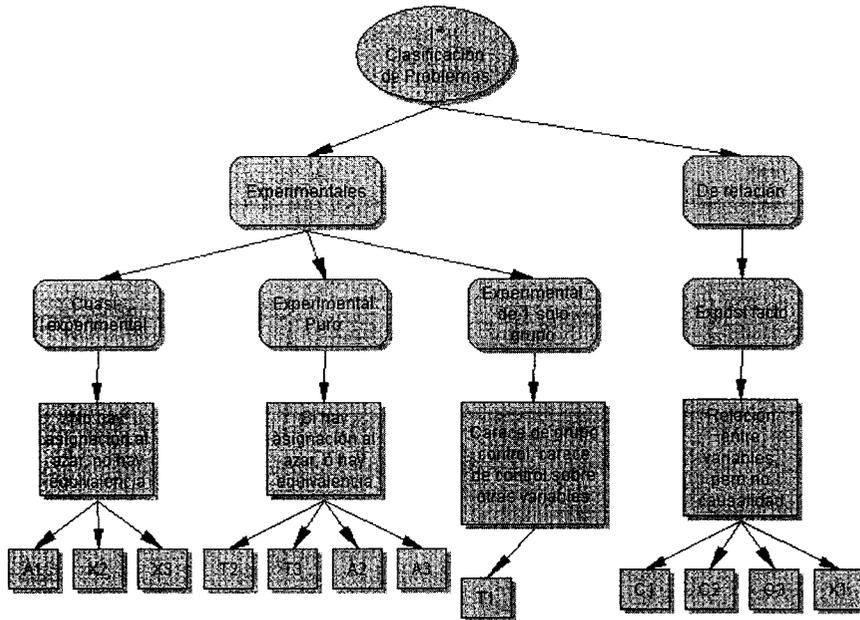


Figura 8. 1ª Clasificación de los problemas estadísticos del experto 1.

El mismo experto hace una segunda clasificación (ver Figura 9) que ahora se fundamenta en criterios estadísticos, en la cual agrupa los problemas por tipo de estadístico a utilizar (Correlación de Pearson, Correlación de Spearman, Ji Cuadrada, t de Student, t de Student para dependientes, ANOVA).

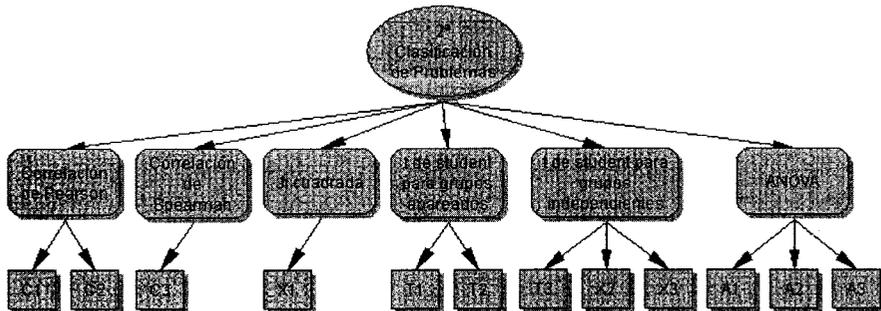


Figura 9. 2ª Clasificación de los problemas estadísticos del experto 1.

En el caso del experto 2, en la figura 10 se observa, una clasificación basada como único criterio en el propósito de investigación del problema (establecer diferencias, relaciones, ó bien causa-efecto). A partir del propósito del estudio deriva tres grupos, a partir de los cuales, se agrupan los diferentes problemas. Esta clasificación le asigna una importancia preponderante a este criterio para entender los diferentes tipos de problemas que visualiza un experto en estadística.

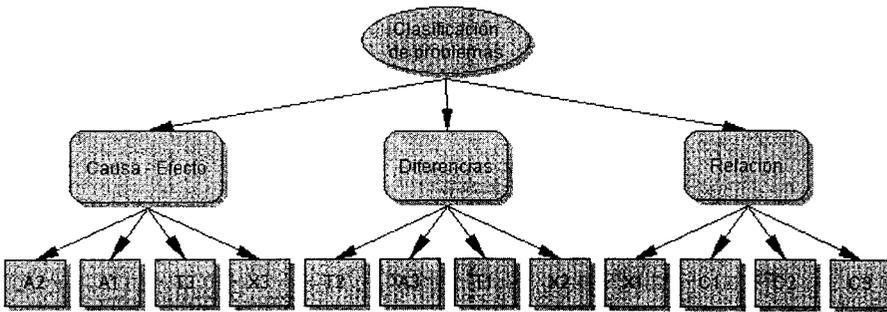


Figura 10. Clasificación de los problemas estadísticos del experto 2.

Por su parte, el experto 3, clasifica los problemas en primera instancia, en muestras independientes, muestras relacionadas (prueba “t” o Wilcoxon), relaciones y prueba t de 2 muestras independientes). Como segundo criterio utiliza la prueba estadística a utilizar y la clasificación de pruebas paramétricas y no paramétricas. Sin embargo, se aprecia un problema de inclusión en clase, debido a que agrupa en dos categorías diferentes a los problemas que se refieren a muestras independientes (ver Figura 11).

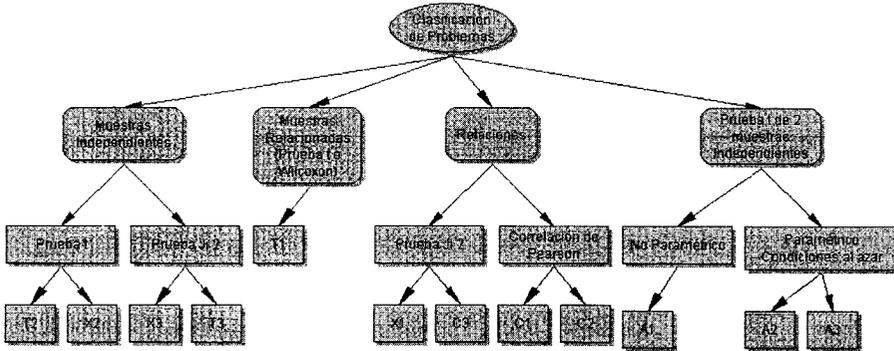


Figura 11. Clasificación de los problemas estadísticos del experto 3.

A diferencia de los 3 expertos anteriores, el experto 4, agrupa los problemas en paramétricos y no paramétricos. En un segundo nivel de clasificación se basa en el número de condiciones o grupos formados y algunas suposiciones sobre posibles pruebas estadísticas a utilizar (ver Figura 12).

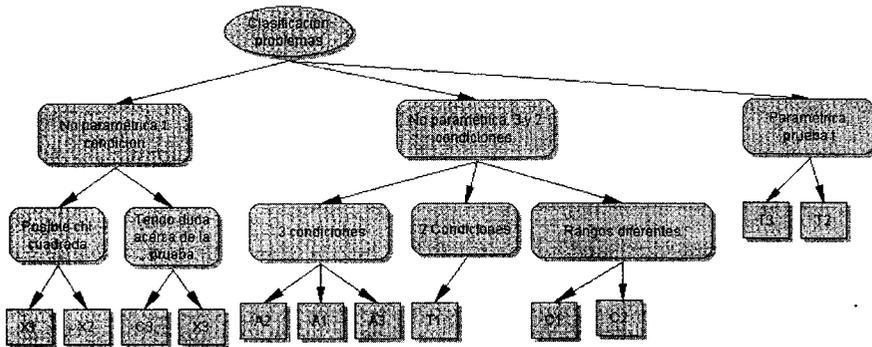


Figura 12. Clasificación de los problemas estadísticos del experto 4.

B2) Estructura de la representación de los problemas estadísticos por parte de los Novatos

Nivel 1a, Doctorado

En este caso, se aprecia que la clasificación que realizan los alumnos de doctorado sigue una estructura similar a la de los expertos, en el sentido de que agrupa los problemas tomando en consideración en primer lugar el propósito del estudio, considerando los que se refieren a diferencias entre grupos de una misma muestra (pre-post), los de diferencias entre grupos (con al menos dos grupos) y los problemas que pretenden establecer relaciones entre variables. Como segundo criterio se apoya en el nivel de medición de las variables dependientes para establecer una subclasificación. Sin embargo, esta clasificación no logra hacer alguna suposición de alguna prueba estadística, ni separar los problemas que tienen tres grupos o condiciones. En suma, se puede afirmar que la forma de representación de los problemas estadísticos en el nivel 1ª, presenta rasgos similares a las formas de representación de los expertos (ver Figura 13). Sin embargo, los alumnos en este nivel fallan en hacer una distinción específica con respecto a las pruebas estadísticas.

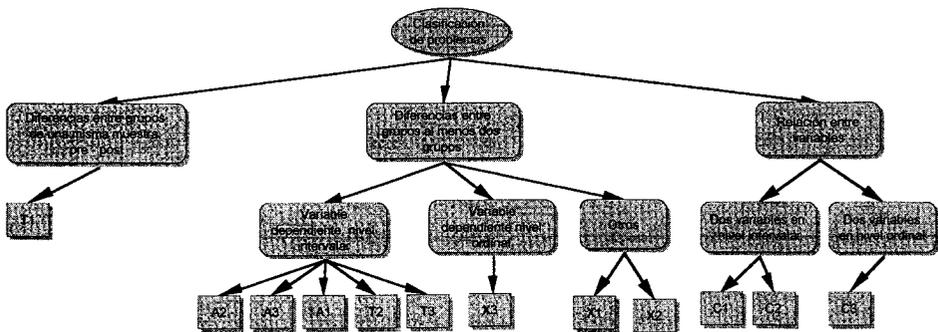


Figura 13. Ejemplo de clasificación de los problemas estadísticos del nivel 1a, doctorado.

Nivel 1b, Maestría

Por lo que respecta a los alumnos de maestría, se observa en el esquema de la figura 14, una clasificación muy interesante de los problemas estadísticos. De la misma manera que en el caso del alumno de doctorado, el alumno de maestría establece claramente una diferencia entre los problemas, de acuerdo al propósito del estudio, pero adicionalmente establece una diferencia entre los problemas que presentan dos grupos o condiciones, de los problemas que presentan más de dos grupos. Como segundo criterio de clasificación, se refiere a la característica de los problemas (cuantitativos o cualitativos), sin embargo en la categoría “más de dos grupos diferentes” incluye el término “ordinal”, para identificar el nivel de medición de las variables relacionadas con el problema. Lo anterior muestra una confusión con los conceptos para referirse al tipo de problemas estadísticos.

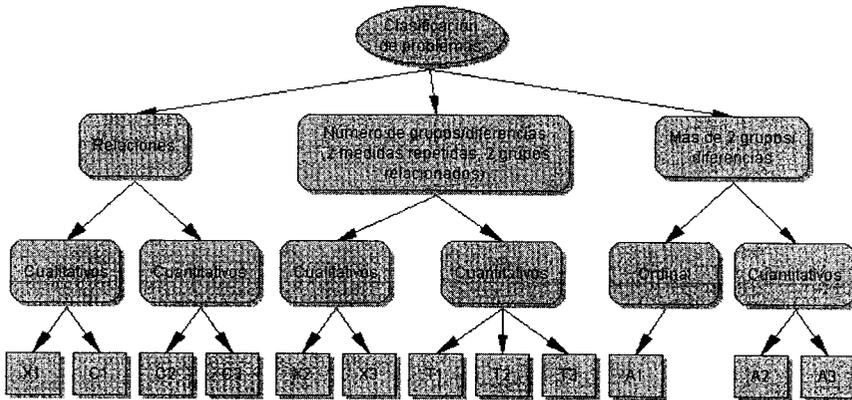


Figura 14. Ejemplo de clasificación de los problemas estadísticos del nivel 1b, maestría.

Nivel 2, alumnos en proceso de elaboración de tesis de licenciatura

Con respecto a los alumnos que han cubierto el 100 por ciento de créditos, es decir, que han cubierto todas las materias del plan de estudio relacionadas con estadística y que se encuentran actualmente en proceso de elaboración de tesis de licenciatura, la clasificación que realizan se basa en criterios poco definidos, debido a que combinan los criterios relacionados con el propósito del estudio (relación entre variables), con criterios relacionados con la planeación de la investigación (planteamiento del problema, establecimiento de hipótesis) y aspectos asociados con el diseño del estudio pre-post (ver Figura 15). En un segundo momento, se subclasifican los problemas del grupo “relación entre variables” en las categorías “comparación de grupos”, los “relación entre variables” y los que se refieren a una temática en particular (identificar el mejor método). Lo anterior indica que existe una confusión con respecto al criterio de inclusión en clase, al colocar en el mismo subgrupo a los problemas con un propósito de comparación entre grupos y los que buscan relación entre variables, en una

misma posición jerárquica se encuentra un criterio superficial (identificar qué método es mejor).

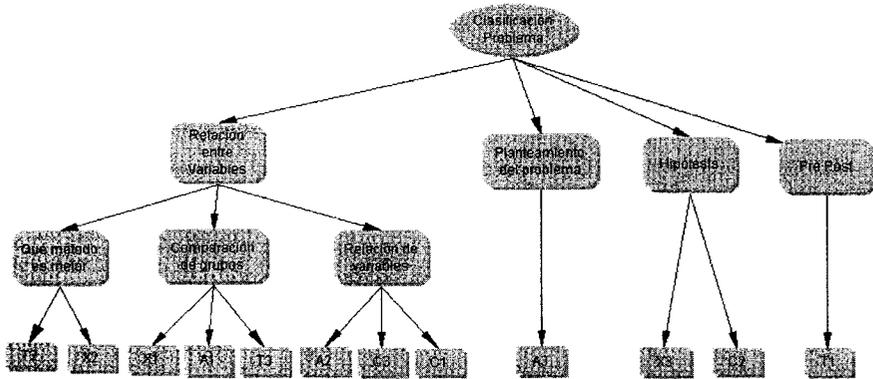


Figura 15. Ejemplo de clasificación de los problemas estadísticos del nivel 2, tesis.

Nivel 3, alumnos de 3º a 5 semestre de licenciatura

Por su parte, los alumnos que cursan los primeros semestres de la carrera de psicología, y por lo tanto los primeros semestres de estadística descriptiva e inferencial, basan su clasificación básicamente en criterios superficiales (como la aplicación de pruebas, aplicación de estímulos y los beneficios en aplicar ciertas condiciones). Por otra parte, no existe dificultad en identificar los problemas de relación y los problemas que incluyen a los mismos sujetos evaluados antes y después. Es de notarse que no utilizan los conceptos usuales en estadística, como “relación entre variables”, “diferencias entre grupos” “muestras independientes”, etcétera (véase Figura 16).

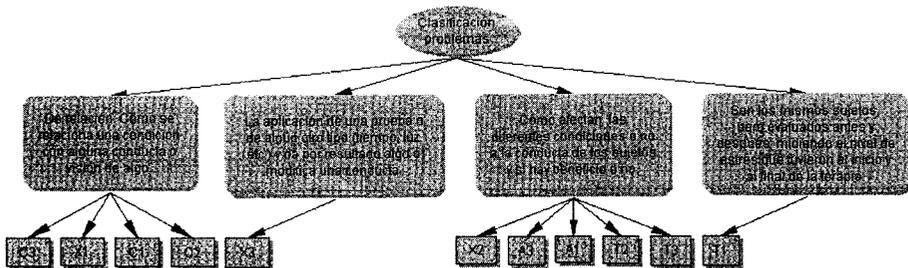


Figura 16. Ejemplo de clasificación de los problemas estadísticos del nivel 3, 3º a 5º semestre de licenciatura.

6.3.2 Tarea de solución de un problema estadístico

Para el análisis de la tarea de solución de problemas estadísticos se solicitó a los 5 expertos y a los 10 novatos seleccionados que intentaran describir los pasos que seguirían durante la solución de un problema de ANOVA y que pensarán en voz

alta durante el proceso de solución, hasta llegar a la decisión de la posible prueba estadística a aplicar, excluyendo el procedimiento de cálculo del estadístico (véase Apéndices C y D). Los protocolos verbales se segmentaron en unidades que correspondan con enunciados que representen una idea particular. Se excluyeron los segmentos de protocolo que se refieren a una segunda lectura del problema; alguna conversación llevada a cabo con el experimentador no relacionada con el problema estadístico; o bien, cuando se tratase de una respuesta a un requerimiento del experimentador (por ej. hablar más fuerte). Con los protocolos verbales de los expertos y novatos se elaboraron árboles de decisión del proceso de solución de cada uno de los expertos, identificando las diferentes formas de abordar el problema. En el caso de los novatos, se realizó un esquema de solución para cada nivel de pericia, estableciendo una comparación entre el espacio del problema experto y las soluciones planteadas en los diferentes grados de pericia.

A continuación se presentan los resultados tanto de los 5 expertos como de los novatos en los tres niveles.

6.3.2.1 Secuencia de solución de los expertos

Es importante notar que la manera de abordar un problema específico varía en cada experto, algunos hacen hincapié en ciertos aspectos clave, mientras que otros enfatizan diferentes aspectos. Todos los expertos coinciden en que la prueba más apropiada para el problema es ANOVA, sin embargo algunos expertos difieren en cuanto al criterio del tamaño de la muestra, al considerar que se trata de una muestra demasiado pequeña (7 sujetos en cada grupo). En este punto algunos expertos sugieren aumentar el tamaño de la muestra, en tanto que otros prefieren aplicar una prueba estadística no paramétrica, tal como Kruskal-Wallis (considerada como ANOVA no paramétrica). En seguida se muestran los esquemas que representan la secuencia de solución de cada uno de los expertos entrevistados.

Experto 1

El experto 1 inicia identificando el propósito del problema como uno de comparación de grupos, a continuación identifica la variable independiente y dependiente (ver Figura 17). Otro aspecto que es relevante para el experto 1 es que se trata de un problema de tres condiciones y la asignación de sujetos a los grupos fue aleatoria. Sin embargo, detecta un problema con el tamaño de la muestra, el cual considera muy reducido, por lo que propone aumentar la muestra en cada grupo ó bien, aplicar una prueba no paramétrica (prueba de Fisher). Posteriormente se enfoca en hacer un análisis de la distribución de las variables para verificar que se cumplen los supuestos de normalidad e igualdad entre varianzas, condiciones necesarias para la aplicación de una prueba paramétrica, tal como ANOVA de 1 vía. Finalmente propone realizar una prueba de Tukey para verificar cuáles son los grupos que difieren, en caso de aplicar un Análisis de Varianza.

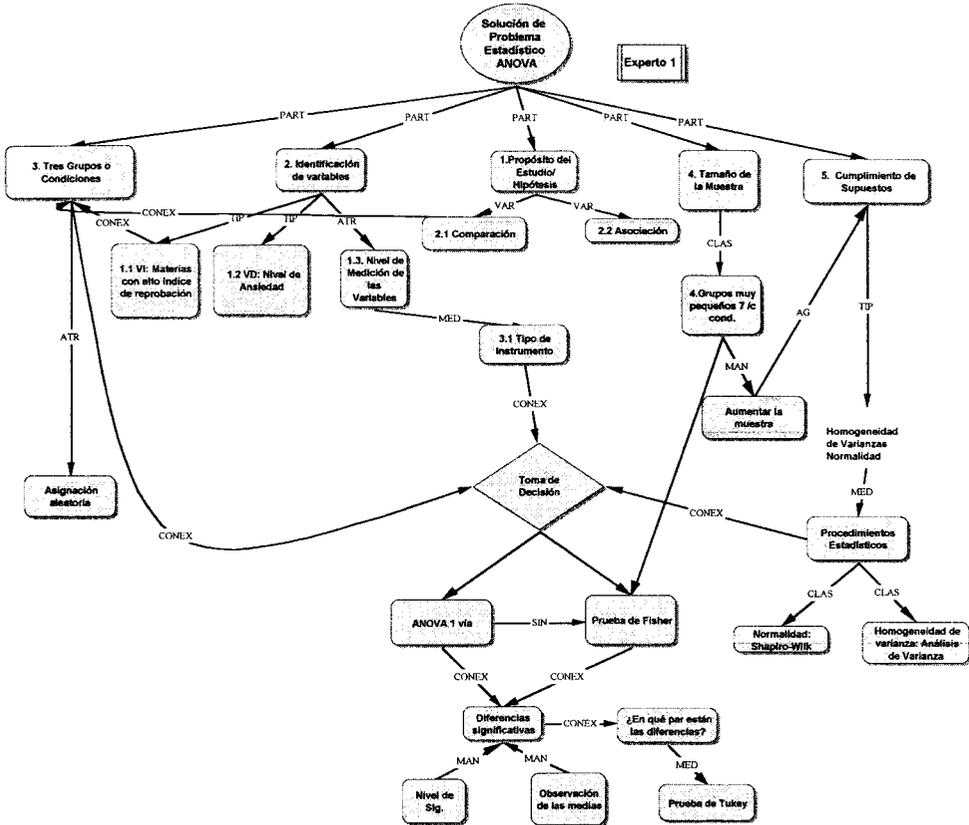


Figura 17. Secuencia de solución del experto 1, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Experto 2

El experto 2 considera como primer aspecto clave para solucionar su problema, la identificación de las variables y la identificación del nivel de medición en que se encuentra cada una (ver Figura 18). En segundo lugar, determina el propósito del estudio, el cual intenta comparar aspectos de la variable independiente (materias con alto índice de reprobación), esto le lleva a identificar tres grupos y que se trata de muestras independientes. En cuanto a la variable dependiente, el aspecto que resalta es el instrumento utilizado (escala Likert), por lo que asume una escala intervalar en cuanto a la suma de puntuaciones. Estos argumentos le llevan a considerar la posibilidad de aplicar la prueba estadística ANOVA de 1 Vía. Para proceder a la aplicación de la prueba propuesta se deben cumplir con ciertos requisitos o supuestos de independencia, normalidad y homogeneidad de varianzas. Para tal efecto se pueden aplicar procedimientos estadísticos (correspondientes al análisis exploratorio de datos) como Shapiro-Wilk. Una vez aplicada la prueba de ANOVA de una vía, en el caso de encontrar diferencias

significativas, es necesario realizar una prueba de comparaciones múltiples para identificar en qué grupo se encuentran las diferencias.

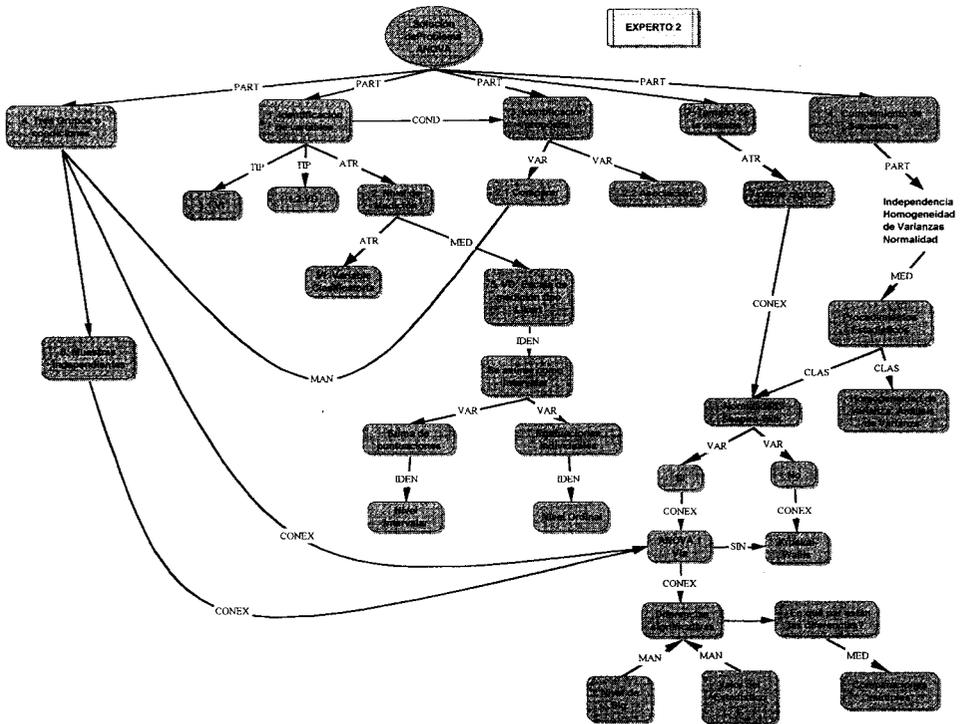


Figura 18. Secuencia de solución del experto 2, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Experto 3

El caso del proceso de solución del problema por parte del experto 3, (ver Figura 19), se puede considerar como una versión simplificada del rastro de solución que siguió el experto número 2, excluyendo el criterio de la identificación del propósito del estudio (relación entre variables ó diferencias entre grupos).

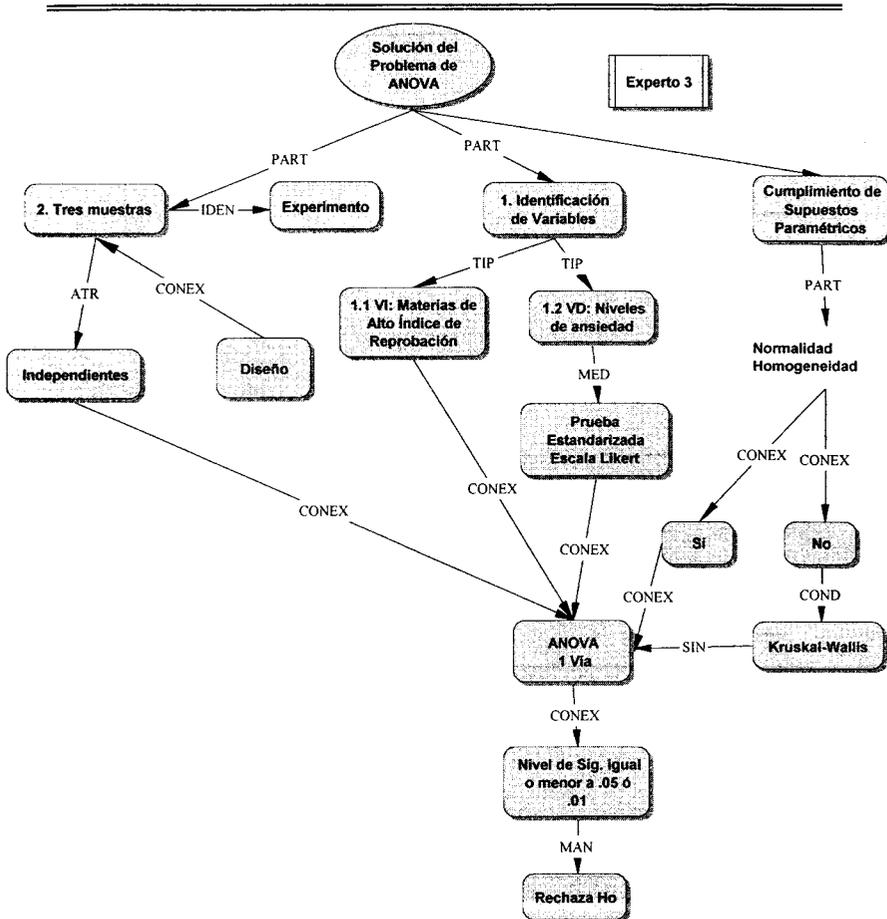


Figura 19. Secuencia de solución del experto 3, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Experto 4

El esquema que se representa en la Figura 20, nos muestra un rastro de solución un tanto diferente por parte del experto 4, en virtud de que se enfoca principalmente en consideraciones relacionadas con las características de la muestra (la representatividad y el tamaño pequeño de la muestra). Otro aspecto importante para el experto 4 es el planteamiento de la hipótesis, que incluye la identificación de variables, el nivel de medición y el número de observaciones o mediciones. El experto considera que el tamaño de la muestra tan pequeño lo conduce a inferir que no existe una distribución normal, por lo tanto considera que lo más apropiado es considerar una prueba estadística no paramétrica, sin especificar cuál. Tampoco menciona en su descripción la posibilidad de aplicar algún análisis de tipo exploratorio.

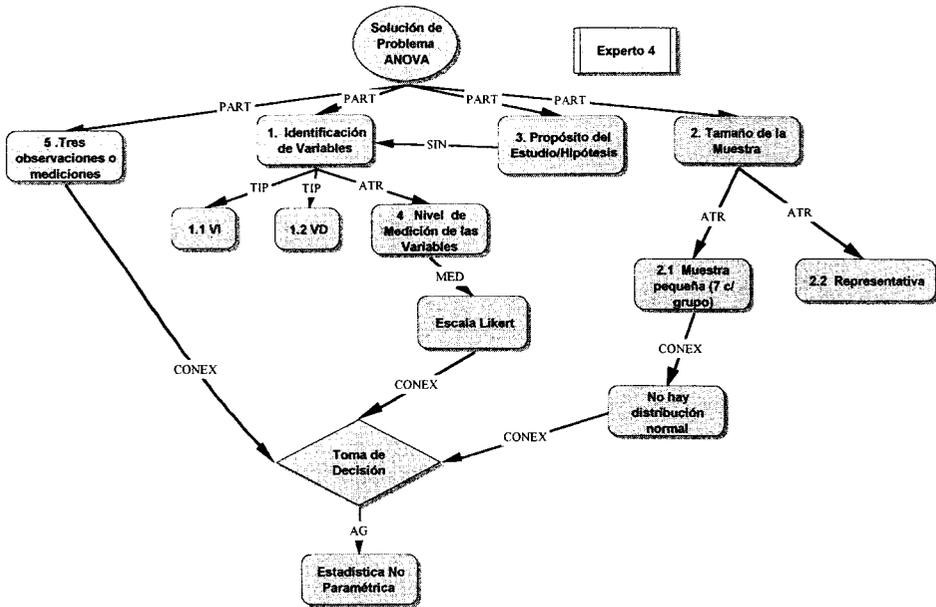


Figura 20. Secuencia de solución del experto 4, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Experto 5

El experto 5 describe con detalle un procedimiento que contempla los aspectos clave mencionados por los expertos anteriores (ver Figura 21). Cabe destacar que el experto agrega el factor del tipo de muestreo (probabilístico o no probabilístico). Además de considerar en primera instancia, al igual que el experto 4, la aplicación de la prueba no paramétrica Kruskal-Wallis, debido al tamaño pequeño de los grupos (7 sujetos por grupo), sin considerar la posibilidad de aplicar algún análisis de normalidad u homogeneidad de varianzas.

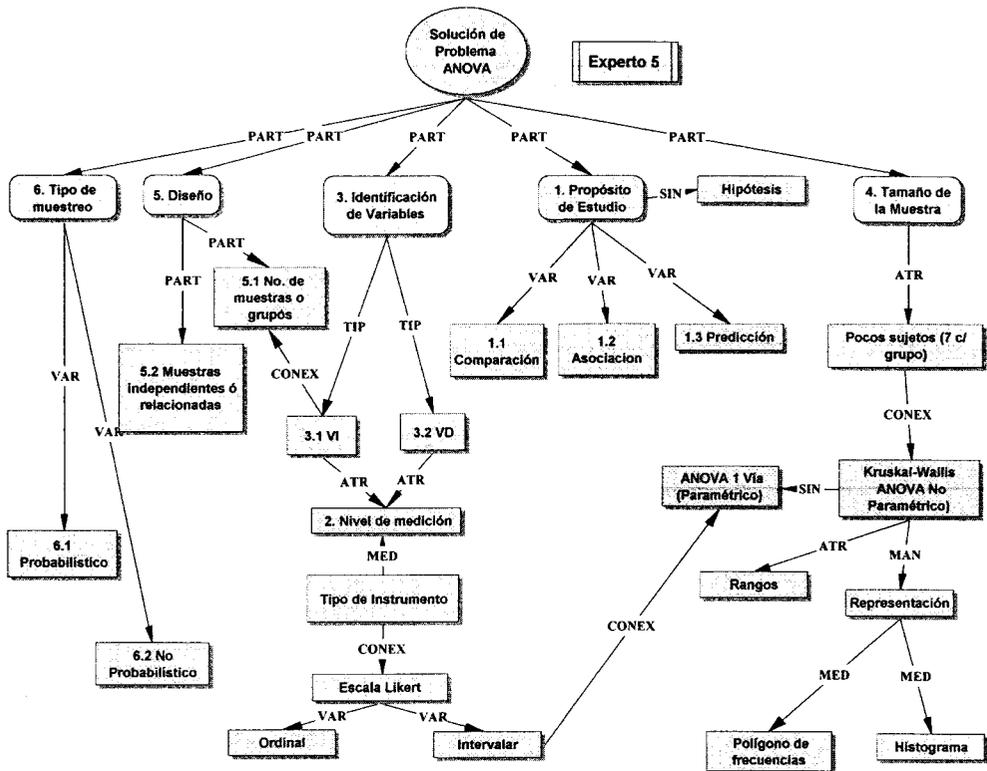


Figura 21. Secuencia de solución del experto 5, de un problema estadístico que implica la aplicación del estadístico ANOVA.

¿Cómo representan los expertos los problemas estadísticos?

En suma podemos afirmar que los expertos coinciden en identificar ciertos aspectos clave al solucionar un problema estadístico específico. Entre los aspectos que se consideran, se encuentran: el propósito u objetivo del estudio (relación entre variables ó diferencias entre grupos), la identificación de variables (VI y VD), así como su nivel de medición, el cual está asociado al tipo de instrumento utilizado y las características de construcción del mismo. Otro aspecto relevante para los expertos es el tipo de diseño, el cual incluye el número de muestras o grupos asociados a la variable independiente y si se trata de muestras independientes o relacionadas. La consideración que sigue es la aplicación de procedimientos apropiados par verificar los supuestos requeridos para las pruebas paramétricas (como normalidad, independencia y homogeneidad de varianzas), antes de proceder a aplicar determinado estadístico de prueba. Un aspecto que también incluyen es la aplicación de procedimientos de comparaciones múltiples para confirmar a favor de qué grupos se encuentran las diferencias (en el caso específico del ANOVA de 1 Vía).

Por otra parte, se observan importantes diferencias entre los expertos, que se relacionan con el tamaño de la muestra, ya que mientras unos expertos consideran la necesidad de aplicar análisis de normalidad y homogeneidad de varianzas, otros dan por sentado que la distribución de una muestra pequeña viola los principios de normalidad y por lo tanto tienden a considerar más apropiadas las pruebas no paramétricas. Cabe hacer notar que los ejemplos mostrados a los expertos y novatos entrevistados fueron extraídos de libros de texto de estadística y en casi todos los ejemplos de ANOVA de 1 Vía encontrados se presentan datos con muestras reducidas, sin incluir ninguna nota aclaratoria al respecto.

Cabe hacer notar que los expertos entrevistados muestran en sus verbalizaciones, durante la solución del problema, escasas referencias a criterios puramente estadísticos, tales como $x = 0$, como se pueden observar en los libros de texto utilizados para apoyar las clases de nivel licenciatura. En su lugar pasan una gran parte del tiempo tratando de extraer la información clave del problema que les permita identificar de qué clase de problema se trata y si cumple con los criterios de una clase de problema específico. Es de notarse que los expertos no utilizan la notación estadística durante la fase de comprensión y representación del problema también característicos de los libros de texto de estadística.

6.3.2.2 Secuencia de solución de los novatos

Doctorado. Nivel 1.

Como se puede apreciar en la Figura 22, correspondiente a un alumno de doctorado, su procedimiento para resolver un problema estadístico se acerca mucho a los procedimientos mostrados por los expertos, al identificar los aspectos claves: número de grupos o mediciones, el propósito del estudio (comparación), identificación de variables (VI y VD) y el nivel de medición de la variable dependiente (intervalar), determinado por la escala Likert utilizada.

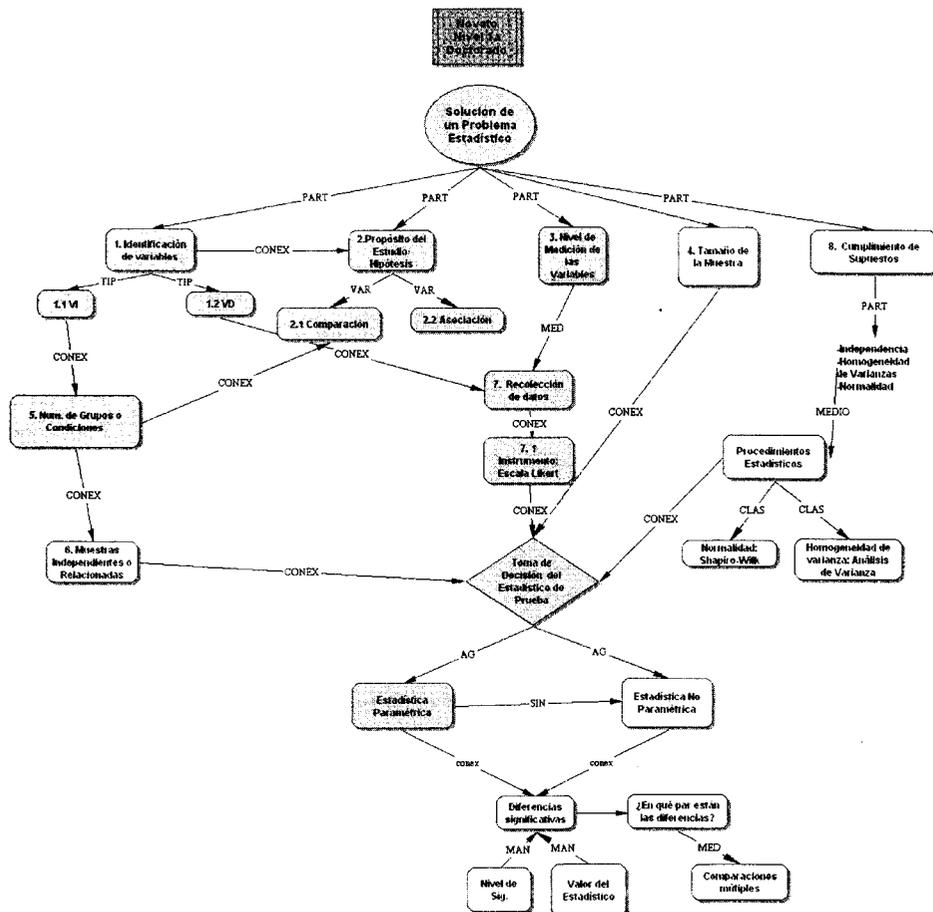


Figura 22. Secuencia de solución del novato nivel 1, doctorado, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Maestría. Nivel 1

En el caso del alumno de maestría, el esquema de la Figura 23, revela un rastro igualmente similar al alumno de doctorado, identificando como elementos clave para tomar su decisión la identificación de las variables (VI y VD). Asociar la variable independiente con los grupos o muestras en que se divide el estudio y la variable dependiente con la condición intervalar de la prueba estandarizada. Lo que conduce al estudiante a inferir la aplicación de una prueba estadística ANOVA.

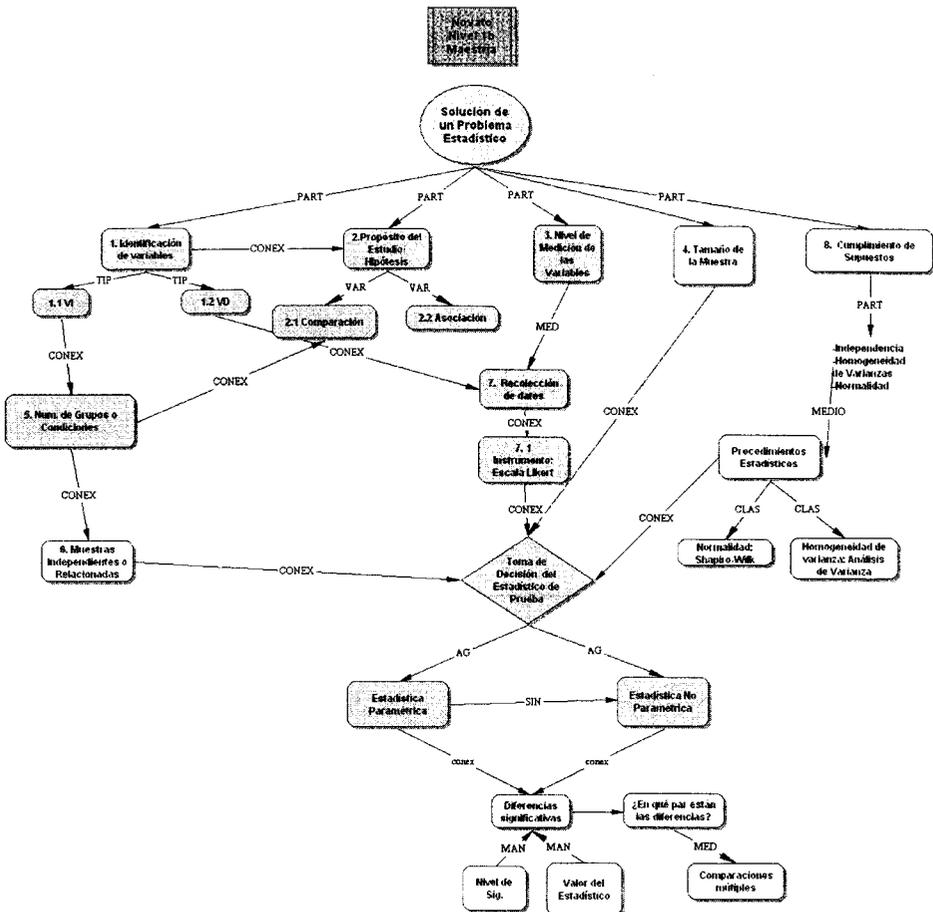


Figura 23. Secuencia de solución del novato nivel 1, maestría, de un problema estadístico que implica la aplicación del estadístico ANOVA.

Tesis de licenciatura. Nivel 2

En el ejemplo mostrado de un alumno en proceso de titulación, se observan mayores diferencias con respecto a los esquemas de los dos niveles anteriores (véanse Figuras 22, 23 y 24). En este ejemplo, el alumno se apoya en tres aspectos clave para su decisión, el tipo de prueba, la identificación de variables y el nivel de medición. De manera general se puede observar en el esquema de la Figura 22, que incluye conceptos en categorías que no corresponden con las clasificaciones de los expertos ni con las de los novatos de nivel posgrado, por ejemplo, incluye en la categoría de tipo de prueba la clasificación de paramétrica y afirma erróneamente que el muestreo probabilístico es una condición para las pruebas paramétricas. Por otra parte, menciona que el nivel de medición es un aspecto clave, sin embargo no logra identificar el nivel de medición de sus

variables y menos aún llega a proponer alguna prueba estadística para el problema en cuestión.

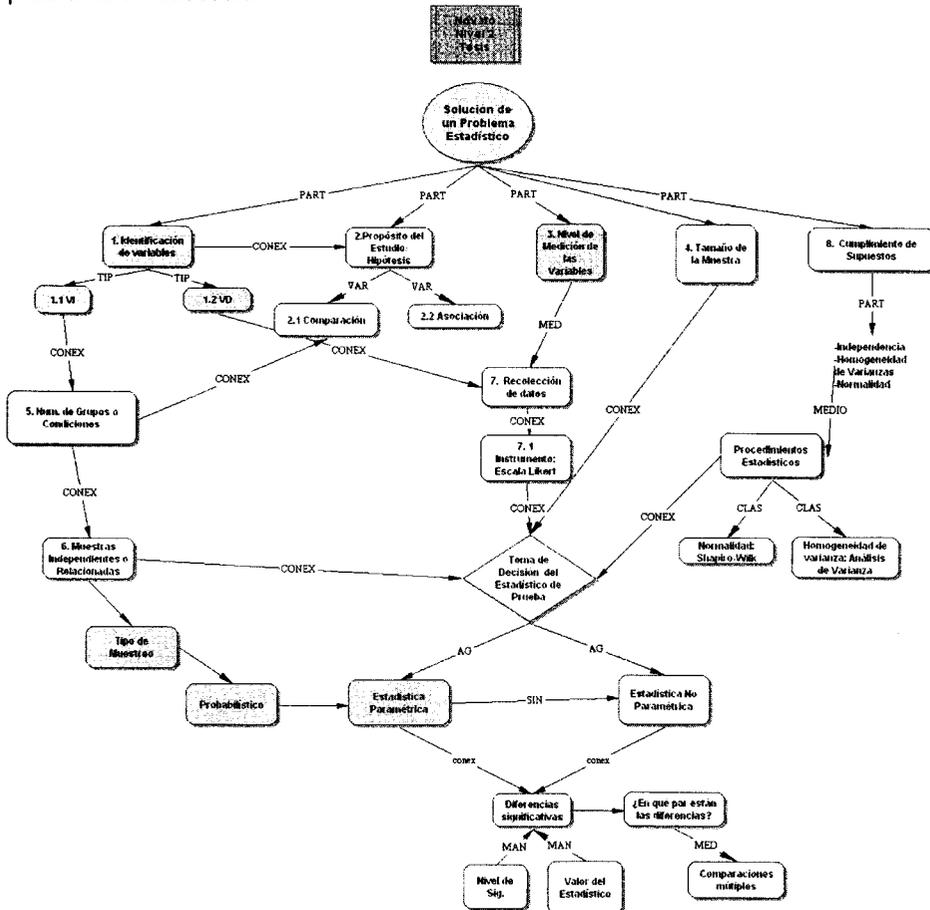


Figura 24. Secuencia de solución del novato nivel 2, tesis, de un problema estadístico que implica la aplicación del estadístico ANOVA.

3º a 5º grado de licenciatura. Nivel 3

Finalmente, tenemos el ejemplo de la ruta de decisiones tomada por el sujeto de licenciatura, quien se basa en dos criterios clave únicamente, el propósito de investigación y el número de muestras utilizadas (véase Figura 25). Sin embargo, cuando se trata de los siguientes niveles de decisión, se puede notar que al referirse al propósito de investigación no se refiere al propósito de buscar diferencias entre grupos, o establecer comparaciones entre condiciones ó muestras como sucede en los niveles de pericia anteriores. Tampoco logra identificar en qué nivel de medición se encuentra la variable dependiente mediante la escala Likert. Aunque el novato menciona la característica de utilizar muestras

aleatorias, no logra identificar algún tipo de prueba asociada. Por otra parte, infiere una posible aplicación de la prueba Wilcoxon (prueba que no se aplicaría en este caso), sin embargo no da argumentos para fundamentar su propuesta.

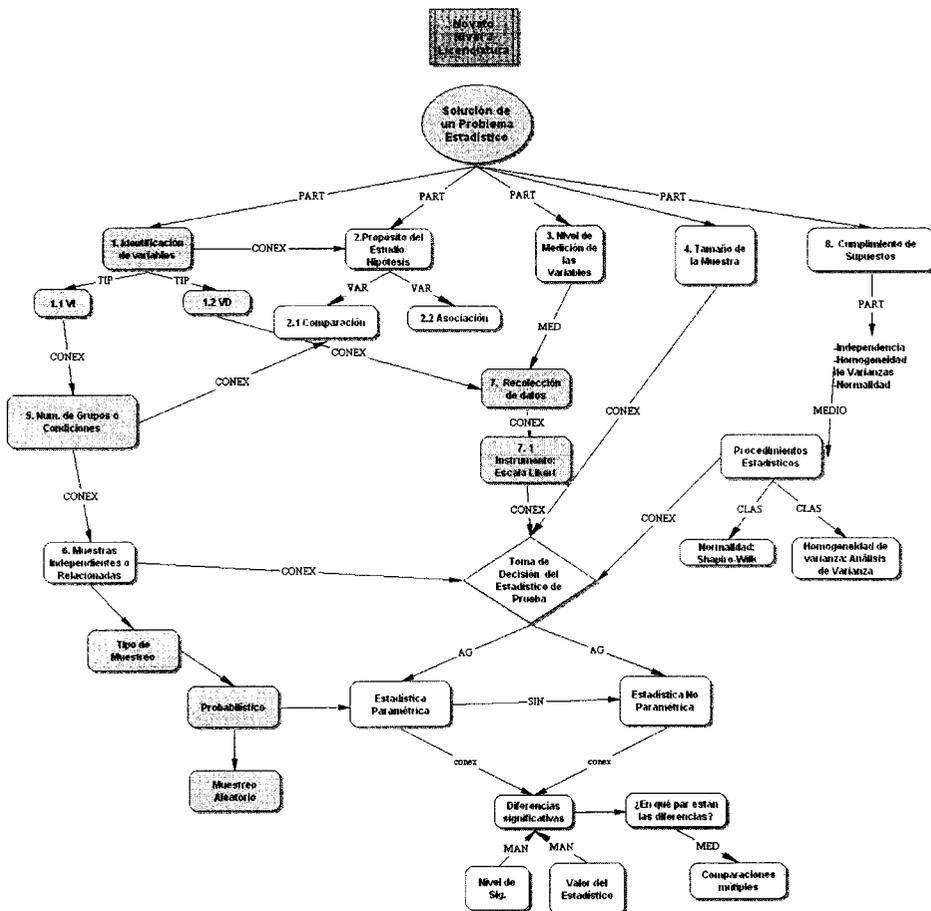


Figura 25. Secuencia de solución del novato nivel 3, 3º a 5º grado de licenciatura, de un problema estadístico que implica la aplicación del estadístico ANOVA.

En la Tabla 9 se puede observar una síntesis de los productos obtenidos de la Fase 3.

Tabla 9.

Productos de la Fase 3 Adquisición de Habilidades y Análisis Refinado

	HABILIDADES	CONOCIMIENTOS	MODELO MENTAL
FASE 3: ADQUISICIÓN DE HABILIDADES Y ANÁLISIS REFINADO	Los novatos en todos los niveles, son capaces de identificar las variables independiente y dependiente.	Los novatos de nivel 3 (licenciatura) poseen los conocimientos necesarios para identificar el tipo de instrumento utilizado, sin embargo no poseen conocimientos adecuados para desempeñar las otras condiciones de la elección de la prueba estadística.	El modelo mental para cada nivel de pericia, en comparación con el modelo experto se representa en las figuras 22, 23, 24 y 25.
	Los novatos de los niveles 2 y 3 (estudiantes en proceso de titulación y de 3er a 5º semestre de licenciatura, respectivamente) no toman en cuenta todos los aspectos clave para resolver problemas que implican la elección de la prueba estadística (número de grupos o condiciones, propósito del estudio, nivel de medición de las variables, tamaño de la muestra, cumplimiento de supuestos, decisión estadística, muestras independientes o relacionadas y la propia elección del estadístico de prueba).	Los novatos de nivel licenciatura poseen los conocimientos necesarios para identificar el nivel de medición de las variables, sin embargo no poseen conocimientos adecuados para desempeñar las otras condiciones de la elección de la prueba estadística.	
	Los novatos de los niveles superiores (1a y 1b) son capaces de identificar las variables, establecer el propósito del estudio y el nivel de medición en que se encuentran las variables.	Los novatos de los niveles 1 y 2 identifican adecuadamente el nivel de medición de las variables, el propósito del estudio y el tipo de instrumento, sin embargo no poseen los conocimientos necesarios para evaluar el cumplimiento de supuestos, el tamaño de la muestra y la decisión estadística.	
	Los novatos de todos		

los niveles, no
identifican
adecuadamente las
muestras
relacionadas e
independientes,
cumplimiento de
supuestos, tamaño
de la muestra y
decisión estadística.

6.4 Cuarta Fase Análisis de Datos y Representación del Conocimiento

6.4.1 Conocimiento conceptual, procedimental y estratégico involucrado en la selección de la prueba estadística

Las tablas a), b) y c) del Apéndice E, describen en detalle los conceptos, procedimientos y principios que están en juego durante el proceso de selección de la prueba estadística. El diseño instruccional de la enseñanza de la estadística que considere la selección de la prueba estadística, desde la perspectiva de la solución de problemas y la enseñanza situada, debería considerar los siguientes elementos conceptuales, procedimentales y estratégicos claves, utilizados por los expertos durante su proceso de decisión en la elección de la prueba estadística.

6.4.2 Principales errores conceptuales durante el proceso de elección de la prueba estadística

Al analizar las argumentaciones verbalizadas por los novatos, se localizaron algunos conceptos que presentan problemas para su comprensión en los diversos niveles de pericia (véase Tabla 10, para más detalle, véase Apéndice G). En seguida se describen algunos conceptos en los que se encontraron errores conceptuales.

Diseño Experimental

Los estudiantes usualmente asocian el concepto de experimento con el control de variables, sin embargo se aprecia que confunden el control de variables con la manipulación de las variables y la observación con la ausencia de control. En otro caso se asume que cuando se trata de sujetos diferentes se logra un mejor control de variables, es quiere decir, cuando estamos hablando de muestras relacionadas no existe control de variables.

Estrategias de Recolección de Datos

Algunos alumnos presentan problemas para determinar el nivel de medición de las variables medidas con escalas Likert, ya que no les queda claro si son medidas en escala nominal, ordinal o intervalar. Se debe mencionar que en sus respuestas,

los novatos no argumentan utilizando criterios estadísticos acerca del nivel de medición de las escalas Likert, tal como lo hacen los expertos. De la misma forma, existen dificultades para comprender por qué las variables que en un inicio están medidas en un nivel intervalar, cuando se forman grupos de comparación se convierten a un nivel de medición nominal.

Propósito del Estudio

Una de las cuestiones que se dificultan a los estudiantes novatos tiene que ver con la identificación del propósito u objetivo del estudio. Los novatos encuentran grandes dificultades para diferenciar entre los estudios que tienen como propósito establecer relaciones entre variables (que los expertos identifican como estudios que requieren un análisis de tipo correlacional) y los estudios que intentan establecer diferencias entre grupos o condiciones (al referirse a los estudios que implican un diseño experimental y que implican la comparación de muestras). Esta confusión de los novatos los lleva a confundir las pruebas estadísticas a aplicar en cada caso. Otro problema en la comprensión del propósito del estudio, se relaciona con la confusión entre los términos afecta y efecto, esto lleva a los novatos a inferir que se está buscando una relación de causa efecto, o bien una correlación entre variables.

Tamaño de la Muestra

El tamaño de la muestra es identificado como un factor importante para la toma de decisión sobre la prueba estadística apropiada para un problema estadístico particular, no obstante, cuando se intenta indagar sobre las razones para tomar alguna decisión al respecto, relacionadas con el tamaño de la muestra, no son capaces de clarificar ese conocimiento muy general. Dicho de otra manera, este concepto tan importante para los expertos en estadística, ya que se relaciona con la inferencia estadística y la probabilidad y el Teorema de Límite Central, no ha sido bien comprendido por parte de los novatos en todos los niveles.

Tipo de Muestreo

Entre los novatos, se aprecia poco conocimiento de las razones por las que los estudios utilizan muestreos aleatorios, ya que no se vincula con el control de variables en los estudios experimentales ni con la inferencia estadística ni la probabilidad, inclusive algunos novatos, llegan a confundir el muestreo probabilístico con la utilización de pruebas paramétricas.

Interpretación de Resultados

Cuando se requiere la interpretación de resultados a partir de una tabla de resultados (output) del programa estadístico SPSS, sobre un problema de ANOVA, los novatos no lograron identificar apropiadamente los elementos que se muestran en la tabla, tales como nivel de significancia, valor del estadístico o las diferencias intra y entre grupos. La única inferencia que pueden identificar los

novatos del nivel superior (doctorado) es que existen diferencias entre grupos, pero no pueden explicar por qué esas diferencias son o no significativas estadísticamente.

Decisión estadística

Así como los novatos no logran identificar los elementos a partir de los cuales se puede establecer una conclusión estadística acerca de los datos, consecuentemente no logran establecer una adecuada decisión estadística. Los novatos poseen alguna idea de que el resultado del proceso de decisión estadística conduce a la conclusión de aceptar o rechazar la hipótesis nula; sin embargo, no pueden explicar las razones estadísticas que los conducen a establecer esas conclusiones a partir de una regla de decisión. Otro error conceptual consiste en asociar el grado de significancia con la debilidad o fortaleza de la hipótesis. También tienen alguna idea de que se requiere un análisis adicional para establecer a favor de qué grupo se encuentran las diferencias (en el caso de un problema de ANOVA); sin embargo, no logran identificar qué prueba estadística es la apropiada para comparaciones múltiples.

Tabla 10
Tipos de errores conceptuales en que incurren los novatos

ERRORES CONCEPTUALES EN LA ELECCIÓN DE LA PRUEBA ESTADÍSTICA
• Diseño experimental (control de variables)
• Estrategias de recolección de datos (nivel de medición, instrumentos de medición, formación de grupos de comparación)
• Propósito del estudio (relación entre variables/comparación entre grupos o condiciones)
• Tamaño de la muestra (inferencia estadística)
• Tipo de muestreo (probabilístico/no probabilístico)
• Interpretación de resultados
• Decisión estadística (estadístico de prueba, nivel de significancia, regla de decisión, análisis pos hoc)

6.4.3 Aspectos problemáticos dónde se debe concentrar la enseñanza

Cuando resuelven problemas estadísticos (véase Tabla 11 y para más detalle véanse Figuras 22, 23, 24 y 25), los novatos en todos los niveles, son capaces de identificar las variables independiente y dependiente adecuadamente, pero los novatos de los niveles 2 y 3 (estudiantes en proceso de titulación y de 3er a 5º semestre de licenciatura, respectivamente) tienen dificultades en tomar en cuenta todos los aspectos clave para resolver problemas que implican la elección de la prueba estadística (número de grupos o condiciones, propósito del estudio, nivel de medición de las variables, tamaño de la muestra, cumplimiento de supuestos, decisión estadística, muestras independientes o relacionadas y la propia elección del estadístico de prueba).

Los aspectos problemáticos para todos los novatos, independientemente del nivel

son: muestras relacionadas e independientes, cumplimiento de supuestos, tamaño de la muestra y decisión estadística. Los aspectos mencionados anteriormente requieren una importancia mayor en la instrucción, ya que presentan mayores dificultades, a pesar de la instrucción y la experiencia. Por otra parte, los novatos de los niveles superiores (1a y 1b) son capaces de identificar las variables, establecer el propósito del estudio y el nivel de medición en que se encuentran las variables, sin embargo, no contemplan el tamaño de la muestra, ni la diferencia entre muestras independientes y relacionadas cuando intentan solucionar un problema estadístico. De la misma manera, no demuestran habilidades para establecer la decisión estadística (proceso que implica el nivel de significancia, el valor del estadístico y las comparaciones múltiples).

Tabla 11.
Aspectos problemáticos dónde se debe concentrar la enseñanza de la elección de la prueba estadística

ASPECTOS CLAVE	ASPECTOS PROBLEMÁTICOS PARA LOS NOVATOS (*)			
	Nivel 1a Doctorado	Nivel 1b Maestría	Nivel 2 Tesis	Nivel 3 Lic.
I. Identificación de variables				
A. Variable Independiente				
1. Num. de Grupos o Condiciones			*	
a. Muestras Independientes/ Relacionadas	*	*	*	*
B. Variable Dependiente				
II. Propósito del Estudio/Hipótesis			*	*
C. Asociación			*	*
D. Comparación			*	*
III. Nivel de Medición de las Variables				*
A. Recolección de Datos			*	
• Tipo de Instrumento			*	
IV. Tamaño de la Muestra	*	*	*	*
V. Cumplimiento de Supuestos	*	*	*	*
A. Independencia	*	*	*	*
B. Homogeneidad de Varianzas	*	*	*	*
C. Normalidad	*	*	*	*

D. Procedimientos Estadísticos	*	*	*	*
a. Normalidad: Shapiro-Wilk	*	*	*	*
b. Homogeneidad de varianza	*	*	*	*
VI Toma de Decisión del Estadístico de Prueba			*	*
A: Estadística Paramétrica/No Paramétrica	*		*	*
• Diferencias significativas	*	*	*	*
• Nivel de Significancia	*	*	*	*
• Valor del Estadístico	*	*	*	*
• ¿En qué par están las diferencias?	*	*	*	*
• Comparaciones múltiples	*	*	*	*

DISCUSIÓN Y CONCLUSIONES

El presente estudio se propuso como objetivo analizar el conocimiento experto de la solución de problemas de estadística inferencial, específicamente en relación con la elección de la prueba estadística. Este análisis implicó la identificación de los conceptos, procedimientos y estrategias involucrados en la elección de la prueba estadística utilizados, tanto por expertos, como por novatos en este campo de conocimiento.

La elección de la prueba estadística es una de las actividades que realizan los estudiantes de las carreras de Ciencias Sociales durante el desarrollo de investigaciones en las materias: estadística, métodos de investigación y durante la elaboración de tesis de licenciatura y posgrado. La elección de la prueba estadística es sólo una parte del proceso de la prueba de hipótesis, pero se eligió como tema central de esta investigación porque involucra una complejidad muy alta para los estudiantes (Ávila, Márquez y de la Rosa, 1999), dado que, para elegir la prueba estadística, los estudiantes deben poner en práctica sus conocimientos acerca de una serie de conceptos, procedimientos y estrategias de solución problemas que implican un razonamiento estadístico. La enseñanza de la estadística en los cursos tradicionales enfrenta muchos problemas para la comprensión de conceptos, el razonamiento, y la solución de problemas en todos los niveles de la enseñanza de las carreras del área de Ciencias Sociales. Los estudiantes en estas carreras utilizan la estadística como una herramienta de investigación, en diferentes momentos de su vida escolar y profesional. Sin embargo, según reportan los estudiantes entrevistados, cuando tienen que resolver un problema estadístico, usualmente recurren a otros alumnos más experimentados o a los profesores que imparten prácticas o asignaturas de estadística, para consultar sobre las pruebas estadísticas apropiadas para su investigación. Por tanto, se supone que la instrucción no ha sido efectiva para formar estudiantes que resuelvan problemas estadísticos y razonen estadísticamente de una manera cercana a la forma en que los expertos lo hacen.

La construcción de un modelo de conocimiento experto desarrollado en esta investigación se vio dificultado por la escasez de investigaciones anteriores sobre este campo. Los estudios sobre la pericia en estadística han sido realmente escasos (Alacaci, 2004; Gal y Garfield, 1997). Probablemente esto se debe a que se considera que este campo de conocimiento se encuentra bien definido, ya que existen una gran cantidad de libros de texto y manuales que dan cuenta de la estructura del conocimiento de la disciplina y su aplicación en las investigaciones en Ciencias Sociales. La mayoría de los profesores de estadística apoyan su enseñanza en la estructura de la disciplina expuesta en los libros de texto y no en estudios sobre la adquisición de conocimiento experto en la materia; sin embargo la estructura que se ha extraído de los libros, no revela la complejidad del proceso de solución, ni las decisiones que toman los expertos cuando se enfrentan a problemas reales. Adicionalmente, cuando enfocamos la enseñanza de la estadística desde la perspectiva de la solución de problemas, encontramos que la

mayoría de los estudiantes novatos no poseen las habilidades ni el conocimiento, que los expertos han desarrollado a través de muchos años de experiencia, y cuando poseen dicho conocimiento, éste no se encuentra estructurado adecuadamente (Lavigne y Glaser, 2001).

Para extraer el conocimiento experto en el campo de la estadística, se utilizó la metodología de Análisis Cognitivo de Tareas. El ACT utilizado para este estudio, se apoyó en el Modelo (MATI), de Análisis Integrado de Tareas (Ryder y Redding, 1993), consistente en tres fases, para ello se adaptó dicho modelo agregándole una cuarta fase denominada: "Análisis de Datos y Representación del Conocimiento", cuyo objetivo es realizar un trabajo sintético, que resume los hallazgos de las fases anteriores. El uso extendido de este modelo puede ser utilizado para extraer el conocimiento experto en disciplinas que presentan problemas de aprendizaje y bajo rendimiento, como es el caso de la estadística. A partir de los modelos cognitivos obtenidos con la utilización del ACT, es posible plantear directrices pedagógicas para mejorar la enseñanza e identificar aquellos puntos en los que es necesario concentrar la enseñanza o buscar mejores maneras de abordar los contenidos problemáticos en los que los alumnos, en su calidad de novatos en el dominio de un conocimiento, encuentran mayores problemas para dominar el conocimiento.

El análisis del conocimiento experto se centró básicamente en la obtención de datos cualitativos y dio como resultado hallazgos similares a los encontrados en los estudios clásicos sobre el conocimiento experto. Es decir, el experto ocupa una cantidad considerable de tiempo representándose el problema a resolver, su conocimiento se encuentra estructurado en unidades de conocimiento o paquetes, que son difíciles de verbalizar en voz alta (Chi, Glaser y Rees, 1982; Larkin, 1980). La razón de no realizar un análisis cuantitativo de las diferencias entre expertos y novatos en estadística, se debe a que la importancia del estudio no radica en confirmar las diferencias generales entre expertos y novatos, sino en identificar cómo se van adquiriendo los conocimientos y habilidades a medida que se avanza en un nivel de pericia, así como en identificar los cuellos de botella en el aprendizaje, para enfocar la enseñanza y buscar nuevas estrategias para mejorar su comprensión y las habilidades para solucionar problemas en el contexto de las disciplinas de las Ciencias Sociales.

Con el conocimiento extraído de los expertos se elaboró el espacio del problema, el cual contiene los aspectos clave o más importantes que componen el proceso de decisión, de acuerdo al criterio de los expertos entrevistados (Means, 1993). Para los expertos, los conceptos básicos involucrados en el proceso de la elección de la prueba estadística son básicamente los siguientes: el objetivo o propósito de investigación (buscar relaciones, comparar grupos, establecer predicciones), hipótesis de investigación, el nivel de medición de las variables (ligadas al instrumento de medición), tipo de muestreo (probabilístico/no probabilístico), el diseño (una sola muestra, dos, tres, cuatro muestras), lo cual coincide con lo planteado por Hawkins, Joliffe y Glickman (1992). A pesar de que la mayoría de expertos se refieren a los mismos aspectos clave para tomar la decisión acerca la

prueba estadística, no todos coinciden en el nivel de importancia de cada uno, si tomamos en cuenta que se mencionan en primero, segundo, tercer lugar, etcétera, durante la solución de los problemas.

Se encontraron algunas diferencias en los expertos en cuanto al énfasis que se le asigna a determinados conceptos y procedimientos. Por ejemplo, sólo dos expertos mencionan la importancia de llevar a cabo un análisis exploratorio de datos para verificar el cumplimiento de supuestos para la aplicación de las pruebas paramétricas, en tanto que para otros, es suficiente con evaluar los aspectos clave que se explicitan en el problema, tales como nivel de medición de las variables, propósito del estudio y tamaño de la muestra para determinar la aplicación o no de una prueba paramétrica. Por tanto, se asume que aunque el análisis exploratorio de datos es un tema de creciente interés en la práctica de la estadística en las ciencias sociales (Hawkins, Joliffe y Glickman, 1992), su uso aún no se encuentra extendido entre los expertos. Por lo tanto, una de las propuestas del presente estudio es analizar la pertinencia de incluir contenidos específicos de análisis exploratorio de datos en los programas de enseñanza de las carreras de Ciencias Sociales.

Sin embargo, no hay que descartar que las diferencias en los criterios expertos para abordar los problemas estadísticos, pudieron deberse a las diferencias intrínsecas en la formación y experiencia personal de cada uno de los expertos, ya que mientras algunos se han dedicado más extensamente a la investigación, otros se han dedicado a impartir principalmente asignaturas prácticas de estadística.

Entre los conceptos asociados a la elección de la prueba estadística, considerados por los expertos como más problemáticos para la comprensión de los alumnos, se encuentran los conceptos relacionados con la probabilidad (como la distribución normal y el teorema de límite central) y otros conceptos relacionados con el método de investigación (planteamiento de hipótesis, la determinación de los niveles de medición de las variables, la definición operacional de las variables, las estrategias de medición (instrumentos observación, entrevistas, etcétera), el tipo de muestreo y el diseño de investigación. Esto coincide de manera importante con lo planteado por Hawkins, Joliffe y Glickman, (1992). Especialmente el concepto de distribución normal es difícil de comprender para los alumnos, de acuerdo con lo que refieren los expertos. Sin embargo, el mayor problema consiste en la actitud negativa con la que ingresan los alumnos a los primeros niveles de estadística de la licenciatura, debido a las experiencias desmotivantes de los años escolares previos, con respecto a los contenidos escolares que involucran temas de matemáticas.

Entre las sugerencias se encuentra la propuesta de simplificar los contenidos de estadística para hacerlos más entendibles a los alumnos, es decir, ir de lo simple a lo complejo. Otra sugerencia consiste en situar la enseñanza en ejemplos más realistas, por ejemplo, dar la oportunidad a los alumnos para que elijan la temática de investigación de acuerdo a sus propios intereses de la vida cotidiana y dentro de la disciplina (en este caso la Psicología). Para los expertos, los conceptos de

probabilidad no se deben enseñar de forma aislada, sino relacionados con la toma de decisiones relacionadas con la prueba de hipótesis.

A partir del análisis de las entrevistas y tareas aplicadas a los expertos, se elaboró un esquema del Espacio del Problema, correspondiente a la Elección de la Prueba Estadística que contenía los pasos típicos que los expertos seguían para representar y resolver los problemas estadísticos. En este esquema se corrobora lo que expresaron los expertos en la entrevista, con respecto a los aspectos claves o más importantes en su decisión sobre la prueba estadística.

· Cuando los expertos intentan resolver un problema estadístico, concentran su atención en diversos aspectos clave, los cuales no tienen un orden de importancia específico; sin embargo, se consideran en la mayoría de los casos de la siguiente forma: La mayoría de los expertos, ocupa gran parte del proceso de análisis del problema estadístico en explorar las características de las variables que intervienen en el estudio, variable independiente y dependiente y reconocer el propósito del estudio, al que dividen en problemas de comparación (entre grupos o condiciones) o de relación (para referirse a estos últimos como casos de correlación). En seguida, los expertos proceden a determinar el nivel en que se han medido las variables, principalmente la variable dependiente, aspecto que está íntimamente relacionado con la estrategia de recolección de datos (instrumento, entrevista, observación, etcétera). El nivel de medición de la variable independiente, se encuentra más relacionado con el número de condiciones o grupos del diseño de investigación.

La consideración de los grupos como muestras independientes o relacionadas se relaciona también con el nivel de medición, es decir, el hecho de que se esté midiendo al mismo grupo en dos o más momentos o bien, se trata de grupos diferentes, influye en la elección de la prueba estadística más apropiada. El tamaño de la muestra, también se considera un aspecto de suma importancia, ya que esto se vincula de manera central con el cumplimiento de supuesto de normalidad (asociado a la idea de que una muestra muy pequeña tiende a constituir una distribución con sesgo importante), decisión que conduce a seleccionar el tipo de estadística a utilizar (paramétrica o no paramétrica). La normalidad es solo uno de los supuestos que se deben cubrir para aplicar una prueba paramétrica, para lo cual se utilizan una serie de procedimientos estadísticos, como la prueba de Shapiro-Wilk. Todas estas decisiones parciales confluyen en la toma de decisión de la prueba estadística, que consiste en ponderar todos los elementos anteriormente señalados para optar por una prueba paramétrica, siempre y cuando se cumplan los supuestos mencionados anteriormente, o en su defecto optar por una prueba no paramétrica. Una vez que se aplica un procedimiento estadístico, el siguiente paso es establecer la decisión estadística, considerando el nivel de significancia y el valor del estadístico y finalmente tomar la decisión de aplicar un análisis post hoc, que permita establecer comparaciones múltiples, y así concluir en qué grupo se encuentran las diferencias.

El análisis de las diferencias entre los expertos y los novatos, con respecto a la representación de los problemas estadísticos revela que los novatos se enfocan considerablemente más en aspectos superficiales y moderadamente en principios que los expertos, quienes tienden a basarse de manera importante en principios, es decir, en considerar en primera instancia los aspectos clave correspondientes al espacio del problema. Sin embargo, los novatos del nivel 1b (nivel maestría), representaron los problemas estadísticos de una forma más cercana a los expertos, que los novatos 1^a (de nivel doctorado), quienes calificaron más bajo en el nivel basado en principios. De la misma forma, los novatos del nivel 2 y 3, no presentan diferencias importantes, ya que representan los problemas de manera muy similar, se debe considerar el hecho de que los alumnos del nivel 2, son alumnos que se encuentran en proceso de elaboración de tesis y los del nivel 3 son alumnos de tercer a quinto grado de licenciatura.

En términos generales se puede afirmar que, existe variabilidad en las habilidades para representarse los problemas estadísticos entre los diferentes niveles de los novatos, por lo que es posible deducir que probablemente existen algunas variables, que afectan el desempeño de los estudiantes y que no fueron contempladas en el presente estudio, tales como experiencias previas o cursos adicionales de estadística o metodología. Otra explicación posible se relaciona con el hecho de que los estudiantes que se encuentran en proceso de elaboración de tesis, han pasado mucho tiempo sin tener contacto con los contenidos de estadística o metodología. Por otra parte, habrá que considerar que los alumnos de maestría recibieron recientemente un curso sobre metodología de investigación, como parte de su formación, mientras que los alumnos de doctorado no lo han tomado.

Los esquemas de representación de los problemas estadísticos, elaborados a partir de las clasificaciones de los expertos, revelan diferencias en cuanto a la forma de agrupar los problemas estadísticos; no obstante, todos coinciden en separar los problemas en dos categorías más grandes, que son los problemas de relación (correlación) y los de diferencias entre grupos o condiciones (experimentales). En un segundo nivel de clasificación se consideran los problemas de muestras relacionadas o independientes y las pruebas paramétricas y no paramétricas. Algunos expertos proceden a clasificarlos de acuerdo a las pruebas estadísticas más apropiadas. Sin embargo, *es notorio que en la representación de los problemas estadísticos por parte de los expertos, está ausente la consideración del tamaño de la muestra y el cumplimiento de supuestos para decidir sobre la clasificación; por tanto, se puede inferir que estos aspectos clave no son centrales para los expertos, o al menos no los consideran en primer lugar en el momento de representarse los problemas estadísticos.*

Por su parte, los novatos de los niveles 1a (doctorado) y 1b (maestría), tienen éxito en clasificar los problemas como problemas de relación o comparación, establecer el nivel de medición de las variables y en considerar el número de grupos o condiciones. Sin embargo, tienen dificultades para clasificar los problemas estadísticos de Ji Cuadrada. Con respecto a los novatos de los niveles

2 y 3, la clasificación carece de una estructura definida, lo que refleja una representación defectuosa de los problemas estadísticos, por ejemplo, *confunden los tipos de problemas que hablan de relaciones entre variables con los que hacen comparación entre grupos*, simplemente basándose en claves verbales cómo: "qué relación existe entre...", "se compararon dos grupos o condiciones", pero sin considerar otros aspectos, tales como tamaño de la muestra, nivel de medición, número de grupos. Otra forma de clasificar los problemas se relaciona con los aspectos considerados superficiales, ya que no se relacionan con conceptos estadísticos, sino conceptos superficiales, por ejemplo: temática del problema (alguna conducta, aplicación de una prueba u otra cosa, que da como resultado la modificación de una conducta, y cómo afectan determinadas condiciones a la conducta).

Con respecto a la actividad de solución de problemas estadísticos, *los expertos* tienden a considerar los mismos aspectos clave, sin embargo, existen algunas diferencias, en cuanto al orden en que consideran prioritario un aspecto u otro. Aunque la mayoría llega a la misma conclusión, *algunos enfatizan más la consideración del tamaño de la muestra para elegir una prueba no paramétrica, en tanto que otros, le dan prioridad a realizar un análisis exploratorio*, para analizar el cumplimiento de supuestos de normalidad, homogeneidad de varianzas, e independencia apoyándose en procedimientos estadísticos.

En el caso de los novatos, se observaron diferencias más evidentes entre los diferentes niveles de pericia, comparados con la secuencia de pasos para solucionar un problema estadístico que siguieron los expertos. Por ejemplo, en los niveles 1a (doctorado) y 1b (maestría), los pasos seguidos incluyen los mismos aspectos claves, a excepción del tamaño de la muestra, el cumplimiento de los supuestos y los procedimientos estadísticos necesarios para valorarlos, la consideración de los diseños que incluyen muestras independientes o relacionadas y los aspectos relacionados con la decisión estadística, considerando el nivel de significancia y el valor del estadístico. Dicho de otra manera, no existen diferencias entre los procedimientos seguidos por los estudiantes de doctorado y maestría, pero sí se observan notables diferencias con el proceso de solución seguido por los expertos.

Las diferencias se hacen aún mayores cuando comparamos a los novatos del nivel 2 (tesis), quienes omiten algunos aspectos considerados de suma importancia para la elección de la prueba estadística por los expertos (además de los aspectos omitidos por los novatos del nivel 1), tales como el propósito del estudio (comparación/asociación), el número de grupos o condiciones, la recolección de datos e instrumento de medición, además de que no llegan a determinar o sugerir una posible prueba estadística para aplicar.

Por su parte, los novatos del nivel 3 (licenciatura), a diferencia de los de nivel 2, sí consideraron la posibilidad de encontrar muestras independientes o relacionadas; sin embargo, no lograron diferenciar entre las pruebas paramétricas y no paramétricas, y más aún no identificaron el nivel de medición de las variables

(demás de los aspectos omitidos por los novatos del nivel 1 y 2).

En suma podríamos afirmar que los novatos difieren en los aspectos claves que toman en cuenta cuando resuelven un problema estadístico que involucra la elección de la prueba estadística, estas diferencias están en relación con los diferentes niveles de pericia y se asume que esta diferencia está relacionada con el nivel de entrenamiento en estadística o experiencia en investigación utilizando métodos estadísticos. Estas diferencias que se muestran son progresivas, por lo que se puede trazar una ruta instruccional de los conceptos y procedimientos que se deben enfatizar en la enseñanza de la estadística en los diferentes niveles de pericia. Cabe aclarar que estas diferencias entre los diferentes niveles de pericia, son casi imperceptibles entre los novatos de los niveles 2 y 3, es decir entre alumnos de los primeros semestres de psicología (en los cuales se toman las materias relacionadas con la estadística) y los alumnos que se encuentran realizando su tesis de licenciatura, una explicación posible a esta situación, es que los estudiantes de los primeros semestres de psicología cursan las materias de estadística descriptiva e inferencial fuera de contexto y transcurren aproximadamente cinco semestres, en el mejor de los casos, para que tengan nuevamente algún contacto con los contenidos de estadística, siempre y cuando realicen su tesis utilizando algún método de análisis estadístico.

Otro hallazgo es que los alumnos de maestría han logrado resultados más cercanos a los expertos, lo cual se puede explicar por un cercano entrenamiento para el examen de selección para el ingreso y el entrenamiento obligatorio en metodología de investigación cuantitativa que cursan dichos estudiantes.

La muestra pequeña de expertos y novatos en la presente investigación no permite generalizar los hallazgos a poblaciones mayores; sin embargo, el estudio constituye un punto de partida exploratorio, para dar lugar a diversos estudios que indaguen posibles relaciones entre los diferentes factores que pueden influir en el conocimiento experto en estadística. Asimismo, estos estudios se pueden extender a otras habilidades que componen el campo de estudio de la pericia en estadística, tales como probabilidad y distribución muestral, conceptos clave problemáticos en la comprensión de los problemas estadísticos.

Entre las limitaciones de la presente investigación, se pueden mencionar dos aspectos principales, uno que se refiere a las dificultades en la aplicación y análisis del ACT, y el otro se relaciona con la selección de los diferentes grados de pericia, dado que los criterios de selección siempre son de alguna manera arbitrarios y una mayor experiencia o grado académico no garantizan un mayor dominio en el campo de conocimiento de que se trate.

La aplicación del ACT es especialmente complejo y laborioso, por tal motivo la mayoría de los estudios se realizan con un escaso número de expertos, debido en parte a la escasa disponibilidad de los sujetos que han logrado una maestría considerable en una disciplina (Quilici y Mayer, 1996). El caso de este estudio no es la excepción, ya que sólo fue posible acceder a cinco expertos que cumplieran

el perfil especial de dominio y años de experiencia, además de la facilidad para disponer de tiempo para la entrevista y las dos tareas del estudio. Por ejemplo, en la Fase de Orientación, sólo fue posible entrevistar a dos de los expertos y no a los cinco como se tenía previsto en un principio. Por otra parte, la selección de los novatos en tres niveles, se realizó con el fin de lograr diferentes grados progresivos de pericia en estadística en Ciencias Sociales, la cual es una tarea muy difícil de definir y delimitar. Por tal motivo, habría que reconocer que los criterios de selección (alumnos de doctorado, maestría, en proceso de elaboración de tesis, 3º a 5º de licenciatura), son de alguna manera arbitrarios y que pudieran no coincidir necesariamente con un nivel de pericia efectivo, en virtud de que pueden existir algunas variables que influyen en el entrenamiento y el grado de estructuración del conocimiento, no consideradas en la selección de los sujetos de estudio.

De la misma forma, la dificultad para acceder y disponer del tiempo de los expertos, además de la extensión del tiempo para las entrevistas y tareas del estudio, no permiten utilizar a un número mayor de sujetos para realizar análisis estadísticos que permitieran comparar de manera cuantitativa los datos obtenidos. Por lo que, paradójicamente se ha realizado un estudio acerca del conocimiento estadístico, sin utilizar procedimientos estadísticos, esta característica puede resultar enriquecedora, si consideramos el potencial del análisis cognitivo de tareas para indagar el conocimiento experto en disciplinas complejas, como la estadística.

Consideraciones Finales

Las implicaciones pedagógicas del presente estudio se apoyan en la idea de que la estadística descriptiva e inferencial se debe enfocar desde la perspectiva de la solución de problemas (Lavigne & Glaser, 2001; Lajoie, Jacobs & Lavigne, 1995), apoyando a los estudiantes de Ciencias Sociales para adquirir habilidades y conocimientos necesarios para representar, razonar y plantear problemas estadísticamente (Garfield & Ahlgren, 1994).

La formación de los repertorios necesarios en los estudiantes requiere replantearse la enseñanza tradicional de la estadística, en la cual se abordan los conocimientos estadísticos de una manera aislada, utilizando ejemplos de libro de texto y datos simulados o problemas artificiales que no tienen sentido para los alumnos. Esta forma tradicional de abordar la enseñanza de la estadística basada en la memorización de conceptos abstractos para los alumnos, así como en el énfasis en el cómputo de datos, carentes de relación con la actividad profesional de los alumnos, afecta directamente la motivación de los estudiantes con respecto al aprendizaje de la estadística. En su lugar, se plantea la alternativa de brindar a los alumnos la posibilidad de realizar sus propios proyectos de investigación, escogiendo las temáticas de acuerdo a sus propios intereses y relacionados con la disciplina de estudio desde los primeros semestres de la carrera. Como plantean Lajoie, Jacobs y Lavigne (1995) y Lavigne y Munsie (1993), un programa alternativo debería involucrar conocimientos acerca del método de investigación,

tales como diseños experimentales, planteamiento de hipótesis, objetivo de estudio, junto con los conocimientos de estadística, ya que están fuertemente vinculados con la solución de problemas estadísticos. Asimismo, se considera de suma importancia proporcionar una práctica extensiva en la representación de los diferentes tipos de problemas estadísticos característicos de la solución de problemas en cada disciplina que utilice la estadística como una herramienta del quehacer profesional.

Como lo señalan los expertos, existe una actitud negativa hacia la estadística en los alumnos que ingresan a las carreras de Ciencias Sociales, que han adquirido a través de sus experiencias negativas con las materias relacionadas con las matemáticas, como lo demuestran los altos índices de reprobación y deserción (Avila, Márquez & de la Rosa, 1999). Estos resultados se ven influidos fuertemente por la aproximación de la enseñanza tradicional de las matemáticas que enfocan la enseñanza como un cómputo de datos, más que con la formación de un repertorio de conocimientos y habilidades que posibilite a los estudiantes hablar el lenguaje de la estadística y razonar para resolver problemas o plantear problemas que le permitan acceder al conocimiento propio de su campo de estudio (Garfield, 1992).

El presente estudio presenta evidencia acerca de la viabilidad y la bondad de utilizar las aproximaciones experto-novato y el método de *Análisis Cognitivo de Tareas* para extraer el conocimiento experto e identificar las áreas de conocimiento y habilidades cognitivas que presentan mayor problema para los aprendices de disciplinas de gran complejidad, tales como la estadística. De esta manera, es posible hacer ajustes en las estrategias de instrucción, secuenciación de los contenidos y aspectos clave, en los cuales se debe concentrar la enseñanza en cada nivel de pericia.

Para futuras investigaciones se sugiere seleccionar a los sujetos novatos mediante una prueba de conocimientos y la aplicación de cuestionarios que permitan identificar más precisamente los niveles de pericia para poder caracterizar de mejor manera la progresión de los modelos mentales. Esto permitirá asimismo, detectar los errores conceptuales y las omisiones de procedimientos, y de esta manera elegir los procedimientos instruccionales más adecuados para cada nivel de pericia.

Se sugiere realizar investigaciones que aborden diversas alternativas para mejorar la enseñanza de la estadística y comparar su eficacia en cada nivel de pericia, poniendo en práctica diseños instruccionales que incorporen las aproximaciones de la solución de problemas, la enseñanza basada en proyectos, y los supuestos de la enseñanza situada, además del apoyo de las nuevas tecnologías y estrategias instruccionales avanzadas como el modelo de tutelaje cognitivo "cognitive apprenticeship", la enseñanza recíproca y la creación de comunidades de aprendizaje presenciales y virtuales.

La conclusión final tiene que ver con una postura optimista acerca de la

enseñanza de la estadística en el ámbito de las Ciencias Sociales, en la cual es posible mejorar y enriquecer las prácticas de enseñanza y aprendizaje actuales, siempre y cuando se adopte un diseño instruccional alternativo, que implique la enseñanza de la estadística como una disciplina poco estructurada, en la cual es importante razonar estadísticamente y hablar el lenguaje de la estadística, antes que aplicar fórmulas y aprender procedimientos en forma memorística, fuera de contexto y poco relacionados con el conocimiento de la disciplina y los métodos de investigación en Ciencias Sociales. Este es el momento de “empoderar” estadísticamente a los estudiantes, como lo define Lojoie et al. (1995), al darles la oportunidad de aprender estadística haciéndola realmente.

REFERENCIAS

- Alacaci, C. (2004). Inferential statistics: Understanding expert knowledge and its implications for statistics education. *Journal of Statistics Education*, 12 (2), www.amstat.org/publications/jse/v12n12/alacaci.html.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*(89), 369-406.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, Ma.: Harvard University Press.
- Avila, C. J., Márquez, R. L., & de la Rosa, R. R. (1999). *Diagnóstico de la enseñanza de la estadística para el desarrollo de un ambiente de aprendizaje basado en la computadora*. Ponencia presentada en el Congreso General de Cómputo 99, México, D.F.
- Bransford, J., Hasselbring, T., Barron, B., Kulewicz, J. S., Guzdial, M., & Palincsar, A. (1989). Uses of macro-contexts to facilitate mathematical thinking. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 127-147). Reston, VA.: National Council of Teachers of Mathematics.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn* Washington D.C.: National Academy Press.
- Bransford, J. D., & Stein, B. S. (1993). *The ideal problem solver* (2a ed.). New York: Freeman.
- Chervany, N. L., Benson, P. G., & Iyer, R. K. (1980). The planning stage in statistical reasoning. *American Statistician*, 34 (4), 222-226.
- Chervany, N. L., Collier, R. O., Fienberg, S. E., Johnson, P. E., & Neter, J. (1977). A framework for the development of measurement instruments for evaluating the introductory statistics course. *American Statistician* (31), 17-23.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science* (5), 121-151.
- Chi, M. T. H., & Glaser, R. (1985). Problem solving abilities. In R. S. Sternberg (Ed.), *Human abilities*. San Francisco: Freeman.
- Chi, M. T. H., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hilldale, NJ: Erlbaum.
- Chi, M. T. H., Glaser, R., & Rees, E. (1982). *Expertise in problem solving*. En R. S. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, N.J.: Erlbaum.
- Chi, M. T. H., & Rees, E. (1983). A learning framework for development. In M. T. H. Chi (Ed.), *Trends in memory development research*. Basilea: Karger.
- Coolidge, F. L. (2000). *Statistics: A gentle introduction*. London : Sage PublicationsLtd.
- Daniel, W. W. (1977). Statistical significance versus practical significance. *Science Education* (61), 423-427. London
- Evans, J., & St., B. T. (1986). *Bias in human reasoning, causes and*

- consequences. Essays in cognitive psychology*. Hillsdale, N.J.: Laurence Erlbaum Associates.
- Gal, I., & Garfield, J. B. (1997). Curricular goals and assessment challenges. En Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 1-13). Amsterdam: IOS Press.
- Gardner, M., & Altman, D. G. (1989). *Statistics with confidence: Confidence Intervals and Statistical Guidelines*. British Medical Journal.
- Gardner, P. L., & Hudson, I. (1999). University students' ability to apply statistical procedures. *Journal of Statistics Education*, 7(1), www.amstat.org/publications/jse/secure/v7n1/gardner.cfm.
- Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3), www.amstat.org/publications/jse/v10n13/garfield.html.
- Garfield, J., & Ahlgren, A. D. (1994). Students reactions to learning about probability and statistics. Evaluating the quantitative literacy project. *School Science and Mathematics*, 94(2), 89-95.
- Garfield, J. B. (1992, Enero). How students learn statistics. Ponencia presentada en (Reunión de Invierno) American Statistical Association, Louisville, Kentucky.
- Garfield, J. B., & Ahlgren, A. D. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal of Research in Mathematics Education*, 19(1), 44-63.
- Garfield, J. B., & Burril, G. (1996). Research on the role of technology in teaching and learning statistics. Ponencia presentada en IASE Round Table Conference, University of Granada Spain.
- Garnham, A., & Oakhill, J. (1996). *Manual de psicología del pensamiento*. Barcelona: Ediciones Paidós Ibérica, S.A.
- Hawkins, A., Jolliffe, F., & Glickman, L. (1992). *Teaching statistical concepts*. New York: Longman Group UK Limited.
- Holland, J. H., Holyoak, K. J., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery*. Cambridge, Mass: The MIT Press.
- Hubbard, R. (1997). Assessment and the process of learning statistics. *Journal of Statistics Education*, 5(1), <http://www.amstat.org/publications/jse/v5n1/hubbard.html>.
- Johnson, L. W. (1981). Teaching hypothesis testing as a six steps process. *Teaching Statistics*(3), 47-49.
- Jonassen, D. (Primavera, 2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education*, 35(3), 362-381.
- Lajoie, S. P., Azevedo, R., Zhao, X., Kinnon, J., & Fleiszer, D. M. (1997). Cognitive tools for assessment in a high information flow environment. Manuscrito no publicado, Montreal, Quebec.
- Lajoie, S. P., Jacobs, V. R., & Lavigne, N. C. (1995). Empowering children in the use of statistics. *Journal of Mathematical Behavior* (14), 401-425.
- Lajoie, S. P., & Vivet, M. (1999). *Artificial intelligence in education '99. Open learning environments: New computational technologies to support learning, exploration, and collaboration*. Amsterdam: IOS Press.
- Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and

- novice performance in solving physics problems. *Science*(208), 1335-1342.
- Lavigne, N. C., & Glaser, R. (2001). *Variability in student representations of statistics problems*. Ponencia presentada en la Reunión Anual de la American Educational Research Association (AERA), Seattle, WA.
- Lavigne, N. C., & Glaser, R. (2003). *Assessing expert knowledge representations of introductory statistics* (CSE Reporte Técnico No. 600). Pittsburg: CRESST, LRDC, University of Pittsburg.
- Lemke, J. (1997). *Aprender a hablar ciencia*. Barcelona: Paidós.
- Lesagold, A., Lajoie, S. P., Logan, D., & Eggan, G. (1990). Applying cognitive task analysis and research methods to assessment. En *Diagnostic monitoring of skill's knowledge acquisition* (pp. 325-350). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Means, B. (1993). Cognitive task analysis as a basis for instructional design. En M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 97-118). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Nisbett, R. E. (1993). *Rules for reasoning*. Hillsdale, NJ: Lawrence Erlbaum.
- Norman, D. A. (1985). *El aprendizaje y la memoria* (M. V. Sebastian y T. d. Amo, Trans.). Madrid: Alianza.
- Norman, D. A., Rumelhart, D. E., & LNR, G. (1975). *Explorations in cognition*. San Francisco: Freeman.
- Pagano, R. (1999). *Estadística para las Ciencias del Comportamiento*. México: Thomson editores, pp. 430.
- Pozo, J. I. (1993). *Teorías cognitivas del aprendizaje*. Madrid: Ediciones Morata.
- Quillici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, 88(1), 144-161.
- Resnick, L. B. (1989). Treating mathematics as an ill-structured discipline. En R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 32-60). Reston, VA: National Council of Teachers of Mathematics.
- Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning and restructuring: Tree modes of learning. En J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, N.J.: Erlbaum.
- Rumelhart, D. E., & Ortony, A. (1982). The representation of knowledge in memory. *Infancia y Aprendizaje*, 19(20), 115-158.
- Ryder, J. M., & Redding, R. E. (1993). Integrating cognitive task analysis into instructional systems development. *Educational Technology Research & Development*, 41(2), 75-96.
- Scheaffer, R. L. (1988). *Statistics in the schools: The past, present and future of the quantitative literacy project*. ASAProStEd, 71-78.
- Shoughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. En G. Doug (Ed.), *Handbook on research in mathematics education*. New York: Macmillan.
- Shute, V., Surgue, B., & Willis, R. (1997, March, 25). Automating cognitive task analysis. Ponencia presentada en el Simposium Cognitive Technologies for

- Knowledge Assessment. AERA Annual Meeting., Chicago, IL.
- Simon, H. A. (1978). Information processing theory of human problem solving (A. Moreno, Trans.). En K. W. Estes (Ed.), *Lecturas de psicología del pensamiento*. Madrid: Alianza.
- Simon, H. A. (1986). Decision making and problem solving. Reporte de investigación. Washington, DC: National Academy of Sciences. National Academy Press.
- Van Lehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology* (47), 513-539.
- Verkoeijen, P., Imbos, T., Van de Wiel, M., Berger, M., & Schmidt, H. (2002). Assessing knowledge structures in a constructive statistical learning environment. *Journal of Statistics Education*, 10(2), <http://www.amstat.org/publications/jse/v10n12/verkoeijen.html>.
- Wilson, J. W., Fernández, M. L., & Hadaway, N. (1993). Mathematical problem solving. En P. S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics*. New York: MacMillan.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw Hill.
- Wonnacott, T & Wonnacott, R. (1999) *Introducción a la Estadística*. México: Ed Limusa.
- Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (2000). Building cognitive task analyses and models of a decision-making team in a complex real-time environment. En S. Chipman & Schraagen (Eds.), *Cognitive task analysis*. New York: Erlbaum.
- Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (En prensa). Cognitive task analysis and modeling of decision making in complex environments. En J. Cannon-Bowers & E. Salas (Eds.), *Decision making under stress: Implications for training and simulation*. Washington, DC: American Psychological Association.

APÉNDICE A. Guía de Preguntas para la Entrevista del Experto

Se empieza con una pregunta general para extraer los posibles subtemas para generar preguntas más específicas. A continuación se utilizará una (I) para indicar la las intervenciones del Investigador y una (E) para señalar las respuestas del Experto.

EJEMPLO:

I: ¿Qué es lo que hace que usted domine la *Selección de la Prueba Estadística*?

E: Yo se como _____

I: ¿La Selección de la Prueba Estadística, se puede dividir en varios pasos o componentes principales? ¿En caso afirmativo: ¿Cuáles son esos pasos o segmentos?

E: Los pasos que se siguen generalmente para la selección de la prueba estadística son: _____.

A continuación se realizan preguntas específicas para extraer conocimiento conceptual (CC) y conocimiento procedimental (CP). Así como el conocimiento simbólico (Bransford et al.), y conocimiento estratégico (CE) asociado, en cada caso.

CONOCIMIENTO PROCEDIMENTAL (CP)

I: ¿Cuál es el primer paso que sigue usted generalmente cuando _____?

I: Describa usted el paso _____.

I: ¿Cuándo realiza este paso?

I: ¿Siempre realiza este paso en el mismo orden?

Para indagar los pasos sucesivos se realiza la siguiente pregunta:

I: ¿Cuál es el siguiente paso que sigue usted generalmente cuando _____?

Se repiten las preguntas anteriores para extraer toda la información pertinente para ese paso. Adicionalmente se realizan las siguientes preguntas para indagar los siguientes tipos de conocimiento asociados:

CONOCIMIENTO CONCEPTUAL (CC)

-
-
- I: ¿Cuál es la definición de _____?
- I: ¿Qué se necesita saber para _____?
- I: ¿Cuál es la razón o razones para _____?
- I: ¿Cuáles son las condiciones que se deben dar para realizar este paso (_____)?
- I: ¿Por qué es importante realizar este paso (_____)?

En caso de que el experto introduzca un nuevo término en cualquier momento, se le pedirá que proporcione la definición correspondiente.

- I: ¿Cómo define este concepto?

CONOCIMIENTO ESTRATÉGICO (CE)

- I: ¿Existe alguna estrategia para _____?
- I: ¿Qué nombre le daría a esa estrategia?
- I: ¿Cuándo se utiliza esa estrategia?
- I: ¿En qué consiste esa estrategia?
- I: ¿Existe otra estrategia para _____?
- I: ¿Cuándo se utiliza esa otra estrategia?

CONOCIMIENTO SIMBÓLICO (Bransford et al.)

- I: ¿Existe alguna forma de representar ese paso (o concepto)?
- I: ¿Cómo representaría este paso?
- I: ¿Qué otra forma de representación existe?

Una vez agotadas las preguntas para el primer subtema expresado por el experto, se realiza nuevamente la pregunta general para indagar si existen otros subtemas en el dominio general de la tarea (en este caso la *Selección de la Prueba Estadística*).

- I: ¿Existe algo más que hace que domine la *Selección de la Prueba Estadística*?

E: Yo se como *identificar los elementos que componen la Selección de la Prueba Estadística*.

La respuesta se considera otro subtema, por lo tanto se deben generar nuevamente las preguntas correspondientes a conocimiento conceptual (CC), conocimiento procedimental (CP), conocimiento simbólico (Bransford et al.), y conocimiento estratégico (CE). Y así sucesivamente hasta agotar todos los subtemas generados por el experto.

APÉNDICE B. Lista de problemas estadísticos

T1

Prueba t (muestras relacionadas)

8 pacientes con alto nivel de estrés ingresaron a una psicoterapia y se midió su nivel de estrés antes y después de tres sesiones de psicoterapia. El propósito del estudio fue determinar si un nuevo tipo de psicoterapia breve tenía éxito después de tres sesiones. Se midió el estrés de los pacientes con la Escala del Inventario Eje II de Coolidge. Los resultados fueron los siguientes:

Paciente	Puntajes de Estrés	
	Antes de la Terapia	Después de la Terapia
1	68	63
2	58	60
3	74	65
4	55	62
5	81	54
6	59	73
7	47	45
8	75	73

(Coolidge, 2000)

T2

Prueba t (muestras relacionadas)

Un investigador desea determinar cuál de dos métodos de entrenamiento para la memoria funciona mejor con los mayores. Se seleccionaron para el estudio 8 gemelos de 65 años, la mitad fueron entrenados para memorizar una lista de palabras utilizando apoyos visuales y la otra mitad pasaron el mismo tiempo simplemente memorizando las palabras, 24 horas más tarde se midió en ambos grupos el número de palabras correctas recordadas (máximo de 20 correctas). A continuación se presentan los resultados:

Gemelos A	Número de palabras correctas		
	Apoyo visual	Gemelos B	Recuerdo simple
1	15	1	12
2	10	2	11
3	12	3	9
4	9	4	7
5	16	5	13
6	18	6	15
7	11	7	12
8	14	8	12

(Coolidge, 2000)

T3

Prueba t (muestras independientes)

Veinte miembros de un club local de Rotarios participaron en un experimento sobre los efectos del alcohol en el tiempo de reacción. Los sujetos fueron asignados al azar a dos grupos. A cada persona en el grupo de "Alcohol" se le dieron tres latas de ron con cola para consumir en un período de una hora; los miembros en el grupo de "No-Alcohol" se les dieron tres latas idénticas, pero que contenían cola mezclada con una bebida no alcohólica con sabor a ron. Se midió el tiempo de reacción en un aparato de simulación. Se desea saber si el alcohol produce una diferencia en el tiempo de reacción de los sujetos.

	Tiempo de reacción
Grupo Alcohol (N=10)	0.37, 0.42, 0.28, ..., 0.45
Grupo No-Alcohol (N=10)	0.29, 0.32, 0.37, ..., 0.22

(Wonnacott, y Wonnacott, 1999)

A1

ANOVA

Se realizó un experimento para probar si la Thorazina, el principal tranquilizador para la psicosis, afecta la habilidad para manejar. Los sujetos fueron 20 sujetos voluntarios. Se asignaron 5 estudiantes a cada una de las cuatro condiciones: a) placebo, b) 100 mg., c) 250 mg., d) 500 mg. Los puntajes obtenidos representan los resultados de una prueba de habilidades de manejo de 10 puntos (10= excelente, 1= pobre), la cual fue realizada una hora después de la ingestión de la droga. Los datos se presentan a continuación:

	Grupos			
	Placebo	100 mg	250 mg	500 mg
8	7	7	4	
9	9	6	5	
6	8	8	6	
9	10	7	5	
9	8	9	7	

(Coolidge, 2000)

A2

ANOVA

Un investigador desea saber si la iluminación ambiental afecta el tiempo de reacción. Para lo cual fueron asignados al azar, doce estudiantes a tres condiciones de un experimento de coordinación perceptivo-motor.

Tiempo de Reacción (en segundos)		
Luz brillante	Luz media	Luz baja
1.1,1.0,0.8,1.5	1.4,1.1,1.6,1.9	0.8,0.7,1.2,1.3

(Winer, Brown y Michels., 1991)

A3
ANOVA

El director de un aeropuerto regional está interesado en las posibles distracciones de los controladores de tráfico aéreo, cuando éste está sobrecargado. Supongamos que debido a su conocimiento experto en el área de investigación en atención y percepción de patrones visuales usted fue contratado como consultor. Se ha instalado en su laboratorio un simulador del ambiente de control de tráfico aéreo y usted tiene que desarrollar tres condiciones de entrenamiento diseñadas para aumentar la atención selectiva cuando el tráfico aéreo es pesado. Además de una cuarta condición – la condición control – en la cual no se administra tratamiento. Se seleccionó una muestra de 44 sujetos, asignados al azar a cada una de las condiciones. Se registraron las puntuaciones de cada uno de los sujetos al detectar potenciales colisiones durante la serie de eventos desplegados en la pantalla.

	Condiciones			
	1	2	3	4
	10.0	11.4	11.9	10.0
	10.6	10.7	12.6	9.7
	9.4	12.2	12.2	10.5
	9.8	11.8	13.2	10.3
Puntuación	11.0	11.7	12.9	9.9
	10.4	11.3	12.8	9.6
	10.4	11.5	11.8	10.1
	9.2	11.0	12.1	9.5
	10.2	10.9	12.4	10.2
	10.8	12.0	13.0	9.8
	9.6	11.1	12.5	10.4

(Winer, Brown y Michels., 1991)

C1
Correlación de Pearson

Un experimentador desea saber si el tabaquismo está relacionado con la longevidad. Se comparó el número de años de vida de una muestra de fumadores fallecidos recientemente con el número de cigarrillos consumidos por día (estimados por sus parientes durante los últimos 5 años).

Sujetos	Cigarros	Años vividos
1	25	63
2	35	68
3	10	72
4	40	62
5	85	65
6	75	46
7	60	51
8	45	60
9	50	55

(Coolidge, 2000)

C2

Correlación de Pearson

Se ha hipotetizado que el patrón de golpes por minuto de la música country-western que se toca en los bares se relaciona con el consumo de alcohol. Los observadores visitaron tres bares donde tocan música country-western durante tres viernes por la noche. Obtuvieron un permiso para grabar la música y hacer observaciones en las mesas seleccionadas. Cuando la música empezaba, se registraba la razón de sorbos de una bebida alcohólica para cada melodía y se obtenía la media de sorbos durante cada canción. Las grabaciones de música fueron analizadas de acuerdo al patrón (golpes por minuto) de cada melodía.

Patrón de golpes por minuto	Media de sorbos
35	1.150
38	1.150
44	0.400
...	...
112	0.750
118	0.625

(N=18)

(Wonnacott, y Wonnacott, 1999)

C3

Correlación de Spearman

Un profesor de estadística universitario desea investigar si existe relación entre el orden de finalización y la calificación de un examen final del curso. A los estudiantes se les instruyó a resolver el examen disponiendo de todo el tiempo que consideren necesario para resolverlo. El profesor registró el orden en que cada uno de los 16 estudiantes completó su examen y posteriormente anotó la calificación obtenida por cada estudiante.

E= Excelente, MB= Muy Bien, B= Bien, S= Suficiente y N/A= No Acreditado

Estudiante	1	2	3	4	...	5	6
Orden de Finalización	3	5	1	8	...	14	10
Calificación	MB	E	MB	B	...	N/A	S

(N=16)

(Winer, Brown y Michels., 1991)

X1

Prueba Ji cuadrada de independencia entre 2 variables

Existe un decreto que propone reducir la edad legal para permitir consumir bebidas alcohólicas a los 18 años, el cual está en discusión en la legislatura local. Un politólogo que vive en esa ciudad, quiere determinar si existe relación entre la afiliación política y la actitud hacia el decreto. Este científico social envía cartas a dos muestras aleatorias; una de 200 miembros al partido Republicano y otra de igual número al Partido Demócrata; el científico pregunta a las personas si están a favor, indecisas o en contra de esta iniciativa de ley. El investigador recibe un total de 400 respuestas entre republicanos y demócratas.

(Pagano, 1999)

X2

Prueba Ji cuadrada de independencia entre 2 variables

La siguiente tabla muestra los resultados de probar dos vacunas para prevenir la fiebre tifoidea, en una ciudad mexicana. Se administró polisacárido capsular (Vi) o polisacárido neumococo (Neumococo). Después de aplicadas las vacunas 71 sujetos desarrollaron síntomas de fiebre tifoidea. Se desea determinar si alguna de las dos vacunas es más efectiva que la otra para prevenir la fiebre tifoidea.

Resultados	Tipo de Vacuna		Total
	Vi	Neumococo	
Con fiebre tifoidea	14	57	71
Sin fiebre tifoidea	3457	3393	6836
Total	3457	3450	6907

(Marques, 2001)

X3

Prueba Ji cuadrada de independencia entre 2 variables

Un psicólogo prueba la hipótesis de que la gente que está fatigada disminuye su estado de alerta mental. Se aplica una prueba para medir la atención, que consiste en que los sujetos se sienten frente al televisor y presionen un botón cada vez que

aparezca un punto en la pantalla. Se presentaron un total de 110 puntos durante un período de 90 minutos, el psicólogo registra el número de errores para cada sujeto. Se seleccionaron veinte sujetos; la mitad fue evaluada después de 24 horas sin dormir y la otra mitad fue evaluada por la mañana después de dormir toda la noche. Cada sujeto fue etiquetado como alto o bajo estado de alerta mental con base en el número de errores cometidos en la prueba.

(Quillici y Mayer, 1996).

APÉNDICE C. Solución de un problema estadístico mediante la aplicación de ANOVA

Se condujo una investigación en una Universidad en la que se trataba de estudiar el efecto de cursar materias de alto índice de reprobación sobre el nivel de ansiedad de los alumnos. Los registros mostraban que tanto Anatómo-Fisiología como Estadística eran materias frecuentemente reprobadas por los estudiantes. Por ello, se escogieron tres muestras aleatorias; la primera de alumnos que entre sus materias cursaban Estadística, una más de aquellos estudiantes que cursaban Anatómo-Fisiología y una tercera de aquellos que cursaban otras materias, pero ninguna de las dos mencionadas, con el propósito de comparar sus niveles de ansiedad. A los estudiantes se les aplicó una prueba de ansiedad estandarizada, cuyos reactivos se respondían mediante una escala Likert, obteniéndose los resultados de la tabla siguiente:

Con Estadística	Con Anatómo-Fisiología	Sin ninguna de las dos
34	43	32
25	51	50
36	55	44
25	45	34
31	57	46
27	59	35
28	64	30

	Sum of Squares	df	Mean Square	F	Sig
Between Groups	2050.381	2	1025.190	22.426	.000
Within Groups	822.857	18	45.714		
Total	2873.238	20			

(Wonnacott, y Wonnacott, 1999)

APÉNDICE D Características generales de los doce problemas presentados.

CLAVE DEL PROBLEMA	ESTADÍSTICO DE PRUEBA	PROPÓSITO DEL ESTUDIO	VARIABLES	NIVEL DE MEDICIÓN	NO. DE GRUPOS O COND.	TAMAÑO DE LA MUESTRA
T1	Prueba t (muestras relacionadas)	Diferencias	VI: Psicoterapia (antes-después) VD: Nivel de estrés (escala)	VI: Nominal VD: Intervalar	2	Pre =8 Post =8 Total =8
T2	Prueba t (muestras independientes)	Diferencias	VI: Método de entrenamiento para la memoria VD: Num. De palabras recordadas	VI: Nominal VD: Intervalar	2	Gem. A=8 Gem. B=8 Total=16
T3	Prueba t (muestras relacionadas)	Diferencias	VI: Consumo de alcohol (si/no) VD: Tiempo de reacción (seg.)	VI: Nominal VD: Intervalar	2	Alcohol=10 No Alcohol= 10 Total=20
A1	ANOVA	Diferencias	VI: Dosis de Thorazina VD: Habilidad para manejar (prueba de 10 puntos)	VI: Nominal VD: Intervalar	4	Plac.=5 100mg=5 250mg=5 500mg=5 Total=20
A2	ANOVA	Diferencias	VI: Iluminación ambiental VD: Tiempo de reacción (seg.)	VI: Nominal VD: Intervalar	3	Brillante= 4 Media=4 Baja=4 Total=12
A3	ANOVA	Diferencias	VI: Tipo de entrenamiento VD: Detección de posibles colisiones	VI: Nominal VD: Intervalar	4	C.1=11 C.2=11 C.3=11 C.4=11 Total=44
C1	Correlación de Pearson	Relaciones	VI: Tabaquismo (No. De cigarrillos) VD: Longevidad (años vividos)	VI: Intervalar VD: Intervalar	1	Fumadores=8
C2	Correlación de Pearson	Relaciones	VI: Patrón de golpes por minuto (música Country-Western) VD: Consumo de alcohol (cantidad de sorbos)	VI: Intervalar VD: Intervalar	1	Melodía=18
C3	Correlación de Spearman	Relaciones	VI: Orden de finalización	VI: Ordinal	1	Estudiantes=16

			VD: Calif. de un examen (MB, B, S, N/A).	VD: Ordinal		
X1	Ji cuadrada	Relaciones	VI: Afiliación política VD: Actitud hacia el decreto	VI: Nominal VD: Nominal	1	Partidos Repub.=200 Democ.=200 Total=400
X2	Ji cuadrada	Relaciones	VI: Tipo de vacuna (Vi/Neumococo) VD: Síntomas de fiebre tifoidea (si/no)	VI: Nominal VD: Nominal	1	Vacuna Fiebre+Vi=14 Fiebre+Neu.=57 S/Fiebre+Vi=3457 S/Fiebre+Neu.=3393 Total=6907
X3	Ji cuadrada	Relaciones	VI: Fatiga (dormir/sin dormir) VD: Estado de Alerta (alto/bajo)	VI: Nominal VD: Nominal	1	S/Dormir+Alto=5 S/Dormir+Bajo=5 Dormir+Alto=5 Dormir+Bajo=5 Total=20

NOTA. Se asignó una clave a cada problema para su posterior identificación en el análisis. T: Prueba t, A: ANOVA, C: Correlación y X: Ji Cuadrada.

APÉNDICE E Taxonomía de relaciones semánticas, adaptada de Lemke, (1997).

RELACIONES SEMÁNTICAS		
CLAVE	TIPO DE RELACIÓN	RELACIONES NOMINALES
ATR	ATRIBUTIVA	Una característica descriptiva, cualidad
CLAS	CLASIFICADORA	Un tipo o clase de; una característica que identifica a una subclase
RELACIONES TAXONÓMICAS		
TIP	TIPO	Un ejemplo individual de tipo o clase
PART	PARTE	Nombre de una parte que pertenece a un todo
SIN	SINÓNIMO	Par equivalente, sinónimos
ANT	ANTÓNIMO	Par contrastante, antónimos, contrastes
RELACIONES DE TRANSITIVIDAD		
AG	AGENTE	La entidad que hace o actúa; la causa o el instigador de un proceso
MET	META	La entidad a la que se hace, sobre la que se actúa, el objeto de la acción
MED	MEDIO	La entidad en relación a la cuál sucede un proceso
IDEN	IDENTIFICACIÓN	Lo que identifica un proceso, evento o entidad
RELACIONES CIRCUNSTANCIALES		
LOC	LOCALIZACIÓN	Expresa la relación espacial de entidades o procesos
TIEM	TIEMPO	Expresa la relación temporal de procesos, eventos, entidades
MAN	MANERA	Como, de qué manera y por qué medios/instrumentos ocurrió el proceso
RELACIONES LÓGICAS		
ADIC	ADICIÓN	A y B
VAR	VARIACIÓN	A o B
CONEX	CONEXIÓN	Una categoría miscelánea que incluye la relación de las partes de varias formas de argumentación

APÉNDICE F Conocimiento declarativo, procedimental y estratégico involucrado en la selección de la prueba estadística

Tabla a)

Conocimiento declarativo involucrado en la selección de la prueba estadística

Conocimiento Declarativo
• Objetivo o Propósito de Investigación
• Comparación o asociación
• Hipótesis
• Hipótesis de Trabajo
• Hipótesis Estadística
• Hipótesis Nula
• Hipótesis Alterna
• Variable Independiente
• Variable Dependiente
• Nivel de Medición de las Variables
• Tipo de Muestreo
• Diseños de Investigación
• Número de Muestras o Grupos
• Muestras Independientes o Relacionadas
• Tamaño de la Muestra
• Tipos de Instrumento
• Distribución Normal
• Error Tipo I y Tipo II
• Probabilidad
• Curva normal
• Teorema de Límite Central
• Medición
• Aritmética
• Independencia
• Homogeneidad de varianzas
• Varianza
• Normalidad
• Linealidad

Tabla b)

Conocimiento procedimental involucrado en la selección de la prueba estadística.

Conocimiento Procedimental

- Planteamiento de Objetivo o propósito de investigación
- Planteamiento de hipótesis
- Aplicar un diseño de investigación a un problema
- Elaboración de un histograma
- Elaboración de un polígono de frecuencias
- Obtención de medias
- Obtención de rangos
- Aplicación de pruebas estadísticas mediante SPSS (ANOVA, Prueba t, KruskalWallis, etc.)
- Aplicación de procedimientos estadísticos para analizar el cumplimiento de supuestos: normalidad (Shapiro Wilk)
- Elaboración de instrumentos de medición

Tabla c)

Conocimiento estratégico involucrado en la selección de la prueba estadística.

Conocimiento Estratégico

- La distribución normal se basa mínimamente en 30 sujetos
- Las variables están asociadas al instrumento
- Si tengo tres grupos, entonces lo que hago es comparar, aplico una prueba de comparación
- En la escala Likert los puntajes individuales se asumen como ordinales, pero la suma de puntuaciones se puede trabajar como intervalares
- Asumo que hay independencia, porque los que están en un grupo, no pueden estar en el otro
- Es importante indicar para poder interpretar de dónde a dónde va la escala, es decir, si los puntajes mínimos se refieren a un nivel bajo de la variable y puntajes altos se refieren a un nivel alto de la variable, porque también puede ser al revés
- Si es grande el valor "F" implica que hay diferencias entre los grupos
- Las medias cuadráticas indican qué tan alejadas están las medias de los grupos de la media global
- Todos los estadísticos que tienden a cero, te hablan de que lo que observas, concuerda con lo que esperas encontrar
- En general la prueba de hipótesis consiste en verificar que lo que observas, concuerda con lo que se esperaría encontrar teóricamente o con lo que se esperaría encontrar por

azar, mientras más se acerquen a cero hay menor discrepancias, mientras más se alejen de cero hay mayor discrepancia

- Para saber en qué par están las diferencias, hay que hacer comparaciones múltiples para definir en qué pares de grupos están las diferencias, en SPSS cuando haces el Análisis de Varianza hay un módulo post hoc
- Lo que te indica si son diferentes los grupos, es el tipo de muestra que vas a usar, entonces tienes que identificar si las muestras son independientes o son relacionadas
- Si tengo el mismo número de sujetos y los voy a medir antes y después, tengo a la misma persona, es decir, muestras relacionadas
- Cuando me hablas de sujetos que emparejaste en edad, sexo, religión, etcétera, se trata de muestras emparejadas; cuando se me trata de la misma persona medida antes y después, son muestras relacionadas
- Tenemos que identificar las variables, como propiedades que quiero medir, diferenciándolas de las características de los sujetos u objetos
- Si vemos rangos de edad 18 a 20, 21 a 25 y 26 a 30, entonces estoy poniendo a la edad como variable y no como característica
- En el caso de los problemas de comparación, forzosamente necesito saber cuál es la variable independiente y la variable dependiente
- La variable independiente es la que te va a permitir hacer la clasificación de grupos
- El objetivo de investigación es el que me permite identificar si es un problema de asociación o de comparación
- Las palabras claves para asociación son : se relaciona con, se asocia
- Las palabras claves para identificar una comparación en un problema estadístico son : quiero saber si hay diferencias entre X y Y; en los casos en que se mencionan grupos de contraste; cuando la variable independiente se divide en varios rangos o niveles; los valores de la variable independiente son clasificatorios
- Puedes tener datos intervalares, pero si no hay normalidad, lo más conveniente es un análisis no paramétrico
- Aunque tengas normalidad, si tu escala es ordinal, lo más conveniente es un análisis no paramétrico
- Para saber si hay normalidad, se puede aplicar la prueba de Shapiro-Wilk, y también con las gráficas que se obtienen al solicitar la prueba ANOVA en SPSS
- Cuando los tamaños de "n" son muy grandes, de 100 en adelante, te dan resultados similares en estadísticas paramétricas y no paramétricas
- La varianza habla de discrepancias, me indica qué tan alejados están los datos con respecto a la media y el dato que tengo es el promedio de esa desviación
- La homogeneidad de varianzas sirve para comparar dos o más grupos, estos deben estar iguales, en términos de la desviación para poder decir que son comparables y así poder atribuir las diferencias a los efectos de la variable independiente; esto lo aplicas antes de aplicar un tratamiento
- Para hacer equivalentes a los grupos, puedes dejar que el azar decida dónde van a estar las diferencias individuales, esto es una forma de hacer equivalentes los grupos

Apéndice G Errores conceptuales en que incurren los novatos en el proceso de elección de la prueba estadística, agrupados por bloques temáticos. Se muestra el nivel de pericia y un extracto de la entrevista realizada. (N= Novato, I= Investigador).

Nivel	Diseño Experimental Control de Variables
1 Doct.	<p>I: Entonces digamos que los separaste en tres y...Ahora ¿me puedes explicar porqué los separaste?</p> <p>N: Si.</p> <p>I: ¿Con qué criterios?</p> <p>N: Si, mira. En éstos (<i>señala un grupo</i>) los seleccioné porque, de acuerdo a la finalidad del problema. Entonces yo lo puedo clasificar en experimento, en el que hay un control de variables.</p> <p>I: ¿Control de variables?</p> <p>N: Control de variables con, que está dada de una serie de sustancias, terapia, alguien se apoya en tarjetas visuales, otro que nada más utiliza palabras. Aquí aplican vacuna, aquí aplican alcohol, a otros no. Aquí aplican mousse brillante y aquí aplican torazina. Están manipulando una o más variables sobre varios sujetos o grupos. Y en ésta solamente los están observando en ambientes, están observando a los sujetos, si están desarrollando o no. no aplican nada,</p> <p>I: O sea ¿aquí más bien sería no hay manipulación, si hay manipulación?</p> <p>N: Si.</p>
1 Maest.	<p>I: ¿Por qué crees que deberían ser sujetos diferentes y no los mismos?</p> <p>N: Porque así hay mayor control de las variables, porque...a ver...déjame pensarle bien... cuando son sujetos diferentes me parece que la comparación es mejor, y el control de las variables, o sea de la variable independiente sobre la variable dependiente se controla mejor</p>

Nivel	Estrategias de Recolección de Datos Nivel de Medición Instrumentos de Medición Formación de Grupos de Comparación
1 Doct.	<p>N: En primera no están cumpliendo... Para hacer una prueba paramétrica, no, porque no cubre los requisitos, para empezar. Son solamente siete. ¿Qué más?... el nivel es nominal porque la escala es tipo Likert, podría meterlo en una prueba ordinal, pero el problema es la cantidad de sujetos: son muy poquitos.</p>
1 Doct.	<p>I: Nada más estás sacando de las variables dependientes.</p> <p>N: Ahorita sí. De hecho ya me empecé a cuestionar esto de la variable independiente en este tipo de problemas.</p> <p>I: Mjm. ¿Qué es lo que te cuestionaste?</p> <p>N: Me dije, bueno, son grupos entonces...este...entonces lo que se me viene a la mente ¿son nominales? ¿es nominal la cosa o cómo? Esto ya se esta convirtiendo en un experiencia de aprendizaje,</p>
1 Doct.	<p>I: ¿Qué te dice que este dato?</p> <p>N: La escala Likert ¿no? Cuando tienes una escala Likert, automáticamente ya no es nominal...este...ni ordinal. Es que ahí esta una cuestión ¿no? Una polémica que incluso de planteaba en los últimos, en el último curso de</p>

	<p>maestría. Es decir, en una escala Likert tú puedes asumir que los intervalos son similares ¿no?, hay intervalos similares o no los hay. Entonces quedó ahí como que bueno, hay un grupo de personas que piensan que sí, hay otros que piensan que no. Yo prefiero pensar que es intervalar.</p> <p>I: Lo que pasa es aquí ya es cuestión de criterio.</p> <p>N: Este..sí, y bueno, me parece que en la practica...este...en la práctica se toma como intervalares.</p>
1 Maest.	<p>I: ¿A qué te refieres?</p> <p>N: A lo que viene siendo la escala de medición. Por la escala Likert, qué tipo de medición me está dando para poder...</p> <p>I: ¿Qué tipo de medición te da?</p> <p>N: ¿Ésta?, creo que es ordinal. No sé. Ahí me confundo mucho con la escala Likert, porque algunos maestros me han dicho que es ordinal, y otros que es intervalar. Ahí, luego, si me..., eso si me confunde mucho.</p>

Nivel	Propósito del Estudio Relación entre Variables/Comparación entre Grupos o Condiciones
1 Doct.	<p>N: Lo que estoy seguro es de la aplicación de una prueba paramétrica, qué digo, no paramétrica, pues para grupos pequeños, con un tipo de nivel nominal y a la vez ordinal. Si aplicamos r de Pearson... Tengo curiosidad, no me acuerdo cuál busca diferencias y cuál busca... tendría que checar... ¿Pearson? ¿La r de Pearson? Se me hace que la r de Pearson busca relaciones. Creo que la aplicaría para saber si hay diferencias entre los grupos y obtener...</p>
1 Doct.	<p>N: Luego, luego se ve por la puntuación: Anatomía y Fisiología. Es mayor que éste y éste. Al menos por la medidas de Tendencia Central. Con ellas sería suficiente. Ahora, para certificarlo ya, sería, bueno aplicar... te digo, ¡es Spearman! Para ver si hay diferencias entre los grupos.</p>
1 Doct.	<p>N: ¡Ah! Ya empiezo a tener dudas (risas).</p> <p>I: ¿Porqué?</p> <p>N: Porque entre éstas..., porque se parecen (señalando) ésta, ésta y ésta; en cuanto a que el consumo de alguna...eh...pues si de alguna..., bueno, básicamente en ésta. Éstas dos, es el consumo de algo que afecta alguna habilidad o una capacidad. Y son, este, pues son los efectos que pueda tener este consumo o no consumo. Pero, éste lo había visto, como que es una relación..., es que por la palabra <u>afecta</u> yo había pensado en algo, en esto que yo decía: esto tiene, va a causar algo. Pero ahorita ya me pregunté si esto de <u>afecta</u> y <u>efecto</u>, puede ser equiparable. O sea si es lo mismo en cuanto a la relación que se quiere evaluar. Y entonces...ya no sé.</p>
1 Doct.	<p>I: Bueno, entonces ¿cómo llegaste a esa decisión?</p> <p>N: Eh... Bueno, lo primero que vi, es si tenía que ver con..., pues si era una relación entre variables o si algo afectaba, causa efecto, vamos. Pues, primero la relación que se está planteando aquí y después el diseño lo que me hace pensar en el tipo de prueba. Aunque ya después, esto de efecto como que me causó un poco de ruido.</p>
1 Doct.	<p>I: ¿Qué otra cosa puedes ver de...?</p> <p>N: Viendo esto también se puede saber, si el modelo está contemplando la diferencia entre grupos, o está haciendo un modelo general de la población, algo así, tengo como la idea, pero no tengo claro.</p>
2 Tesis	<p>N: Es un estudio comparativo. Es que se están comparando los resultados en cuanto a los tratamientos. Tal vez haya la necesidad de un análisis, tal vez correlacional de las respuestas en una posterior etapa, pero por lo pronto</p>

	<p>en un estudio de comparación, comparativo se podría decir. <i>Cae dentro de los comparativos.</i> Aunque el nivel de medición los situaría del de razón ¿no? Si, creo que si, el de razón.</p>
3 Lic.	<p>N: Es que... estoy confundida. I: ¿Por qué? N: Porque se supone que <i>la hipótesis de estudio es... saber si el tabaquismo está relacionado con el antiviral, pero como dice se comparó, eso me confundió un poco.</i></p>

Nivel	Tamaño de la Muestra Inferencia estadística
1 Doct.	<p>I: ¿Y qué otra cosa tomarías en cuenta para decidir la prueba estadística aplicar? N: La prueba...este...pues esto que ya automáticamente lo pienso como diferencias entre grupos... Está interesado en indagar sobre relaciones de causa – efecto. Entonces ahí hay una serie de estadísticos ¿no?, que son los más apropiados para...este...creo que la cantidad, ¿no?. <i>La cantidad de sujetos que tienen, pero la verdad, no sabría ahorita, decidir en función de eso. Esta es la cantidad...la n que tienes me parece que también es importante, pero te digo, no. No recuerdo particularmente.</i></p>
1 Maest.	<p>N: Porque bueno, lo primero que tendría que tomar en cuenta es este, la... ¡ah!, bueno, algo también importante, el número de alumnos que fueron..., el <i>número total de alumnos.</i> I: ¿Por qué sería importante eso? N: Pues, porque también nos da una pauta para decidir qué tipo de estadística utilizar para analizar los datos. I: Bueno. N: Por lo menos, eso fue lo que me enseñaron mis maestros, que <i>si es importante saber el número de la muestra, qué tan grande o pequeña es, la manera en que...</i> I: ¿Qué informaciones te hubiera dado la muestra? N: Deja me acuerdo: pues que tanto se puede, de alguna manera, <i>no sé si generalizar el efecto que está teniendo este tipo de materias en la ansiedad no, de los alumnos para poder saber este..., si puede ser generalizable a toda la población.</i></p>
2 Tesis	<p>N: Siete alumnos por cada grupo. Ah, pues el tamaño de la muestra, también... I: ¿Y eso que implicaría? N: Porque, creo que, bueno, no me acuerdo bien, pero <i>si hay muestras pequeñas se utilizan algunas pruebas,</i> bueno, un tipo, si un tipo de pruebas; y si son más datos, se utilizan, si son más de... no sé... treinta datos, se utilizan otro tipo de pruebas. <i>Mientras más chica sea la prueba, o la muestra, perdón, la muestra, se utiliza un tipo de estadística, y si es más grande se utiliza otro tipo de estadística.</i></p>

Nivel	Tipo de Muestreo Probabilístico/No Probabilístico Muestra Aleatoria
2 Tesis	<p>N: ¿Y porqué de las muestras? Republicanos y demócratas; aleatorios porque siempre se sabe que..., pues que, cuando tu agarras <i>una muestra aleatoria te va a dar mayores resultados y menos contaminados que una muestra que no</i></p>

	lo es.
2 Tesis	N: Son muestras aleatorias, entonces supongo que el muestreo es de tipo probabilístico , y si no mal recuerdo porque no estoy muy segura, este sí el muestreo es probabilístico, bueno había una clasificación de las pruebas que son paramétricas y las no paramétricas y me parece que el muestreo fuera probabilístico tenía que ver con si podías escoger una prueba paramétrica o no paramétrica , me parece que las no paramétricas son las que, no al revés, la paramétricas son las que necesitan un muestreo de tipo probabilístico , o sea hay como ciertos requisitos que se tienen que cubrir ese sería uno.

Nivel	Interpretación de Resultados
1 Doct.	<p>I: Bueno, pues ahorita esto es el output con el resultado de este problema. Bueno, aquí ¿qué te dice este resultado? ¿Qué es lo que puedes obtener de él?</p> <p>N: Pues bueno, lo que veo es que, efectivamente, hay una diferencia entre los grupos. No sé exactamente entre cuáles, pero sé que hay una diferencia y también esto es algo como que apenas estoy empezando a entender. Todavía no lo comprendo muy bien, entre éstos puntajes que obtenemos aquí que tiene que ver con la cantidad de varianza que estamos explicando ¿no?</p>
3 Lic.	<p>I: Bueno, aquí tengo, es un resultado del SPSS y...aquí te muestran, precisamente los resultados de ésta prueba, de éste problema. Y de éste... ¿qué te dice éste SPSS? ¿Qué representa ese resultado? ¿Qué se puede interpretar de ahí?</p> <p>N: No, jamás he manejado SPSS. Supongo que aquí agruparon dos grupos.</p>

Nivel	Decisión Estadística Estadístico de Prueba Nivel de Significancia Regla de decisión Análisis Post Hoc
1 Doct.	<p>N: Pero sé que ésta información me da el total de varianza que yo explico y que si es significativa, la diferencia entre los grupos. Y sí efectivamente, hay una diferencia, pero no se puede saber entre qué grupos están las diferencias, habría que hacer un análisis adicional.</p> <p>I: ¿Sabes que análisis podría ser?</p> <p>N: Pues el de..., o bueno, hay varios, pero también depende de la exigencia de cada uno y que también tiene que ver, bueno, con otras variables. Sé que tiene que ver, tiene su justificación, cada uno de ellos, pero no los tengo claros. Sé que hay unas más exigentes que otras.</p>
1 Doct.	<p>I: ¿Y qué podrías concluir acerca de este problema con los resultados?</p> <p>N: Ah, ok. Que si influye la materia ¿no? Es que no recuerdo, obviamente aquí lo que...bueno, lo siguiente es saber pues en dónde está el mayor índice de, el mayor nivel de ansiedad. No me acuerdo si aplicas una prueba posterior para saber en dónde están las diferencias. Porque creo que también sacas la media, y bueno, ahí puedes darte una idea, pero no me acuerdo si en este caso sacas una prueba posterior para ver en dónde están, este, las diferencias ¿no? Porque éste (refiriéndose al output) es general. O sea hay diferencias entre los grupos, pero no sabría, necesitaría yo otro...</p>

<p>1 Maest.</p>	<p>I: ¿Qué es lo que estás viendo en el problema? N: Busca hacer una comparación, se valora mediante una escala tipo Likert, y pues ya, no sé que prueba estadística utilizar.</p>
<p>1 Maest.</p>	<p>I: Ese es el resultado de este problema. Es el resultado del SPSS ¿Qué crees que sea lo más importante? N: Pues, según yo, la media, y ¿Éste es el nivel de significancia, no? I: Si. N: Y el nivel de significancia. I: Y eso ¿Qué te dice?, ¿Qué te dice ese valor? N: Así como que, me imagino que cuál es la diferencia entre los grupos, qué tan grande es esa diferencia, o si hay más bien diferencia entre los grupos. I: Y eso ¿Qué te dice? N: Y ese valor, pues no me acuerdo. Porque yo nada más de acuerdo que es .05, .01 y punto nada más. I: Pero ¿Cómo tomas en cuenta ese .05, .01 con respecto a esto? ¿Te acuerdas? N: No.</p>
<p>2 Tesis</p>	<p>I: Bueno este es el Output de ese problema que es lo que te dice allí N: Una vez que haces la prueba, creo que tienes que..., de acuerdo a lo que te arroja lo de la prueba para tomar la decisión, de si rechazas o no rechazas la hipótesis nula, hay una regla de decisión, que tampoco me la sé. Hay una regla de decisión que no me la sé, y bueno que tiene que ver, si no mal recuerdo, con el intervalo de confianza que se establece para ver esto de la región de rechazo y aceptación de la hipótesis nula, que hay como varios ¿no?... de 5 % o 0.5% una cosa así, y entonces eso es lo que utilizas como para ver si es una regla si es menor o mayor a no sé que cosa a esto el nivel de significancia entonces ya dices si rechazas o no rechazas la hipótesis nula pero como la verdad no me acuerdo de la regla de decisión bien.</p>
<p>3 Lic.</p>	<p>I: ¿Eso qué te dice? N: El grado de significancia. Es con el que se tiene que comparar para aceptar o rechazar la hipótesis ¿no? para ver si es débil o no. I: ¿Y qué te dice que la debes aceptar o rechazar? N: Mayor o igual (risas), ay no sé.</p>