

z	Una cola		Dos colas		z	Una cola		Dos colas	
	$\pi$	$\pi$	$\pi$	$\pi$		$\pi$	$\pi$	$\pi$	$\pi$
	frontera	resto	frontera	resto		frontera	resto	frontera	resto
0,88	0,1894	0,8106	0,3789	0,6211	1,32	0,0934	0,9066	0,1868	0,8132
0,89	0,1867	0,8133	0,3735	0,6265	1,33	0,0918	0,9082	0,1835	0,8165
0,90	0,1841	0,8159	0,3681	0,6319	1,34	0,0901	0,9099	0,1802	0,8198
0,91	0,1814	0,8186	0,3628	0,6372	1,35	0,0885	0,9115	0,1770	0,8230
0,92	0,1788	0,8212	0,3576	0,6424	1,36	0,0869	0,9131	0,1738	0,8262
0,93	0,1762	0,8238	0,3524	0,6476	1,37	0,0853	0,9147	0,1707	0,8293
0,94	0,1736	0,8264	0,3472	0,6528	1,38	0,0838	0,9162	0,1676	0,8324
0,95	0,1711	0,8289	0,3421	0,6579	1,39	0,0823	0,9177	0,1645	0,8355
0,96	0,1685	0,8315	0,3371	0,6629	1,40	0,0808	0,9192	0,1615	0,8385
0,97	0,1660	0,8340	0,3320	0,6680	1,41	0,0793	0,9207	0,1585	0,8415
0,98	0,1635	0,8365	0,3271	0,6729	1,42	0,0778	0,9222	0,1558	0,8444
0,99	0,1611	0,8389	0,3222	0,6778	1,43	0,0764	0,9236	0,1527	0,8473
1,00	0,1587	0,8413	0,3173	0,6827	1,44	0,0749	0,9251	0,1499	0,8501
1,01	0,1562	0,8438	0,3125	0,6875	1,45	0,0735	0,9265	0,1471	0,8529
1,02	0,1539	0,8461	0,3077	0,6923	1,46	0,0721	0,9279	0,1443	0,8567
1,03	0,1515	0,8485	0,3030	0,6970	1,47	0,0708	0,9292	0,1416	0,8584
1,04	0,1492	0,8508	0,2983	0,7017	1,48	0,0694	0,9306	0,1389	0,8611
1,05	0,1469	0,8531	0,2937	0,7063	1,49	0,0681	0,9319	0,1362	0,8638
1,06	0,1446	0,8554	0,2891	0,7109	1,50	0,0668	0,9332	0,1336	0,8664
1,07	0,1423	0,8577	0,2846	0,7154	1,51	0,0655	0,9345	0,1310	0,8690
1,08	0,1401	0,8599	0,2801	0,7199	1,52	0,0643	0,9357	0,1285	0,8715
1,09	0,1379	0,8621	0,2757	0,7243	1,53	0,0630	0,9370	0,1260	0,8740
1,10	0,1357	0,8643	0,2713	0,7287	1,54	0,0618	0,9382	0,1236	0,8764
1,11	0,1335	0,8665	0,2670	0,7330	1,55	0,0606	0,9394	0,1211	0,8789
1,12	0,1314	0,8686	0,2627	0,7373	1,56	0,0594	0,9406	0,1188	0,8812
1,13	0,1292	0,8708	0,2585	0,7415	1,57	0,0582	0,9418	0,1164	0,8836
1,14	0,1271	0,8729	0,2543	0,7457	1,58	0,0571	0,9429	0,1141	0,8859
1,15	0,1251	0,8749	0,2501	0,7499	1,59	0,0559	0,9441	0,1118	0,8882
1,16	0,1230	0,8770	0,2460	0,7540	1,60	0,0548	0,9452	0,1096	0,8904
1,17	0,1210	0,8790	0,2420	0,7580	1,61	0,0537	0,9463	0,1074	0,8926
1,18	0,1190	0,8810	0,2380	0,7620	1,62	0,0526	0,9474	0,1052	0,8948
1,19	0,1170	0,8830	0,2340	0,7660	1,63	0,0516	0,9484	0,1031	0,8969
1,20	0,1151	0,8849	0,2301	0,7699	1,64	0,0505	0,9495	0,1010	0,8990
1,21	0,1131	0,8869	0,2263	0,7737	1,645	0,05	0,95	0,10	0,90
1,22	0,1112	0,8888	0,2225	0,7775	1,65	0,0495	0,9505	0,0989	0,9011
1,23	0,1093	0,8907	0,2187	0,7813	1,66	0,0485	0,9515	0,0969	0,9031
1,24	0,1075	0,8925	0,2150	0,7890	1,67	0,0475	0,9525	0,0949	0,9051
1,25	0,1056	0,8944	0,2113	0,7887	1,68	0,0465	0,9535	0,0930	0,9070
1,26	0,1038	0,8962	0,2077	0,7923	1,69	0,0455	0,9545	0,0910	0,9090
1,27	0,1020	0,8980	0,2041	0,7959	1,70	0,0446	0,9554	0,0891	0,9109
1,28	0,1003	0,8997	0,2005	0,7995	1,71	0,0436	0,9564	0,0873	0,9127
1,282	0,10	0,90	0,20	0,80	1,72	0,0427	0,9573	0,0854	0,9146
1,29	0,0985	0,9015	0,1971	0,8029	1,73	0,0418	0,9582	0,0836	0,9164
1,30	0,0968	0,9032	0,1936	0,8064	1,74	0,0409	0,9591	0,0819	0,9181
1,31	0,0951	0,9049	0,1902	0,8098	1,75	0,0401	0,9599	0,0801	0,9199

(continúa)

blanca que para los de raza negra. Los autores (Cronk y Roches, 1982; Harsha *al.*, 1974; Johnston *al.*, 1974) afirman que hay diferencias en el grosor del pliegue cutáneo entre los niños de raza negra y blanca, aunque los patrones de la AAHPERD no establecen distinciones. Si los patrones son válidos, deben ser representativos de la población que se está estudiando. Además, puede haber diferencias mayores en la adiposidad entre los de raza negra y blanca en distintas edades.

De manera parecida es posible que una combinación distinta de pliegues cutáneos sea más válida para las niñas que para los niños. No parece ser una gran ventaja en las mediciones utilizar el mismo lugar para ambos sexos si hay otros puntos que sean indicadores de la adiposidad total con igual validez.

Unas palabras finales de advertencia. Cuando te formulan la inevitable pregunta, durante la lectura de la tesis, sobre la importancia del estudio, no contestes “Ha sido necesario para obtener el título”. El silencio sepulcral será crispante.

## Diferencias entre una tesis y un artículo de investigación

Una ojeada a los artículos de investigación de las revistas demuestran que han desaparecido algunos de los apartados que se encuentran en las tesis o tesinas, descritos en este capítulo. Hay por lo menos dos motivos que explican esto. El primero es económico: las revistas están sujetas a unos costes de publicación por lo que exigen brevedad. Segundo, parece que hay algún tipo de ritual entre noveles y expertos. Se exige a los noveles que, de forma explícita, formulen las hipótesis, definan los términos, establezcan los supuestos, identifiquen las limitaciones y justifiquen el motivo del estudio que están redactando. En realidad, estos pasos forman parte de la definición y delimitación del problema a investigar, y, sin duda, representa un gran trabajo expresar formalmente cada uno de los puntos.

Por otro lado, los autores de artículos de investigación no necesitan explicar, paso a paso, el método utilizado para estudiar el problema. Normalmente, un artículo de investigación tiene una introducción que incluye una breve revisión bibliográfica. La extensión puede variar mucho y algunas revistas insisten en la brevedad de las introducciones.

Casi siempre se especifica el objetivo del estudio aunque no suele incluirse en un subapartado; en vez de esto, acostumbra a ser la última frase de la introducción. Por ejemplo, en 30 artículos de un volumen de la *Research Quarterly for Exercise and Sport*, sólo uno tenía un subapartado titulado “Objetivo del estudio”. Veinticuatro finalizaron la introducción con frases que empiezan con las palabras, “El objetivo de este estudio es...”. Cuatro indicaron el objetivo con frases que empezaban con “Este estudio se ha diseñado para...” o bien “La finalidad del estudio ha sido...”. Sólo uno de los trabajos no especifica el objetivo. En 29 casos, los autores y los editores opinaron que el objetivo era evidente en el título y la introducción.

A veces, las hipótesis de trabajo se especifican aunque con poca uniformidad. Son raros los casos en los que se especifican, en artículos de investigación, las definiciones operacionales, supuestos, limitaciones e importancia del trabajo. Aparentemente, para el investigador “experto”, estos ítems se cumplimentan en el desarrollo del problema y pueden deducirse o sobreentenderse, o ambas cosas, sin que sea necesario especificarlos. Si el artículo está bien redactado, se pueden discernir las definiciones operacionales, los supuestos y limitaciones, además de las variables independiente, dependiente y categórica aunque no se hayan definido específicamente. Además, la importancia del estudio debe quedar implícita si el autor ha redactado una buena introducción.

## Algunas bases de datos disponibles para búsquedas informatizadas

**ERIC (Centro de Información de Recursos Docentes) (*Educational Resources Information Center*) (1966 — )**

**Guía de descriptores.** *Thesaurus de los descriptores de ERIC.*

**Tipo de información.** Se encuentran todos los temas relacionados con la educación desde preescolar hasta formación de adultos, clasificados en áreas de administración, currículos, enseñanza y aprendizaje. A los artículos publicados se accede desde el *Current Index to Journals in Education*, y el material no publicado puede obtenerse de los *Resources in Education que forma parte del ERIC*. Estas dos bases de datos pueden consultarse simultáneamente o por separado. La búsqueda informatizada del ERIC también está disponible en CD-ROM.

**Medline (1966 — )**

**Guía de descriptores.** *Medical Subjects Headings.*

**Tipo de información:** Puede accederse a través de MEDLINE a la amplia cobertura de artículos de revistas de biomedicina, cuidado de la salud, gerontología y otras áreas del *Index Medicus*. También está disponible en CD-ROM.

**PsycINFO (1967 — )**

**Guía de descriptores.** *Thesaurus of Psychological Index Terms.*

**Tipo de información.** Artículos de revistas, tesis y publicaciones técnicas (las citas de libros y de capítulos de libros están disponibles desde 1992) de psicología y ciencias del comportamiento relacionadas. Esta base de datos proviene del *Psychological Abstracts*. El *Psyclit* (1974 — ) es la versión en CD-ROM del PsycINFO.

**Sociological Abstracts (1963 — )**

**Guía de descriptores.** *Thesaurus of Sociological Indexing Terms.*

**Tipo de información.** Esta base de datos contiene una amplia cobertura de artículos de revistas, publicaciones de conferencias y libros de psicología social, del desarrollo y clínica. Varios archivos de bases de datos están combinados en esta cobertura. *SOCIO-FILE* (1974- ) es la versión en CD-ROM.

**Dissertation Abstracts Online (1861– )**

**Guía de descriptores.** Ninguna. Las palabras clave las establece el que realiza la búsqueda.

**Tipo de información.** Resúmenes de tesis y proyectos de másters (desde 1962) del *Dissertation Abstracts International* y del *American Doctoral Dissertation*. Incluye tesis de casi todas las instituciones de América, Canadá y otros países. Este servicio está disponible en CD-ROM.

**Uncover (1988 — )**

**Guía de descriptores.** Ninguna. Las palabras clave las establece el que realiza la búsqueda.

**Tipo de información:** Artículos e índices de más de 15.000 revistas con miles de citas añadidas diariamente. Es un servicio de pago que ofrece la oportunidad de recibir los artículos por fax en 24 horas.

**SPORTDiscus (1975 — )**

**Guía de descriptores.** Ninguna. Las palabras clave las establece el que realiza la búsqueda.

**Tipo de información.** Amplia cobertura de la literatura de temas sobre deporte, forma física y educación física. También incluye un índice de tesis y tesinas, además de citas sobre investigación de los aspectos sociales y psicológicos del deporte y del ocio. Este servicio está disponible en CD-ROM.

educación física. Se desarrollaron y refinaron las técnicas observacionales, lo que permitió una evaluación precisa del diseño de las acciones docentes y la interacción estudiante-profesor. El área de observación en el aula, llamada *investigación de la enseñanza* se desarrolló para su utilización en educación física en el gimnasio y en el terreno de juego. La investigación en teoría curricular en educación física, medida y evaluación, y la historia del deporte y de la educación física recibieron una atención renovada.

Muchos investigadores del ámbito de la actividad física empezaron a establecer grupos profesionales fuera de las organizaciones tradicionales y que habían estado en la AAHPERD. Algunos grupos, como el Colegio Americano de Medicina Deportiva, la Sociedad Norteamericana de Historia del Deporte, la Sociedad Americana de Psicología del Deporte y Actividad Física, además de otros grupos, contribuyeron a la investigación básica. Como respuesta a estos cambios entre sus miembros y en las actividades, la AAHPERD creó varias academias e hizo del Consorcio de la Investigación su brazo investigador, además de cambiar el nombre de la *Research Quarterly* en su 50º aniversario por el de *Research Quarterly for Exercise and Sport*. Para su primer número con el nuevo nombre (Safrit, 1980), el editor y el consejo de redacción solicitaron artículos de distintas áreas de estudio de la actividad física: biomecánica, fisiología del ejercicio, psicología del deporte, medidas y diseños de investigación, sociología del deporte, desarrollo motor y comportamiento motor.

Esta creciente actividad investigadora tanto en el estudio de la actividad física (ver Massengale y Swanson, 1966) como en la base profesional de la educación física sólo pudo dejar un mejor conocimiento y unos estándares académicos más altos. Mientras que la disciplina basada en el estudio de la actividad física siguió cada vez más dolorida y en crisis de identidad (p. ej., ¿cómo nos llamaremos a nosotros mismos: kinesiología, ciencia del ejercicio y del deporte, ciencias del movimiento?), eran manifiestos los programas de alta calidad de doctorado, investigación y pregrado. Por ejemplo, a finales de 1987, se publicó un número especial de *Quest* dedicado por entero a la docencia de posgrado. Aparecieron nuevas revistas especializadas sobre todo en subáreas de la disciplina (p. ej., *Journal of Exercise and Sport Psychology*, *Journal of Sport Sociology*, *Pediatric Exercise Science*, *Journal of Sport Management*).

Actualmente se están refinando los objetivos de programas de ejercicios y de habilidades en deportes basados en escuelas y en la comunidad, además de los de habilidades deportivas (desarrollo de habilidades de forma física y motoras) con un sólido conocimiento de la actividad física y una mayor sofisticación en la planificación y docencia de niños y adultos. Las nuevas revistas especializadas han desarrollado estas áreas, como la *Journal of Teaching in Physical Education*, para ubicar el mayor conocimiento generado. Todos estos factores deberán producir, durante los 20 o 30 años siguientes, una población más fácil de conocer, en mejor forma física y más habilidosa.

Pueden adelantarse hipótesis según las investigaciones previas y, a veces, algún modelo teórico. Además, las definiciones operacionales sirven para informar al lector del por qué el investigador utiliza ciertos términos (ver más adelante una definición operacional del críquet). Las definiciones operacionales deben describir fenómenos observables y deben relacionarse, en general, con las variables dependiente e independiente. También pueden establecerse supuestos básicos para especificar determinadas premisas que deben estar presentes para poder llevar a cabo el estudio. El investigador debe hacer referencia a las limitaciones y a posibles problemas del estudio que normalmente se deben a las restricciones impuestas por el propio investigador. El primer capítulo concluye con una referencia a la significación del estudio, el cual puede ser interpretado como un objeto de investigación básica o aplicada. El objetivo del capítulo es poner de manifiesto los resultados contradictorios y las limitaciones de investigaciones previas y la forma en que el estudio que se propone contribuye a un mejor conocimiento del tema de investigación.

### Definición operacional del críquet tal como nos lo explicaría un amigo australiano

Hay dos lados: uno al lado del campo y el otro fuera. Cada jugador del lateral sale y cuando está fuera vuelve a entrar, y entra el siguiente jugador hasta que sale.

Cuando están todos fuera, los del lado de fuera entran y los del lateral que están dentro salen e intentan echar a los que han entrado. A veces pueden quedar jugadores dentro y no salir.

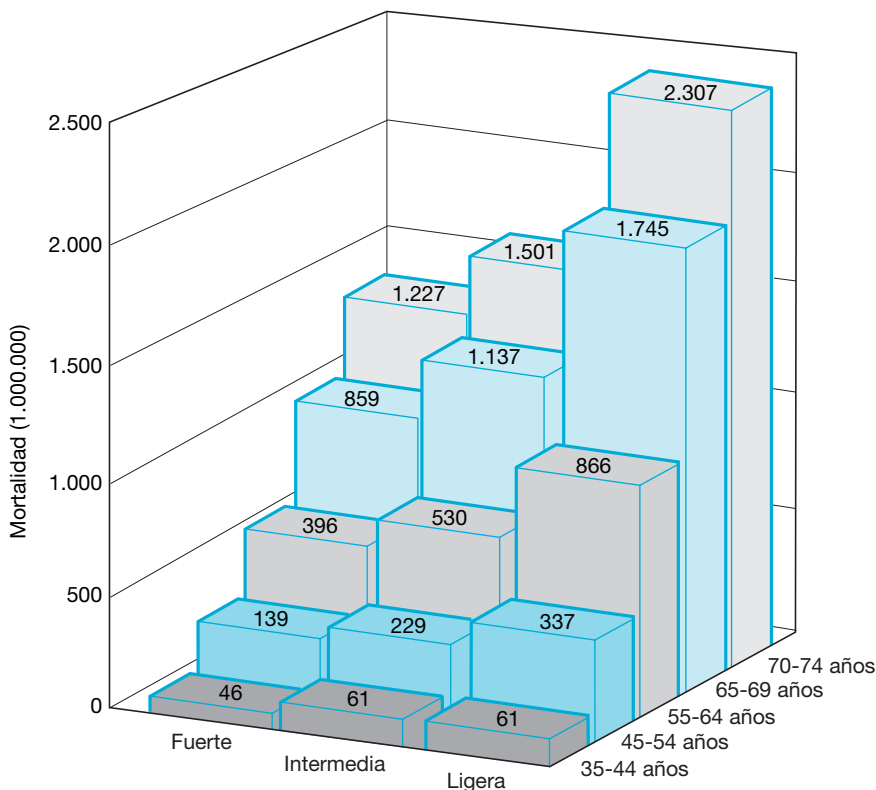
Cuando los de ambos laterales han estado dentro y fuera, incluyendo los que no han salido, ¡se acaba el partido!

El capítulo 2 de este libro hace referencia a la revisión bibliográfica e incluye una explicación de los procesos de razonamiento inductivo y deductivo utilizados en el desarrollo del problema y en la formulación de hipótesis. El capítulo 3 expone el resto de apartados requeridos normalmente en la sección o capítulo de introducción de un proyecto.

Se necesitan muchas horas para preparar el primer capítulo (introducción) del proyecto, sobre todo en la preparación de la sección bibliográfica y en la formulación del problema. Normalmente, el estudiante tiene una fuerte dependencia del tutor y de estudios completos con ejemplos de formato y descripción. Antes de seguir, hay que decir que hay que escribir el proyecto en futuro simple, que se seleccionarán algunos participantes y que se llevarán a cabo determinados procedimientos. En teoría, si el proyecto está planificado en detalle y está bien redactado, sólo será necesario cambiar el tiempo verbal futuro a pasado simple para obtener los dos primeros apartados o capítulos de la tesis o tesina. Sin embargo, y siendo realistas, es probable que se hagan muchas revisiones entre el proyecto y la versión definitiva. Insistimos en que la importancia del estudio y de su aportación al ámbito profesional son los puntos principales del primer apartado del proyecto y que esto constituye la base de su aprobación o desaprobación.

## Descripción del método

El apartado de la metodología del proyecto con frecuencia determina la mayoría de las preguntas de los miembros del tribunal en la reunión de proyectos. **En el apartado o capítulo de metodología, el estudiante debe describir claramente la recogida de datos para resolver el problema expuesto en el primer apartado.** Debe especificar quiénes serán los participantes y cómo se seleccionarán, el número de participantes planificado, algunas características concretas de importancia, cómo se protegerán los derechos y la privacidad de los participantes, y cómo se realizará el consentimiento



**Figura 16.4.** Mortalidad por cardiopatía en las categorías de actividad ocupacional fuerte, intermedia y ligera por edades entre varones: Inglaterra y Gales, 1930-1932.

Reimpresa con permiso de Morris *et al.*, 1953, "Coronary Heart disease and physical activity of work", *The Lancet* 2:111-120.

to, se ha ido disponiendo de un gran volumen de investigaciones con diseño de estudios metodológicamente superiores que no han sido capaces de demostrar la fuerte relación entre el consumo de grasas y el cáncer de mama que se había observado en los primeros estudios ecológicos (Willet, 1990, 311-340). Según lo expuesto, los estudios ecológicos deben considerarse sólo como una primera línea de investigación encaminada al desarrollo de hipótesis demostrables.

En resumen, los estudios descriptivos son útiles para desarrollar y comprobar de manera algo tosca las hipótesis iniciales sobre las relaciones entre la exposición y la enfermedad. Sin embargo, pueden obtenerse resultados más definitivos y de manera más científica con el diseño de estudios analíticos o experimentales.

## Diseños analíticos

Los diseños analíticos, estudios de caso-control y de cohortes sirven para probar hipótesis concretas según relaciones causales entre distintas exposiciones, y la mortalidad e incidencia de sucesos, utilizando sólo métodos observacionales.

Para comprender mejor los diseños analíticos, hay que considerar un modelo simplificado de la historia natural de la enfermedad (ver fig. 16.5). Debido a que la mayoría de estudios epidemiológicos que han evaluado la actividad física como la exposición de interés principal se realizaron para enfermedades crónicas, este modelo es para enfermedades crónicas. Estas enfermedades, como la cardiopatía, osteoporosis o cáncer, pueden tardar entre 10 y 40 años en desarrollar un problema suficiente para que pueda diagnosticarse clínicamente.

Normalmente, los epidemiólogos de la actividad física están interesados en ver cómo la exposición a la actividad física en fases tempranas de la enfermedad modifica la evolución de la historia natural de la enfermedad. Dicha evolución puede

con la evolución de una cardiopatía concreta. Segundo, los estudios ecológicos limitan mucho las posibilidades de control de los efectos de otros factores que pueden enmascarar la relación que se está estudiando.

Mientras que los análisis ecológicos pueden proporcionar estimaciones útiles y válidas de una relación entre exposición y enfermedad, tal como se ha expuesto en el ejemplo anterior, también es sabido que producen resultados espurios. Por ejemplo, los primeros estudios realizados en la década de los setenta estudiando la relación entre el consumo de grasa de la dieta y la mortalidad por cáncer de mama revelaron una fuerte relación lineal entre esta exposición y el cáncer de mama. Las mujeres de países con una mayor ingesta de grasas tuvieron mayores tasas de mortalidad por cáncer de mama. Sin embargo, y desde este momento,

## Muestra de carta de portada #1

### *Garantía de confidencialidad*

Apreciado \_\_\_\_\_:

Es necesaria su participación en una encuesta nacional como autoridad reconocida en educación física. Como estudiante de doctorado en educación física en la Universidad de Georgia, estoy dirigiendo este estudio para comparar y contrastar la capacidad de percepción entre hombre y mujer de los comportamientos de liderazgo de los hombres y mujeres jefes de departamento. Esta institución se ha seleccionado al azar para tomar parte en este proyecto de investigación. Su jefe de departamento nos ha comunicado su aceptación para participar en el estudio. Estoy preguntando a miembros de la facultad seleccionados al azar dentro de vuestra institución si quieren participar. Su nombre es uno de los seleccionados para consultarle si participará respondiendo a los cuestionarios adjuntos.

La participación requiere unos 15 minutos de su tiempo para contestar los dos cuestionarios que se utilizarán para completar un banco de datos. Los cuestionarios y las instrucciones de cómo cumplimentarlos se han incluido con esta carta con la esperanza de que acepte participar. En este estudio no hay juicios de valor explícitos ni implícitos. Los instrumentos utilizados describen la percepción del comportamiento de líder del administrador. El análisis de los datos se realizará con los valores medios del grupo. Después de la encuesta y el análisis estadístico de los datos, me gustaría remitirle un resumen de los resultados. Todos los datos se tratarán con confidencialidad y no se identificarán las instituciones o los individuos que participen en el estudio.

Esperamos que encontrará tiempo en su apretada agenda para participar en este estudio. Gracias por su tiempo y colaboración. Espero su pronta respuesta.

Un saludo:

Kay A. Johnson.

Clifford G. Lewis  
Catedrático

Copias: 5

Reimpresa con permiso de C.G. Lewis

## Seguimiento

No debe ser un gran problema, pero es improbable que se obtenga un 100% de respuestas del envío inicial. Casi siempre es necesaria una carta de seguimiento que puede hacerse de muchas maneras. Una probabilidad es esperar unos 10 días después del envío inicial y luego enviar una carta a cada sujeto de la muestra, recordando la importancia de su participación (y agradeciendo también a los que ya han enviado el cuestionario). Unos 10 días después de la carta, puede enviarse una segunda carta con otra copia del cuestionario y otro sobre con la dirección del envío y un sello a quienes no hayan respondido. Esto es caro, pero el seguimiento es eficaz. En algunos casos, la persona se ha olvidado de responder, y lo hace al recordárselo. Otros que no responden, que habían decidido no contestar al cuestionario, pueden verse influidos por los esfuerzos del investigador por reiterar la importancia del estudio y de la aportación del encuestado.

### Validez de criterio

Grado en el que las puntuaciones de una prueba se relacionan con algún estándar o criterio reconocido.

### Validez de constructo

Grado en el que una prueba mide un modelo ideal; se establece normalmente relacionando los resultados de la prueba con algún comportamiento.

### Validez aparente

Ver *validez lógica*.

### Equilibrio estático

Capacidad para mantenerse en una postura.

### Validez concurrente

Tipo de validez de criterio donde se correlaciona un instrumento de medida con algún criterio valorado de forma concurrente o simultánea.

apoyo monopodal tiene validez lógica. Un test de velocidad de movimiento, en el cual se cronometra a una persona que corre una distancia determinada, debe considerarse una prueba que posee validez lógica. A veces, la validez lógica se utiliza en estudios experimentales, aunque los investigadores prefieren tener pruebas más objetivas de la validez de la medida.

## Validez de contenido

La validez de contenido pertenece al ámbito del aprendizaje en pedagogía. Una prueba tiene validez de contenido si son adecuados los temas que se desarrollan durante el curso. Como en la validez lógica, no hay prueba estadística que confirme la validez de contenido. El diseñador de la prueba prepara una tabla de especificaciones (a veces llamada esquema del test) antes de realizar la prueba. Los temas y los objetivos del curso, así como la importancia relativa de cada uno, pueden codificarse con un número correspondiente de cuestiones de cada área.

## Validez de criterio

Las medidas utilizadas en investigación suelen validarse contrastándolas con algún modelo. Hay dos tipos principales de validez de criterio: validez concurrente y validez predictiva.

### *Validez concurrente*

La **validez concurrente** es un tipo de validez de criterio que implica la correlación de un instrumento con algún criterio que actúa al mismo tiempo (p. ej., concurrentemente). Muchas medidas de actividad física se validan de esta manera. Las medidas de criterio incluyen medidas validadas o aceptadas, como las valoraciones de los jueces y las técnicas de laboratorio. En general, se utiliza la validez concurrente cuando el investigador quiere sustituir un criterio difícil de medir por una prueba más corta y fácil de realizar.

Como ejemplo, el consumo máximo de oxígeno es la medida más válida de la condición cardiorrespiratoria. Sin embargo, se necesita un laboratorio, instrumentación cara y mucho tiempo de prueba; además sólo puede ser valorada una persona a la vez. Supongamos que un investigador, Douglas Bay, quiere clasificar a los individuos según su forma física antes de asignarles los tratamientos experimentales. En vez de utilizar una prueba sofisticada como la del consumo máximo de oxígeno, Douglas afirmó que podría ser ventajoso utilizar una prueba de subir escalones que había diseñado. Para determinar en qué medida era una prueba válida del estado cardiorrespiratorio, realizó la prueba de consumo máximo de oxígeno a la vez que la prueba de subir escalones a un grupo de participantes (de la misma población que utilizaría para su estudio) y contrastó los resultados de ambas pruebas. Si hay una relación satisfactoria, Doug puede concluir que su prueba de subir escalones es válida. Las pruebas escritas también pueden validarse de esta manera. Por ejemplo, un investigador puede preferir utilizar una prueba de ansiedad de diez ítems que puede realizarse en menos tiempo en lugar de una versión más larga. Las puntuaciones de jueces actúan como medidas criterio de algunas pruebas (a veces, la habilidad deportiva y los rasgos de comportamiento se validan de esta manera). Se necesita mucho tiempo y esfuerzo para reclutar a jueces competentes, enseñarles a utilizar la escala de puntuaciones, realizar las pruebas de concordancia entre jueces, establecer un número suficiente de intentos, etc. Por esto, las puntuaciones de jueces no se utilizan de forma habitual para valorar acciones. Es más económico usar determinadas pruebas de habilidad. Además, éstas proporcionan información sobre los resultados y medidas del progreso de los estudiantes. Las pruebas de habilidad deben validarse inicialmente, realizando las pruebas y con jueces que puntúen a los individuos en estas habilidades. Puede calcularse un coeficiente de validez mediante la correlación de las puntuaciones de las pruebas de habilidad con las valoraciones de los jueces.

informado. Se detallarán los métodos de medida y se documentarán la validez y fiabilidad de estas medidas. A continuación se describe la metodología. Si, por ejemplo, el estudio es una encuesta, el estudiante debe exponer los pasos a seguir y la carta de presentación, la forma de envío de los cuestionarios y el seguimiento. Si el estudio es experimental, se describirán explícitamente los tratamientos (o los programas experimentales) y los métodos de control que se realizarán. Para terminar, el estudiante debe explicar el diseño experimental y la planificación del análisis estadístico de los datos.

Hemos insistido en la importancia de realizar un estudio piloto antes de la recogida de los datos. Si se ha realizado un estudio piloto, debe describirse e informarse de los resultados.

A menudo, los miembros del tribunal muestran un gran interés sobre cuestiones como los tipos de tratamiento que producen cambios significativos, hasta qué punto son precisas las medidas y si pueden discriminar entre los participantes, y de qué manera satisfactoria puede tomar las medidas el investigador y administrar los tratamientos. El estudio piloto debe responder a estas cuestiones.

En el capítulo 4 recomendamos que el estudiante utilice la literatura como ayuda para determinar la metodología. Respuestas a cuestiones como si ciertas condiciones de tratamiento son suficientemente largas, intensas y frecuentes para provocar cambios pueden ser defendidas por los resultados de estudios previos.

## Proceso del proyecto

Insistimos en los contenidos de un proyecto: la introducción (incluida la revisión bibliográfica) y la metodología a utilizar.

El proyecto propuesto —en conjunción con la información de fondo pertinente, las hipótesis probables, las definiciones operacionales y las acotaciones— determina la viabilidad del estudio. Por esto, el primer apartado (o capítulo) es vital para estimular el interés sobre el problema y establecer la motivación y el significado del estudio. La decisión de aprobar o no el proyecto por parte del tribunal recae principalmente en la capacidad de persuasión del primer apartado.

De hecho, la decisión básica sobre la importancia del tema debe haberse hecho antes de la reunión del tribunal de proyectos. **El estudiante debe consultar al tutor y a la mayoría (si no a todos) de los miembros del tribunal para conseguir un consenso sobre el valor del estudio antes de la reunión del tribunal de proyectos.** Si no puedes convencer a la mayoría del tribunal de que tu estudio es viable, no lo envíes de manera formal a la reunión de proyectos. Puedes tener problemas si te devuelven el proyecto con una nota como la de la página 392.

## Qué hay que esperar de un tribunal de proyectos

Dedicaremos unos momentos a explicar la composición del tribunal del proyecto. La estructura y el número de miembros de los tribunales varía entre las distintas instituciones. La mayoría de los tribunales de una tesina suelen constar por lo menos de tres miembros, y la mayoría de los tribunales de tesis de por lo menos cinco. El director y el ayudante están incluidos en este número, aunque para los estudiantes de máster no es necesario que esté el ayudante. Otros miembros deben seleccionarse basándose en sus conocimientos sobre el tema o por su experiencia en otros aspectos de la investigación como el diseño y análisis estadístico. A veces, la institución o el departamento especifican cuántos miembros del tribunal deben ser de dentro y de fuera del departamento.



Mi tutor no tiene mucha fe en mi proyecto de tesis.

## Técnica de las mitades

Método de prueba de la fiabilidad en que la prueba se divide en dos, con los ítems de numeración impar en una mitad y los de numeración par en la otra, y se comparan las dos mitades.

## Fórmula de pronóstico de Spearman-Brown

Ecuación que valora la fiabilidad del total de la prueba cuando se utiliza la técnica de las mitades para probar la fiabilidad.

## Método de Flanagan

Un proceso de estimación de la fiabilidad en que la prueba se distribuye en dos mitades, y las varianzas de las mitades de la prueba se analizan en relación con la varianza total de la prueba.

## Método de la equivalencia racional de Kuder-Richardson (KR)

Fórmula que calcula la fiabilidad de una prueba efectuándola sólo una vez.

## Coefficiente alfa

Técnica para determinar la fiabilidad de las pruebas de intentos múltiples; también se denomina *coeficiente alfa de Cronbach*.

La **técnica de las mitades** se ha usado mucho en pruebas escritas y a veces en pruebas de acciones que necesitan muchos ensayos. La prueba se divide en dos y, entonces, se correlacionan ambas mitades. Una prueba puede dividirse en dos mitades, aunque no suele resultar satisfactorio. A veces, alguien se cansa cerca del final de la prueba, y a veces las preguntas más fáciles se colocan en la primera mitad. Normalmente, se comparan las preguntas con numeración impar con las de numeración par. Así, el número de preguntas bien realizadas con numeración impar se correlaciona con el número de respuestas correctas de las preguntas con numeración par.

Ya que la correlación se establece entre dos mitades de la prueba, el coeficiente de fiabilidad representa sólo la mitad del total de la prueba; esto significa que del comportamiento se ha muestreado sólo la mitad del total. Entonces, se utiliza un método en escalones, la **fórmula del pronóstico de Spearman-Brown**, para estimar la fiabilidad del total de la prueba ya que éste se basa en dos veces la muestra del comportamiento (el doble del número de ítems).

La fórmula es:

$$\text{Coeficiente de fiabilidad corregido} = \frac{2 \times \text{fiabilidad para } 1/2 \text{ prueba}}{1,0 + \text{fiabilidad para } 1/2 \text{ prueba}}$$

Si, por ejemplo, la correlación entre los ítems de numeración par y los ítems de numeración impar es de 0,85, el coeficiente de fiabilidad corregido será:

$$\frac{2 \times 0,85}{1,00 + 0,85} = \frac{1,70}{1,85} = 0,92$$

Otro método de las mitades es el **método de Flanagan**, que analiza la varianza de las mitades de la prueba en relación con la varianza del total. En este caso no se realiza ni método de correlación ni método de los escalones de Spearman-Brown.

En el **método de la equivalencia racional de Kuder-Richardson (KR)**, se utiliza una de las dos fórmulas conocidas como KR-20 y KR-21 para los ítems con valores dicotómicos (p. ej., bien o mal). Sólo se necesita hacer una vez la prueba, y se calcula la no correlación.

El coeficiente resultante representa un promedio de los coeficientes de todas las biparticiones posibles. Según numerosos expertos en pruebas, el KR-20 indica las proporciones de estudiantes que responden de manera correcta a cada ítem y de manera incorrecta en relación con la varianza total de los valores. El KR-21 es una versión simplificada y menos precisa del KR-20.

La técnica del **coeficiente alfa** a veces recibe el nombre de **coeficiente alfa de Cronbach** (ver Cronbach, 1951). Es una generalización en el cálculo de coeficientes más versátil que otros métodos. Una de las ventajas particulares del coeficiente alfa es que puede utilizarse con ítems que tengan escalas de valores distintas, como las pruebas de ensayo y las escalas de actitud que tienen como posibles respuestas “muy de acuerdo”, “de acuerdo”, etc.

El método consiste en el cálculo de las varianzas de las partes de una prueba. Las partes pueden ser ítems, mitades de pruebas, ensayos, o una serie de pruebas cortas y rápidas, como los concursos. Cuando los ítems son dicotómicos (p. ej., bien o mal), el coeficiente alfa tiene la misma fiabilidad estimada que la de KR-20 (en realidad, KR-20 es un caso particular del coeficiente alfa). Cuando las partes son mitades de una prueba, el resultado es el mismo que el del método de las mitades de Flanagan. Y cuando las partes son ensayos o pruebas, el resultado es el mismo que el de la correlación interclase. Es posible que el coeficiente alfa sea el método más utilizado de estimación de la fiabilidad en pruebas estandarizadas.

- El investigador tiene una relación de confianza y de colaboración con los participantes
- Los métodos de recogida de datos no son invasivos
- Los métodos de recogida de datos permiten la descripción de los sucesos
- Se dedica el tiempo suficiente al trabajo de campo

### Análisis de datos

- Análisis realizado durante y después de la recogida de datos
- Triangulación de las fuentes de datos y búsqueda de la convergencia
- Búsqueda de casos negativos
- Aporta interpretación y teoría, así como la descripción de los sucesos
- Da oportunidad a los participantes de corroborar la interpretación
- Preparación para una evaluación detallada por colegas de los métodos y su interpretación

### Preparación de la memoria

- Descripción completa del entorno
- Descripción completa de los métodos
- Incluye la descripción de los valores y supuestos del investigador
- Utiliza viñetas y valoraciones para apoyar las conclusiones y la interpretación

### Valoración general

- Validez interna: ¿qué confianza te inspira la calidad de la descripción y de la interpretación de los sucesos en el entorno específico de la investigación?
- Validez externa: ¿cuál es tu opinión sobre la posibilidad de aplicar los resultados de este estudio a distintos entornos que conozcas?

Reproducido por cortesía de Linda L. Bain.

## Resumen

Los métodos de investigación cualitativa incluyen trabajo de campo, estudio de casos, etnografía y memorias narrativas. El investigador recoge los datos en un entorno natural como el gimnasio, el aula, una sala de musculación o una instalación deportiva.

La investigación cualitativa no tiene las hipótesis preconcebidas, características de la investigación cuantitativa. Destaca el razonamiento inductivo, según el cual el investigador intenta desarrollar hipótesis a partir de las observaciones. El centro de atención se localiza en la “esencia” de los fenómenos. El investigador debe demostrar sensibilidad y capacidad de percepción en la recogida y análisis de los datos.

Insistimos en la importancia de tener acceso a los datos en su entorno natural. Es esencial conseguir la compenetración y la sinceridad de los participantes. Los métodos más comunes de recogida de datos son la entrevista y la observación. Los datos deben analizarse durante y después de su recogida. El investigador debe clasificar y ordenar los datos además de desarrollar posibles hipótesis que conduzcan a otras fuentes y tipos de datos.

El análisis de los datos implica organización, abstracción, integración y síntesis. La narrativa analítica es la base de la investigación cualitativa. La viñeta narrativa da al lector la sensación de estar presente durante la observación; confiere, de alguna manera, un significado holístico a la situación. No es infrecuente incluir en los estudios cualitativos análisis cuantitativos.

La triangulación de los datos se utiliza para establecer validez y fiabilidad. La triangulación de los datos incluye múltiples fuentes de datos, distintos investigadores y diferentes métodos. En la teoría de la triangulación, la situación se estudia desde distintos puntos de vista teóricos.

## Resumen

Este capítulo explica la información que suele presentarse en el primer apartado o capítulo de una tesis o tesina (excluida la revisión bibliográfica). En primer lugar, exponemos la extensión y contenido del título del trabajo. A veces se olvida la importancia de un título bueno, corto y descriptivo para su ordenación y búsqueda bibliográfica.

La introducción de un trabajo científico suele ser difícil de redactar. Requiere un gran aporte de inteligencia, esfuerzo y habilidad para convencer al lector de la importancia del estudio. Si no se realiza bien, el lector no estará motivado para leer el resto del trabajo.

La definición del problema y de las hipótesis de trabajo suelen aparecer en muchos estudios de investigación, que pueden ser tesis o tesinas, artículos de revistas o proyectos de investigación. Normalmente, las definiciones operacionales, supuestos, limitaciones y delimitaciones, así como el interés del trabajo, suelen especificarse sólo en tesis y tesinas. El objetivo es ayudar (o forzar) al investigador a definir y delimitar de forma clara el planteamiento de la investigación. Las definiciones operacionales describen específicamente cómo se utilizarán ciertos términos (sobre todo las variables dependientes) en un estudio concreto. Los supuestos definen las condiciones básicas cuya existencia debe asumirse para que los resultados tengan credibilidad. Las delimitaciones hacen referencia al enfoque del estudio impuesto por el investigador como el número y las características de los participantes, las condiciones de tratamiento y las variables dependientes específicas que se han utilizado, además de cómo se han medido. Las limitaciones son posibles influencias en los resultados como consecuencia de las delimitaciones o de la imposibilidad para controlarlas completamente.

El apartado sobre el interés del trabajo obliga al investigador a plantear la cuestión inevitable sobre la motivación del estudio. En términos de flujo contextual debe continuarse con la introducción. Suele poner de manifiesto la relación (y diferencias) entre el estudio actual y los anteriores, controversias y vacíos de la literatura, y la contribución que este estudio puede hacer al profesional, a los modelos teóricos existentes o a ambos.

### Evaluación de lo aprendido

1. Para cada una de las breves descripciones de estudios, redacta un título, el objetivo u objetivos, y tres hipótesis de investigación.

Los investigadores valorarán lo siguiente:

- a. La adquisición de habilidades de tres grupos de niños y niñas de cuarto grado que han aprendido mediante distintas estrategias docentes (A, B y C).
- b. Autoconcepto de dos grupos de niños (un grupo de poca fuerza y un grupo de mucha fuerza) antes y después de un programa de entrenamiento de fuerza.
- c. Composición corporal (estimada como porcentaje de adiposidad) determinada con métodos de análisis de impedancia eléctrica en participantes con hidratación normal (normohidratados) y posteriormente, cuando se hayan deshidratado.
- d. Medias de las clasificaciones de deportistas varones y mujeres de deportes principales y minoritarios (club) procedentes de universidades y pequeñas escuelas privadas

2. Localiza cinco artículos de revistas científicas, y para “cada uno” determina: (a) la hipótesis, si no se ha establecido; (b) la(s) variable(s) independiente(s); (c) la(s) variable(s) dependiente(s); (d) una definición operacional para la variable dependiente; (e) un mínimo de dos delimitaciones; (f) un mínimo de dos limitaciones, y (g) un supuesto básico.

Jerga	Uso recomendado	Jerga	Uso recomendado
En términos de	Sobre	Primero de todo	Primero
En un número de casos	Algunos	Primero que	Antes
En un sentido muy real	En este sentido	Prueba definitivamente	Prueba
Era de la opinión de que	Creía	Puede ser que	Creo
Es algo dudoso que	Posiblemente	Queda ampliamente señalado	
Es de interés observar que	(eliminarlo)	en este contexto que	Creo
Es evidente que	Evidentemente	Queremos agradecer	Agradecemos
Es evidente que a genera b	a genera b	Realizar	Hacer
Escaso en número	Poco	Referenciado como	Llamado
Está claro que	Claramente	Relativo a	Sobre
Está claro que se necesitarán		Resultado final	Resultado
muchos trabajos adicionales antes		Se cree que	Creo
del conocimiento completo	No lo entiendo	Se debe al hecho de que	Debido a
Este resultado parece indicar	Este resultado indica	Se define como	Es
Explicado por el hecho	Porque	Se fundamenta en que	Desde, porque
Falta de habilidad para	No se puede	Se ha observado en el curso	
Finalización	Final	de los experimentos que	Hemos observado
Finalmente	Último	Se ha propuesto que	Creo
Ha llamado nuestra atención	Descubrimos	Sin embargo, puede observarse que	Pero
el hecho de que	tardíamente	Sin necesidad de decir	(eliminarlo, y procurar eliminar cualquier frase que le siga)
Ha sido afirmado por Smith	Smith afirmó		
Habiendo contemplado a	Sobre	Son de la misma opinión	Están de acuerdo
Hay motivos para creer	Creo	Subsecuente a	Después
Iniciar	Empezar	Suficiente	Bastante
La cuestión es como si	Si	Tenemos un conocimiento insuficiente	No sabemos
La gran mayoría de	La mayoría	Terminación	Final
La opinión relevante es que	Creo	Tiene la capacidad de	Puede
Las determinaciones de proteínas	Las proteínas fueron	Tomar en consideración	Considerar
fueron realizadas	determinadas	Totalmente único	Único
Llevar a cabo	Hacer	Un número de	Muchos
Luchar contra	Prohibir	Una cantidad considerable de	Mucho
Más que interesante	Interesante	Una magnitud de orden	10 veces más
Más rápido	Rápido	Una mayoría de	La mayoría
No tanto como	Como	Unidos juntos	Unidos
Para el objetivo de	Para	Utilizar	Usar
Por el motivo que	Desde, porque	Vías y sentidos	Vías, sentidos (una de las dos)
Por el sentido de	Por, con		

Reimpresión de *How to Write and Publish a Scientific Paper*, 3.ª edición, de Robert A. Day. Utilizada con el permiso de Oryx Press, 4041 N. Central Ave. Suite 700, Phoenix, AZ, 85012. (800) 279-6799.

información dentro de los subapartados es desarrollando un guión. Si el guión es muy preciso, la redacción será más fácil. Es útil seleccionar un artículo de revisión de una revista o una revisión bibliográfica de una tesis o tesina y reconstruir el guión que ha utilizado el autor. Buscando en tesis y tesinas antiguas, encontramos que la revisión bibliográfica suele ser una revisión histórica normalmente presentada en orden cronológico. Recomendamos no seleccionar estos estudios antiguos ya que el estilo suele ser distinto y poco sintético.

Para redactar una buena revisión bibliográfica, debes redactarla como si estuvieras leyendo. Nadie quiere leer resúmenes de estudios después de ser presentados en orden cronológico. Un enfoque más interesante y más fácil de leer es la presentación de un

inhibirse mediante exposiciones preventivas (un estilo de vida activo) o activarse por exposiciones adversas (un estilo de vida sedentario).

El *período de inducción* de una enfermedad se define como el período entre el momento en que la exposición de interés altera la fisiología del proceso de la enfermedad y el momento en el que la enfermedad se diagnostica clínicamente. El diagnóstico clínico siempre es más tardío que las primeras fases de la enfermedad. Cuando se diagnostica clínicamente una enfermedad, el paciente deviene un caso incidente.

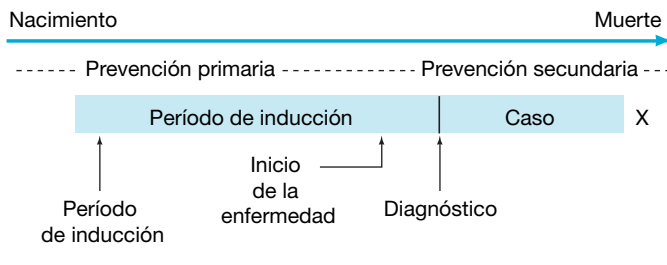


Figura 16.5. Historia natural de la enfermedad crónica.

La prevención primaria (1<sup>aria</sup>) hace referencia a la prevención de la aparición del primer síntoma de la enfermedad, mientras que la prevención secundaria (2<sup>aria</sup>) hace referencia a la prevención de la recidiva. El objetivo habitual de los estudios analíticos es cuantificar el efecto de las exposiciones que tienen lugar durante el período de inducción de un posible riesgo de enfermedad. Esto es, el objetivo consiste en identificar predictores del desarrollo de la enfermedad durante las fases tempranas de la historia natural de la enfermedad.

### Estudios de cohorte

Los términos *seguimiento*, *prospectivo* y *estudios longitudinales* se han utilizado para describir el diseño del estudio de cohortes. En este diseño se define una población amplia libre de enfermedad, de la que se obtiene una valoración de las exposiciones relevantes. Los datos de la línea base se utilizan para clasificar la cohorte en distintos niveles de exposición (p. ej., baja, media y alta). Después de establecer la línea base, empieza el período de seguimiento. Debido a la relativa infrecuencia de las enfermedades crónicas como el cáncer de colon o de mama, el período de seguimiento puede durar desde un mínimo de 2 años hasta más de 20 años. Al concluir el período de seguimiento, se tabulan los individuos de la cohorte que han sido diagnosticados de la enfermedad y que han muerto durante el seguimiento.

Los estudios de análisis de cohortes son relativamente sencillos. Como estamos interesados en si la exposición a distintos niveles de la línea base es predictiva de la incidencia de la enfermedad, el análisis básico consiste en el simple cálculo de las tasas de enfermedad para los distintos niveles de exposición. Por ejemplo, los tasas de mortalidad entre los individuos que practican ejercicio con regularidad pueden compararse con las tasas de mortalidad de los que no practican ejercicio. En análisis más refinados, pueden calcularse las tasas de mortalidad de tres a cinco niveles de actividad física y se pueden establecer comparaciones entre el nivel de baja actividad con cada uno de los niveles más altos. En los estudios de cohortes, las tasas de enfermedad a menudo se expresan en relación a persona-año de seguimiento. Una persona-año significa un año de observación de una persona durante el período de seguimiento. Por ejemplo, en la figura 16.6, se ha observado una cohorte de cinco individuos desde el momento 0 hasta el momento 5 (seis años completos). La contribución persona-año de un individuo a la cohorte empieza cuando se incluye al individuo en el seguimiento. En la figura, los tres primeros individuos se incluyen en el momento 0, el cuarto individuo se incluye en el momento 1 (un año después), mientras que el quinto individuo se incluye en el momento 2 (dos años después). El seguimiento termina cuando ocurre un suceso (x), cuando se pier-

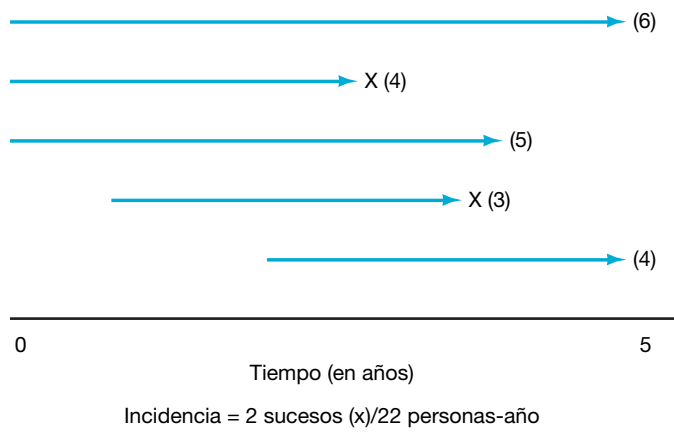


Figura 16.6. Ejemplo de persona-año.

observador no conoce el resultado esperado. La técnica experimental de doble ciego descrita en el capítulo 17 es útil para el control de los errores de expectativa. En el método de doble ciego, los observadores no saben los tratamientos que recibe cada individuo. Tampoco deben saber qué acciones corresponden al pretratamiento y cuáles al postratamiento.

En resumen, los errores de clasificación siempre pueden aparecer. El investigador debe reconocerlos y procurar eliminarlos. **Una forma de minimización de los errores de clasificación consiste en definir la acción a valorar de la manera más objetiva posible. En otras palabras, evitar que el observador realice muchas valoraciones de juicio.** Otra propuesta es que el investigador desconozca la hipótesis y qué sujetos reciben un tratamiento determinado. Pueden reducirse los sesgos y las expectativas si el observador no tiene información alguna sobre los logros, inteligencia, posición social y otras características del participante. La precaución más importante que debe tomar el investigador es entrenar a los observadores de manera adecuada para que alcancen niveles altos de precisión y fiabilidad interobservadores.

## Medida del conocimiento

Es cierto que la medida del conocimiento es una parte fundamental de los aspectos educativos de la educación física, ciencias del ejercicio y ciencias del deporte. Sin embargo, la construcción de pruebas de conocimiento también es importante con fines de investigación. Los métodos para establecer la validez y la fiabilidad de la mayoría de los instrumentos de medida de papel y lápiz utilizados en investigación son similares a los métodos descritos anteriormente. Sin embargo, también hay que determinar en qué medida la capacidad y la habilidad de cada uno de los ítems actúan como discriminantes entre los niveles de habilidad de la manera requerida; esto es, hay que realizar un análisis de ítems. (Para una exposición en profundidad de la metodología de las medidas en las pruebas de conocimiento, ver Mood, 1989.)

## Análisis de los ítems de la prueba

El objetivo del **análisis de ítems** es determinar qué ítems de la prueba son válidos y cuáles necesitan reescribirse o eliminarse. Los dos aspectos básicos del análisis de los ítems es la determinación de la dificultad de los ítems de la prueba y determinar su potencia en la discriminación entre diferentes niveles de logros.

### **Dificultad del ítem**

El análisis de la **dificultad del ítem** suele calcularse fácilmente. Sólo hay que dividir el número de personas que responden correctamente al ítem entre el número total de personas que han respondido al ítem. Por ejemplo, si 80 personas han respondido al ítem y 60 lo han hecho correctamente, el ítem tiene un índice de dificultad de 0,75 (60/80). El ítem de mayor dificultad es que tiene el índice de dificultad más bajo. Por ejemplo, si sólo contestan correctamente al ítem 8 de 80, el índice de dificultad es de 8/80, o 0,10. La mayoría de expertos en estas pruebas recomiendan eliminar las preguntas con índices de dificultad inferiores a 0,10 o superiores a 0,90. Son preferibles las preguntas que tienen índices de dificultad de alrededor del 0,50. A veces, los diseñadores de las pruebas pueden querer establecer un índice de dificultad concreto con fines de búsqueda. Por ejemplo, si sólo se quiere seleccionar el 30% del grupo que realiza la prueba, esto puede conseguirse utilizando preguntas con índices de dificultad de 0,30. Las preguntas que todo el mundo contesta correctamente o las que no contesta nadie no aportan información sobre diferencias individuales en las escalas de medida patrón.

### **Discriminación del ítem**

La **discriminación del ítem**, o el grado al cual los ítems de la prueba discriminan entre las personas que realizan bien la prueba y las que la realizan mal, es una consi-

### **Error de expectativa del observador**

Tendencia de un observador a ver pruebas de ciertos comportamientos esperados y a interpretar las observaciones en el sentido esperado.

### **Análisis de ítems**

Proceso de análisis de las pruebas de conocimiento en que se evalúa la adecuación de la dificultad y discriminación de los ítems de la prueba.

### **Dificultad del ítem**

Análisis de la dificultad de cada ítem de las pruebas de conocimientos, calculada con el cociente del número de personas que han respondido correctamente al ítem dividido por el total de personas que han respondido al ítem.

### **Discriminación del ítem**

El grado en que el ítem de una prueba discrimina entre quienes hacen bien la prueba y quienes la hacen peor; también recibe el nombre de *índice de discriminación*.

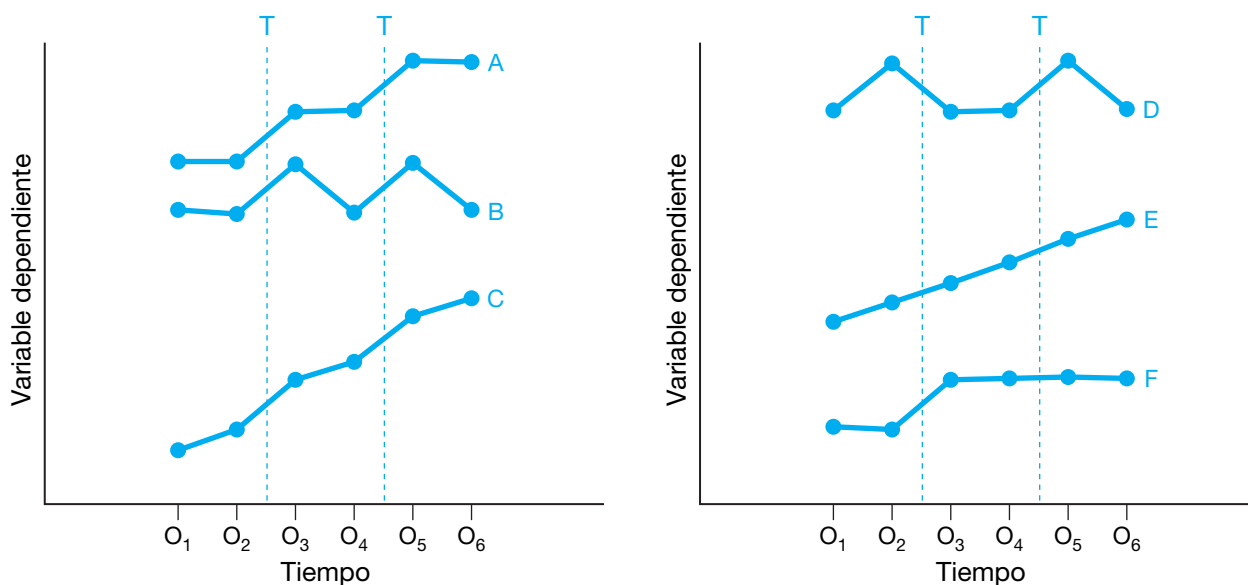
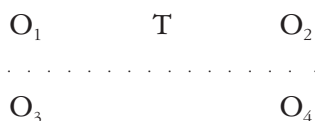


Figura 17.2. Ejemplos de cambios temporales en diseños inversos.



Esto se puede identificar como un diseño de pretratamiento y postratamiento sin aleatorización. Con frecuencia, los investigadores comparan O<sub>1</sub> con O<sub>3</sub> y establecen así la equivalencia de los grupos, si esta comparación no es significativa. Por desgracia, el no encontrar diferencias en los grupos de pretratamiento no significa que no sean diferentes en determinadas características no ponderadas que pueden influir en la evolución del trabajo. Si los grupos difieren cuando se comparan O<sub>1</sub> y O<sub>3</sub>, entonces se puede utilizar el ANCOVA para ajustar O<sub>2</sub> y O<sub>4</sub> en cuanto a las diferencias iniciales. Aunque este diseño se utiliza con mucha frecuencia, pensamos que los diseños de series temporales y los diseños inversos son diseños cuasi experimentales más potentes para la investigación de grupos completos.

### Diseño de ex post facto

En su forma más simple, el diseño *ex post facto* consiste en la comparación de un grupo estático, pero el tratamiento no está controlado por el investigador. Por ejemplo, comparamos con frecuencia características entre deportistas y no deportistas, en buenas condiciones y en malas condiciones físicas, mujeres y hombres, y expertos y noveles. En realidad, estamos buscando variables que discriminen entre estos dos grupos. Normalmente nos interesamos por la siguiente pregunta, ¿estas variables ejercen su influencia en el sentido de que sean distintos estos grupos? En verdad, este diseño no puede responder a esta pregunta, pero puede proporcionar aclaraciones y características para su manipulación en otros diseños experimentales. Este diseño también se ha llamado diseño de comparación causal.

Los diseños cuasi experimentales antes mencionados se han referenciado con frecuencia en estudios de investigación en actividad física. Sin embargo, hay varios diseños adicionales con una potencia mayor, pero se ha observado un menor uso en nuestra investigación. Esperamos que las siguientes presentaciones de dos diseños prometedores aumente el interés y el uso de estos diseños.

# Investigación histórica en la actividad física

*Nancy L. Struna*

Universidad de Maryland

*Es una pena que Noah y su grupo  
no pierdan la barca. Mark Twain*



La historia es la observación sistemática y la explicación del cambio, o de la ausencia de éste, en las situaciones humanas. En nuestro ámbito, este concepto de situación humana se aplica a cualquier aspecto relacionado con el movimiento y el cuerpo humano. Por este motivo, cuando los estudiosos hacen referencia a la “historia del deporte”, en realidad lo que significa es un subdominio que incluye observaciones sistemáticas y demostraciones de un conjunto de prácticas que incluyen deporte, salud, el cuerpo, medicina deportiva, aficiones, ocio y otras. Más allá de las conexiones físicas entre el movimiento y el cuerpo humano, lo que une todos estos temas y es el centro de atención académica son las modificaciones o la persistencia del comportamiento y actitudes a lo largo del tiempo.

En la búsqueda de la prueba y la significación del cambio o la persistencia del comportamiento y actitudes humanas frente al deporte, salud, ocio y otros aspectos, los historiadores actúan en terrenos similares a los de los científicos. En primer lugar, deben tener acceso a los laboratorios, que normalmente reciben el nombre de bibliotecas y archivos. Segundo, en estos laboratorios, trabajan para identificar y comprender los patrones en las pruebas o datos. Tercero, como los buenos científicos, los buenos historiadores pretenden elaborar generalizaciones significativas de la prueba histórica o dato. Las generalizaciones son “extensas”, extractos de síntesis ofrecidos por el sentido histórico a partir de múltiples fragmentos de pruebas o datos. Se expone una interpretación de los datos, normalmente a partir de un conjunto de hechos de un momento o a través del tiempo. Numerosas generalizaciones históricas están enmarcadas o se apoyan en teorías de las ciencias sociales, y los historiadores prueban estas teorías con datos de otros períodos u otros conjuntos de hechos. En realidad, pocos historiadores actúan a partir de una prueba del pasado sin tener en cuenta la teoría y sin valorar las generalizaciones teóricas. Si un trabajo sobre el pasado no obtiene ningún tipo de generalización, no es historia (Mandelbaum, 1977).<sup>1</sup>

## Paradigmas de la investigación

La buena historia y la buena ciencia comparten otra característica: un paradigma significativo o soporte. **Un paradigma es un instrumento teórico que incluye los objetivos académicos y los supuestos acerca del mundo, el pasado y la evidencia; una visión**

- Hunter, J.E., & Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Husserl, E. (1962). *Ideas: General introduction to pure phenomenology* (W.R. Boyce, Trans.). New York: Collier Books.
- Hyde, J.S. (1981). How large are cognitive gender differences? A meta-analysis using  $\omega^2$  and  $d$ . *American Psychologist*, 36(8), 892-901.
- Hyland, D. (1990). *Philosophy of Sport*. New York: Paragon House.
- Jacks, P., Chubin, D.E., Porter, A.L., & Connally, T. (1983). The ABCs of ABDs: An interview study of incomplete doctorates. *Improving College and University Teaching*, 31, 74-81.
- Jackson, A.W. (1978). The twelve-minute swim as a test for aerobic endurance in swimming. Unpublished doctoral dissertation, University of Houston.
- Jacob, E. (1987). Qualitative research traditions: A review. *Review of Educational Research*, 57(1), 1-50.
- Jacob, E. (1988). Clarifying qualitative research: A focus on tradition. *Educational Researcher*, 17, 16-19, 22-24.
- Jacobs, D., Anderson, J., & Blackburn, H. (1979). Diet and serum cholesterol: Do zero correlations negate the relationship? *American Journal of Epidemiology*, 110, 77-87.
- Jacobs, D.R., Jr., Luepker, R.V., Mittelmark, M.B., Folsom, A.R., Pirie, P.L., Mascioli, S.R., Hannan, P.J., Pechacek, T.F., Bracht, N.F., Carlaw, R.W., et al. (1986). Community-wide prevention strategies: Evaluation design of the Minnesota Heart Health Program. *Journal of Chronic Disease*, 39, 775-778.
- Jarusch, K.H., & Hardy, K.A. (1991). *Quantitative methods for historians: A guide to research, data, and statistics*. Chapel Hill: University of North Carolina Press.
- Johnson, B.L., & Nelson, J.K. (1986). *Practical measurements for evaluation in physical education* (4th ed.). Minneapolis: Burgess.
- Johnson, R.L. (1979). The effects of various levels of fatigue on the speed and accuracy of visual recognition. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Jones, E.R. (1988, winter). Philosophical tension in a scientific discipline: So what else is new? *NASPSPA Newsletter*, 14(1), 10-16.
- Joynt, C.B., & Rescher, N. (1961). The problem of uniqueness in history. *History and Theory*, 2, 150-162.
- Kavale, K., & Mattson, P.D. (1983). One jumped off the balance beam: Meta-analysis of perceptual-motor training. *Journal of Learning Disabilities*, 16, 165-173.
- Kendall, M.G. (1959). Hiawatha designs an experiment. *American Statistician*, 13, 23-24.
- Kennedy, J.J. (1983). *Analyzing qualitative data: Introductory loglinear analysis for behavioral research*. New York: Praeger.
- Kennedy, M.M. (1979). Generalizing from single case studies. *Evaluation Quarterly*, 3, 661-679.
- King, H.A., & Bandy, S.J. (1987). Doctoral programs in physical education: A census with particular reference to the status of specializations. *Quest*, 39(2), 153-162.
- Kirk, D. (1992). *Defining physical education: The social construction of a school subject in postwar Britain*. London: Falmer Press.
- Kirk, J., & Miller, M.L. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: Sage.
- Kirk, R.E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Belmont, CA: Brooks/Cole.
- Kleinman, S. (1968, May). Toward a non-theory of sport. *Quest*, 10, 29-34.
- Kohl, H.W., III, Dunn, A.L, Marcus, B.H., & Blair, S.N. (1998). A randomized trial of physical activity interventions: Design and baseline data from Project Active. *Medicine and Science in Sports and Exercise*, 30, 275-283.

demasiado lejos o bien se quedaron por el camino) encontramos filósofos que utilizan técnicas de crítica, poéticas y especulativas.

Los filósofos que han utilizado *razonamiento crítico y poético* parten de alguno de los puntos siguientes. En el apartado IIIA, Dixon supone que un valor importante de la competición deportiva es el “conocimiento preciso y total” de la mayor capacidad. Pero, éste es sólo un valor entre muchos otros que pueden adoptar los atletas de distintas culturas, en momentos distintos y con distintos objetivos personales. ¿Qué permite a Dixon hacer de esto un valor absoluto o incluso darle algo de importancia?

¿No asume Dixon la validez de la cultura occidental orientada al éxito y a la excelencia? ¿No será necesario ampliar el ámbito del deporte para encontrar ejemplos de otros valores como la armonía y la tranquilidad (encontrados en determinadas artes marciales del oriente) o de cooperación (en la cultura hindú)? ¿Qué ocurre al competir por encima de la clasificación en estos contextos tan distintos?

Los *filósofos especulativos* pueden imaginarse las ramificaciones de la tesis de Dixon en distintos contextos —algo que no ha respondido—. Por ejemplo, pueden especular que competir por encima de la clasificación puede ser más problemático en los deportes de contacto en que este desequilibrio suele acompañarse de agotamiento físico y dolor. También pueden especular que competir por encima de la clasificación es menos humillante en actividades paralelas (llamadas deportes cerrados) como los bolos, en los que el encuentro consiste en alternar la realización de las pruebas. Puede ser más humillante en las actividades interactivas (llamadas deportes abiertos) como el baloncesto, en las que parte de las pruebas están proporcionadas por el contrincante y en las que la prueba (si es demasiado difícil) puede mostrar a una persona como histriónica, desequilibrada o incompetente.

Estos filósofos también pueden especular que Dixon ha establecido una definición demasiado restrictiva de lo que es una persona. Él afirma que la humillación más profunda proviene de fallos del carácter y no de fallos en la acción deportiva.

¿Podría no ser una visión holística el que las personas tienden a dividir al ser humano en una parte espiritual y otra física, y entonces afirmar que los fallos espirituales son los determinantes, mientras que los físicos no lo son?

## Bibliografía

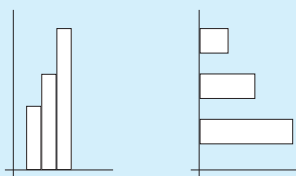
- Dixon, N. (1992). On sportsmanship and “running up the score”. *Journal of the Philosophy of Sport*, 19, 1-13.
- Dixon, N. (1998). Why losing by a wide margin is not itself a disgrace: response to Hardman, Fox, MacLanghlin, and Zimmerman. *Journal of the Philosophy of Sport*, 25, 61-70.
- Dixon, N. (2000). The inevitability of disappointment: Reply to Feezell. *Journal of the Philosophy of Sport*, 27.
- Feezell, R.M. (1999). Sportsmanship and blowouts: Baseball and beyond. *Journal of the Philosophy of Sport*, 26, 68-78.
- Hardman, A., Fox, L., MacLanghlin, D. y Zimmerman, K. (1996). On sportsmanship and running up the score: Issues of incompetence and humiliation. *Journal of the Philosophy of Sport*, 23, 58-69.

**Tabla A.6. Valores críticos de F**

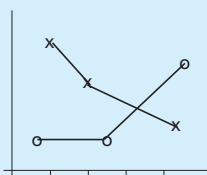
$n_1$		$n_2$ , grados de libertad (del numerador de F)																								
		1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	$\infty$	
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	253	254	254	254	254
	4,052	4,999	5,403	5,625	5,764	5,859	5,928	5,981	6,022	6,056	6,082	6,106	6,142	6,169	6,208	6,234	6,258	6,286	6,302	6,323	6,334	6,352	6,361	6,366	6,366	
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,48	19,48	19,49	19,49	19,50	19,50	
	98,49	99,00	99,17	99,25	99,30	99,33	99,34	99,36	99,38	99,40	99,41	99,42	99,43	99,44	99,45	99,46	99,47	99,48	99,48	99,49	99,49	99,49	99,49	99,50	99,50	
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74	8,71	8,69	8,66	8,64	8,62	8,60	8,58	8,57	8,56	8,56	8,54	8,54	8,53	
	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	26,92	26,83	26,69	26,60	26,50	26,41	26,35	26,27	26,23	26,18	26,14	26,12	26,12	
4	7,71	6,94	6,49	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,68	5,66	5,65	5,65	5,64	5,63	
	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37	14,24	14,15	14,02	13,93	13,83	13,74	13,69	13,61	13,57	13,52	13,48	13,46	13,46	
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,42	4,40	4,38	4,37	4,36	4,36	
	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89	9,77	9,68	9,55	9,47	9,38	9,29	9,24	9,17	9,13	9,07	9,04	9,02	9,02	
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67	3,67	
	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,60	7,52	7,39	7,31	7,23	7,14	7,09	7,02	6,99	6,94	6,90	6,88	6,88	
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23	3,23	
	2,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47	6,35	6,27	6,15	6,07	5,98	5,90	5,85	5,78	5,75	5,70	5,67	5,65	5,65	
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28	3,23	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93	2,93	
	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67	5,56	5,48	5,36	5,28	5,20	5,11	5,06	5,00	4,96	4,91	4,88	4,86	4,86	
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07	3,02	2,98	2,93	2,90	2,86	2,82	2,80	2,77	2,76	2,73	2,72	2,71	2,71	
	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11	5,00	4,92	4,80	4,73	4,64	4,56	4,51	4,45	4,41	4,36	4,33	4,31	4,31	
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54	2,54	
	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71	4,60	4,52	4,41	4,33	4,25	4,17	4,12	4,05	4,01	3,96	3,93	3,91	3,91	
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40	2,40	
	9,65	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40	4,29	4,21	4,10	4,02	3,94	3,86	3,80	3,74	3,70	3,66	3,62	3,60	3,60	
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30	2,30	
	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16	4,05	3,98	3,86	3,78	3,70	3,61	3,56	3,49	3,46	3,41	3,38	3,36	3,36	
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,28	2,26	2,24	2,22	2,21	2,21	
	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	3,85	3,78	3,67	3,59	3,51	3,42	3,37	3,30	3,27	3,21	3,18	3,16	3,16	
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13	2,13	
	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	3,70	3,62	3,51	3,43	3,34	3,26	3,21	3,14	3,11	3,06	3,02	3,00	3,00	
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,15	2,12	2,10	2,08	2,07	2,07	
	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,56	3,48	3,36	3,29	3,20	3,12	3,07	3,00	2,97	2,92	2,89	2,87	2,87	
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,10	2,09	2,07	2,04	2,02	2,01	
	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55	3,45	3,37	3,25	3,18	3,10	3,01	2,96	2,89	2,86	2,80	2,77	2,75	2,75	
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,04	2,02	2,00	1,97	1,96	1,96	
	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45	3,35	3,27	3,16	3,08	3,00	2,92	2,86	2,79	2,76	2,70	2,67	2,65	2,65	
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,29	2,25	2,19	2,15	2,11	2,07	2,04	2,00	1,98	1,95	1,93	1,92	1,92	
	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,44	3,37	3,27	3,19	3,07	3,00	2,91	2,83	2,78	2,71	2,68	2,62	2,59	2,57	2,57	

Grados de libertad del denominador de F

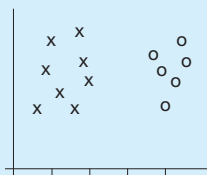
Tabla 19.4. Gráficos y diagramas



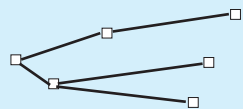
**Diagrama de barras y columnas.** Las barras (horizontal) son mejores para comparación de cantidades; se ordenan por tamaños, de menor a mayor o de mayor a menor. Las columnas (vertical) son idóneas para comparar cantidades en relación con el tiempo sobre todo si hay una tendencia evidente. El sombreado puede ayudar a distinguir entre barras y columnas.



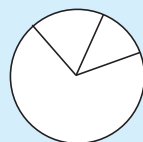
**Gráfico de curva.** El mejor para mostrar cambios en el tiempo; el tiempo está en la horizontal y la cantidad en la vertical. Permite comparar más de una curva. A veces, puede sombreadarse el área entre curvas para mostrar la cantidad de cambio; para diferenciar las líneas pueden utilizarse sombreados, líneas discontinuas, símbolos o líneas coloreadas.



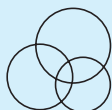
**Gráfico de puntos.** Muestra patrones de valores individuales; cada punto indica un valor tanto en el eje vertical como en el horizontal. Para diferenciar los grupos se pueden utilizar puntos distintos o diferentes símbolos de puntos.



**Diagrama de flujo.** Muestra las relaciones en un proceso; es útil para mostrar los pasos de un proceso cuando hay más de una opción (p. ej., si sí, entonces esto; si no, entonces esto).



**Diagrama de sectores.** El círculo representa el 100% del sector; un máximo de cinco o seis segmentos. Se ordenan los segmentos de mayor a menor empezando a las 12 horas; se pueden destacar los segmentos con sombreados o poniendo el segmento menor el más oscuro.



**Esquemas.** Relaciones entre variables o conceptos (p. ej., el solapamiento entre dos variables relacionadas).

Desarrollado por White, J.V. (1984). *Using charts and graphs: 1.000 ideas for visual persuasion*. New Providence, NJ: RR Bowker Co.

mente uno de otro. En la retención, el grupo experimental 2 redujo el número de errores mientras que el grupo experimental 1 y el control permanecieron en el mismo nivel". Cuando los resultados de grupos diferentes siguen un patrón similar, la figura suele tener un aspecto abigarrado. Si las figuras son abigarradas debido al solapamiento de las barras de desviación estándar, puede intentarse colocar las barras de un grupo hacia la parte superior del gráfico y las barras del grupo más próximo hacia la parte inferior.

Una última consideración es la construcción del eje  $x$ . En general, son necesarios entre 8 y 12 intervalos para abarcar el rango de valores. No hacer el eje  $y$  sobrepasando el rango de valores. Esto desaprovecha espacio. De nuevo, hay que considerar si es necesaria una figura. A veces, las tesis y tesinas incluyen ejemplos como el de la figura 19.3. Si miramos la figura, puede verse una interacción fuerte y significativa entre el conocimiento de los resultados (CR) y la consecución del objetivo. Ahora miremos el eje  $y$  en el que se muestra la variable dependiente. Hay que observar que los valores se muestran aproximados a la centésima de segundo. En realidad, hay



riesgo de generar problemas. Lo que hay que poner en la balanza es el grado de riesgo, los derechos de los participantes y el valor potencial de la investigación en su contribución al conocimiento, al desarrollo tecnológico y a una mejor calidad de vida.

## ¿Qué deben esperar los participantes de una investigación?

Tuckman (1978) ha resumido los derechos de los participantes que han de tener en cuenta los investigadores:

- **El derecho a la privacidad o no participación.** Significa que el investigador no puede solicitar información innecesaria y que debe obtener el consentimiento de los adultos y el consentimiento de los tutores en el caso de los niños (así como el consentimiento de los propios niños cuando sea adecuado).
- **El derecho a permanecer en el anonimato.** El investigador debe explicar que el estudio se llevará a cabo con datos del grupo que se registrarán mediante un número identificador (en vez del nombre de los participantes).
- **El derecho a la confidencialidad.** Los participantes deben conocer quiénes tendrán acceso a los datos originales por los que pueden ser identificados (el acceso se facilita al menor número de personas posible).
- **El derecho a exigir responsabilidades al investigador.** El investigador debe ser bienintencionado y sensible con la dignidad humana. Si el participante no conoce el objetivo del estudio (o lo ha malinterpretado), hay que informar al participante inmediatamente después de concluir la prueba.

La investigación cualitativa (explicada con detalle en el cap.18) plantea por sí misma la posibilidad de problemas debido a la estrecha interacción personal con los participantes. El investigador suele dedicar mucho tiempo a los participantes intentando conocerlos y solicitándoles que expongan sus pensamientos y percepciones. Griffin y Templin (1989) proponen aspectos éticos sobre cómo obtener notas de campo: cómo preservar la autoestima de los participantes sin comprometer la precisión del trabajo en caso de conflicto y qué debe hacerse en caso de denuncia (o amonestación) de algún aspecto ilegal o inmoral de la recogida de datos.

No es fácil responder a las consideraciones éticas en el **trabajo de campo**. A veces, la investigación cualitativa tropieza con generadores de desviaciones como los adictos a drogas y los amantes de la transgresión. En algunos casos, es imposible el consentimiento informado. Punch (1986, 36) hace esta afirmación cuando describe su trabajo con la policía. El coche patrulla se dirigió hacia una pelea, y cuando los policías saltaron del coche y empezaron a reducir a los que se peleaban, Punch pensó si debería quedarse *helado* o empujar su cabeza entre los miembros entrelazados y como si fuera Miranda empezar a cantarles sus derechos a los participantes. De forma parecida, cuando Powermaker (citado por Punch) estuvo cara a cara con una multitud de linchadores, ¿debería haber mostrado su placa de identificación académica y explicar a la gente el motivo de su presencia? Con estos dos ejemplos, no estamos diciendo que los investigadores cualitativos están exentos de consideraciones sobre el consentimiento informado y el fraude. Sólo ponemos de manifiesto que determinadas situaciones en investigación cualitativa plantean problemas éticos particulares. Te invitamos a leer la discusión de Punch (1986) y a consultar algunas de las fuentes citadas sobre el tema.

Las personas con discapacidad representan un tema especial como participantes de una investigación. El participante con alguna discapacidad está protegido por el derecho a la privacidad. Así, las instituciones no pueden facilitar los nombres de las personas con discapacidades que pudieran ser participantes potenciales de la investigación. El investigador debe contactar con la institución para solicitar participantes posibles. Entonces, la institución debe pedir permiso al participante o a sus familiares para facilitar sus nombres a los sujetos que pueden participar y el diagnós-

### Trabajo de campo

Metodología utilizada en investigación cualitativa en que los datos se recogen en sus entornos naturales.

cuatro pistas en el número de derrotas). Las diferencias observadas pueden atribuirse al azar. Es probable que Roger sienta en su corazón que dice la verdad (investigación cardiológica).

En algunos casos, las frecuencias esperadas pueden obtenerse de fuentes de información preexistentes como en el siguiente ejemplo. A una nueva profesora asociada, Nancy Niceperson, se le asigna la docencia del amplio programa de introducción a la kinesiología. Pasados algunos semestres, al director del departamento le llegan rumores de que la Dra. Niceperson es demasiado indulgente en sus notas, por lo que se utiliza la ji cuadrado para comparar las notas de los 240 estudiantes con la curva de distribución normal de notas propuestas por el departamento: 3,5% A y F, 24% B y D, y 45% C. Si la Dra. Niceperson siguiese esta distribución normal, cabría esperar que se obtendrían 8 A y 8 F ( $3,5\% \times 240$ ), 58 B y 58 D ( $24\% \times 240$ ), y 108 C ( $45\% \times 240$ ). Se comparan las frecuencias observadas (notas de la Dra. Niceperson) con las frecuencias esperadas (la distribución impuesta por el departamento) mediante ji cuadrado.

### Ejemplo 9.2

#### Valores conocidos

		Notas					Total
		A (3,5%)	B (24%)	C (45%)	D (24%)	F (3,5%)	
Número de notas observadas	$O =$	21	75	114	28	2	240
Número de notas esperadas	$E =$	8	58	108	58	8	240

#### Cálculos auxiliares

	Notas				
	A	B	C	D	F
$(O - E)$	13	17	6	-30	-6
$(O - E)^2$	169	289	36	900	36
$(O - E)^2 / E$	21,13	4,98	0,33	15,52	4,50

$$\chi^2 = \sum [(O - E)^2 / E] = 46,46$$

El director, siempre ecuánime no quiere tomar una mala decisión, por lo que decide utilizar un nivel de probabilidad del 0,01. La tabla A.7 muestra que para 4 gl (hay cinco grados o celdas) se necesita un valor de ji cuadrado de 13,28 para que sea significativa a un nivel de 0,01. La ji cuadrado obtenida de 46,46 es mayor, lo que indica una desviación significativa de la distribución de las notas esperadas. En las notas de la Dra. Niceperson hay demasiadas A y B y pocas D y F. El director hizo la única cosa prudente, ponerla en su sitio.

#### Tabla de contingencia

Clasificación en dos dimensiones de las observaciones y grupos utilizados para calcular la significación de las diferencias entre los valores observados y los esperados.

### La tabla de contingencia

A menudo, un problema implica dos o más categorías de eventos y dos o más grupos (una clasificación bidimensional). Un ejemplo común es el análisis de los resultados de cuestionarios o de inventarios de actitudes en los que hay varias categorías de respuestas (p. ej., de acuerdo, indiferente, en desacuerdo) y dos o más grupos de encuestados (p. ej., aficionados y no aficionados al ejercicio). Este tipo de clasificación bidimensional recibe el nombre de **tabla de contingencia**.

Como aclaración, supongamos que un grupo de atletas y de no atletas responden a la siguiente frase de un diario de un deportista: "Un jugador de béisbol que atrape

- Preparación y presentación de los proyectos de investigación cualitativa .....394
- Redacción de proyectos de becas .....395
- Solicitud de proyectos internos .....396
- Finalización de la tesis o tesina .....396
- Resultados y discusión .....397
- Cómo referenciar experimentos múltiples en una memoria .....400
- Uso de tablas y figuras .....401
- Resumen .....408

**Capítulo 20    Cómo publicar trabajos de investigación    .....411**

- Pautas básicas de redacción .....411
- Unas palabras acerca de los agradecimientos .....412
- Formatos de tesis y tesinas: clásico versus artículo .....412
- Consejos útiles para la redacción de un buen artículo .....419
- Redacción de resúmenes .....421
- Presentación oral y carteles .....424
- Resumen .....428

**■ Apéndices    .....429**

- A    Tablas estadísticas    .....429**
- B    Breve revisión histórica de la investigación en actividad física en Estados Unidos    .....443**
- C    Informática en investigación y ciencias    .....449**
- D    Modelos de formularios de consentimiento    .....455**

Bibliografía    465

Índice de autores    481

Índice alfabético    484

A propósito de los autores    491

cusión de la investigación que estás realizando. También explicamos cómo preparar tablas, figuras e ilustraciones y dónde ubicarlas en la memoria de investigación.

Para terminar, en el capítulo 20 proponemos formas para utilizar los estilos de artículo y el tradicional para estructurar y redactar tesis y tesinas. También presentamos un breve apartado de redacción para revistas científicas y una explicación corta sobre la preparación y exposición oral, además de las presentaciones de carteles.

# Medición de las variables en investigación

*“Un piano es un piano es un piano”. Gertrude Steinway*



Un paso básico del método científico de solución de problemas es la recogida de datos; por esto, es necesario comprender la teoría básica de la medición. (Debemos decir que la medición se expone aquí como una herramienta de investigación, aunque la propia medición constituye un área de investigación.) En este capítulo se exponen los criterios fundamentales de valoración de la calidad de las medidas utilizadas en la recogida de datos científicos: validez y fiabilidad. Se explican distintos tipos de validez y diferentes formas de demostrar la validez y la fiabilidad. (La validez y la fiabilidad de la investigación cualitativa se exponen en el cap. 18.) Concluimos con algunos puntos sobre la medida del movimiento, la medida de las respuestas escritas en soporte papel, la medida del comportamiento afectivo y la medida del conocimiento.

## Validez

Cuando se reúnen los datos sobre los que se basan los resultados, también hay que prestar atención a la validez de las medidas utilizadas. Si, por ejemplo, en un estudio se quieren comparar métodos de entrenamiento de ganancia de fuerza, los investigadores deben tener una medida válida de fuerza para valorar los efectos de los métodos de entrenamiento. La **validez** de la medida indica el grado en que la prueba, o el instrumento, mide lo que se supone que mide. *Así, la validez hace referencia a la seguridad en la interpretación de una prueba, el aspecto más importante en medición.*

Existen diferentes propósitos para el uso de ciertas medidas. Consecuentemente, hay diferentes tipos de validez. Nosotros consideramos cuatro tipos de validez: **lógica**, de **contenido**, de **criterio** y de **constructo**.

Aunque se ha establecido la validez lógica como un tipo individualizado de validez, la Asociación Americana de Psicología y la Asociación Americana de Investigación Pedagógica consideran la validez lógica como un caso especial de la validez de contenido.

## Validez lógica

La validez lógica también se conoce como **validez aparente**, aunque este término no guste a muchos especialistas en medidas. Se hace referencia a la validez lógica cuando la medida refleja la acción que se va a medir. En otras palabras, que la prueba es válida por definición. Una prueba de **equilibrio estático** que consiste en realizar

### Validez

Grado en que una prueba o instrumento de medida pondera lo que intenta medir; puede clasificarse en *validez lógica, de contenido, de criterio y de constructo.*

### Validez lógica

Grado en que una medida implica de forma obvia la acción que está midiendo; también se llama *validez aparente.*

### Validez de contenido

Grado en que una prueba (normalmente en estudios de pedagogía) representa de forma adecuada lo que se ha impartido durante un curso.

bien redactada de una situación puede dar la impresión de un significado holístico y constituye en definitiva una ayuda para el investigador ya que aporta pruebas a sus aseveraciones. Al definir la investigación cualitativa, Locke (1989) afirmó que el investigador debe describir el escenario de educación física con tanto realismo que “tú puedas oler las taquillas y oír el roce de los pies corriendo” (pág. 4). Griffin y Templin (1989, 399) aportan el siguiente ejemplo de una viñeta:

En el segundo período de clases de educación física en la Big City Middle School, se juega a fútbol. El profesor había colocado dos montones de camisetitas en cada extremo de un amplio campo abierto como porterías. No había marcas en el campo. Cuatro chicos corrían arriba y abajo detrás de la pelota. Los demás estudiantes estaban de pie en las posiciones asignadas hasta que les llegaba la pelota, entonces se desplazaban hacia la pelota para chutar. Tres chicas estaban de pie hablando formando un círculo compacto cerca del límite más alejado. Se sobresaltan cuando la pelota cae dentro de su grupo y dos chicos les gritan para que se aparten de la trayectoria. Ellas lo hacen y luego se reagrupan una vez que la pelota y los chicos se dirigen hacia el otro lado del campo. Dos chicos que no habían tocado la pelota durante la clase empiezan un encuentro amistoso de lucha libre cerca de una portería. El profesor está de pie en el centro del campo con un silbato en los labios. No había dicho nada desde que repartió a los estudiantes en equipos al empezar la clase. Ha tocado dos veces el silbato para pitar falta. Los estudiantes juegan a su alrededor como si no estuviese allí. Suena una campana, y todos los estudiantes se quitan los dorsales y se dirigen al edificio de la escuela. Más tarde, el profesor toca el silbato para concluir el partido y da una vuelta por el campo recogiendo los dorsales.

Después de la clase y mientras vuelve al edificio, el profesor dice, “Estos niños son salvajes. Si se puede agotar algo de su energía, no harán mucho escándalo en la escuela. Este grupo en particular no es demasiado listo (se toca la sien), y no tiene mucha estrategia de juego”. (Mira a un chico y a una chica de la clase que están hablando cerca de la puerta de los vestuarios de las chicas). Grita, “Johnson, mueve el culo hacia la ducha y deja de molestar a las chicas”. Me sonrío. “Deberías estar encima de ellos todo el tiempo”. Mira hacia arriba y señala, “Bien, dos [clases] abajo, tres para irse”.

Sin embargo, hay que tener en cuenta que Locke y otros especialistas no sólo contemplan la riqueza de detalles para dar como válida una viñeta narrativa. Siedentop (1989) advierte de que el que un dato sea el esperado no debe depender de la habilidad narrativa del investigador. Según Erickson (1986), una narración válida no sólo es una descripción sino que también es un análisis: “Una historia puede ser una memoria precisa de una serie de sucesos, pero no contener el significado de las acciones desde las perspectivas tomadas por los actores del suceso... Es la combinación de la riqueza y de la perspectiva de interpretación lo que hace que la narración sea válida” (pág. 150). Las viñetas no se mantienen solas. El investigador debe establecer conexiones interpretativas entre las viñetas narrativas y las otras formas de descripción, como las valoraciones directas y el material cuantitativo.

Las valoraciones directas de entrevistas con los participantes tomadas a partir de notas de campo y cintas magnetofónicas o cintas de vídeo son formas distintas de viñetas que enriquecen el análisis y completan la documentación desde el punto de vista del investigador. La valoración directa de distintos individuos puede demostrar acuerdo (o desacuerdo) sobre algún fenómeno. Las valoraciones directas de la misma persona en distintos momentos pueden aportar pruebas de que determinados sucesos son típicos o pueden demostrar un patrón o tendencia de las observaciones a lo largo del tiempo.

Por ejemplo, K.R. Nelson (1988), para demostrar diferencias en los procesos cognitivos de los estudiantes a quienes les dieron clase profesores expertos y noveles, uti-

Lee (1999, cap. 10) sobre este problema en el aprendizaje y acción motora, o un libro clásico sobre el tema, C. Harris (1963).

Este diseño también puede extenderse a formas más complejas. Primero, pueden utilizarse más de dos (pretratamiento y postratamiento) medidas repetidas. Éste es frecuente en ámbitos de fisiología del ejercicio, comportamiento motor y psicología del ejercicio. Dos grupos de participantes formados aleatoriamente en un experimento de comportamiento motor pueden medirse 30 veces o más mientras aprenden una tarea. Los dos grupos pueden diferir en la información que se les haya dado. El diseño puede utilizar como análisis estadístico un ANOVA de medidas repetidas en el segundo factor (ensayos) de 2 (grupos)  $\times$  30 (ensayos). Hay que recordar del capítulo 8 que es muy difícil establecer supuestos para un ANOVA de medidas repetidas con muchas repeticiones. Así, en un diseño como éste, los ensayos pueden estratificarse (p. ej., muchos ensayos promediados, reducción del número de medidas repetidas) en 3 bloques de 10 ensayos o en 5 bloques de 6 ensayos.

A veces, el diseño se amplía de otra manera. Por ejemplo, en el ejemplo 17.1 (un factorial de 3  $\times$  2) podemos añadir un tercer factor con un pretratamiento y un postratamiento. Esto generará un factorial de tres entradas con medidas repetidas en el tercer factor. Todas las versiones de este diseño son susceptibles de la primera amenaza a la validez externa: efectos reactivos o interactivos de la prueba. El pretratamiento puede sensibilizar al participante al tratamiento y esto reduce la posibilidad de extrapolar los resultados a una población no tratada.

### ***Diseño de cuatro grupos de Solomon***

El diseño de cuatro grupos de Solomon es el único diseño experimental verdadero que evalúa de forma específica una de las amenazas a la validez externa: los efectos de la prueba reactivos o interactivos. El diseño es de la forma siguiente:

R	O <sub>1</sub>	T	O <sub>2</sub>
R	O <sub>3</sub>		O <sub>4</sub>
R		T	O <sub>5</sub>
R			O <sub>6</sub>

Esto combina los diseños de grupos aleatorios con los de pretratamiento y postratamiento de grupos aleatorizados. El objetivo concreto es determinar la causa de si el pretratamiento aumenta la sensibilidad de los participantes al tratamiento. Este diseño permite una replicación del efecto del tratamiento (es  $O_2 > O_4$  y es  $O_5 > O_6$ ), calcular la cantidad de cambio debido al tratamiento (es  $O_2 - O_1 > O_4 - O_3$ ), una evaluación del efecto de la prueba (es  $O_4 > O_6$ ) y establecer de qué manera el pretratamiento interactúa con el tratamiento (es  $O_2 > O_5$ ). Así, éste es un diseño experimental muy sólido. Por desgracia, también es un diseño ineficaz, ya que se necesita el doble de participantes. Esto limita mucho su utilización, sobre todo para los doctorandos. Además, no hay una buena vía de análisis estadístico del diseño. La mejor alternativa (una de las que no utilizan todos los datos) es un ANOVA 2  $\times$  2 agrupado de la siguiente manera:

	No T	T
Preevaluado	O <sub>4</sub>	O <sub>2</sub>
No preevaluado	O <sub>6</sub>	O <sub>5</sub>

Así, VI<sub>1</sub> tiene dos niveles (pretratado y no pretratado), y VI<sub>2</sub> tiene dos niveles (tratamiento y no tratamiento). En el ANOVA, el cociente  $F$  de VI<sub>1</sub> determina los efectos del pretratamiento, el  $F$  de VI<sub>2</sub> establece los efectos del tratamiento y el  $F$  de la interacción evalúa la amenaza a la validez externa de la interacción del pretratamiento con el tratamiento. La tabla 17.2 resume el control de las amenazas a la validez en los diseños experimentales verdaderos.

Feltz y Landers (1983) realizan un metaanálisis sobre los efectos del ejercicio mental en el aprendizaje y acción de habilidades motoras. De 60 estudios, se calculó un tamaño del efecto promedio de 0,48, menor que la mitad de una desviación estándar. Se concluyó que el ejercicio mental mejora algo las habilidades motoras, frente a no hacer ninguna práctica.

En un estudio de diferencias de género en la acción motora de niños y adolescentes, Thomas y French (1985) aportan resultados de 64 trabajos realizados en 31.444 individuos. Encontraron diferencias en la acción motora en relación con la edad en 12 de 20 actividades motoras. Estas 12 actividades seguían cuatro formas características de curvas en relación con la edad. La figura 13.1 muestra la curva característica de tres actividades (salto de longitud, carrera de vaivén y fuerza de prensión). Las diferencias son moderadas (entre 0,5 y 0,75 desviaciones estándar antes de la adolescencia, pero luego hay un gran incremento durante y después de la adolescencia (más de 1,5 desviaciones estándar). Se concluye que las diferencias antes de la adolescencia pueden deberse a factores ambientales (tratamiento diferencial por los padres, profesores, entrenadores y compañeros), pero existe una interacción de la biología y del ambiente al principio de la adolescencia. Observa las diferencias entre las curvas de las figuras 13.1 y 13.2. El tamaño del efecto para el lanzamiento es de 1-5 desviaciones estándar en las edades de 3-4 años y aumenta de forma constante a lo largo de la infancia y la adolescencia hasta una diferencia de 3,5 desviaciones estándar a la edad de 18 años. Thomas y French (1958) apuntan que las diferencias tempranas pueden deberse al sustrato biológico y a la influencia cultural y expectativas para niñas y niños.

Carron, Hausenblas y Mack (1996) presentan un metaanálisis sobre la presión social en el ejercicio. Encontraron 87 trabajos con 224 TE calculados a partir de 49.948 individuos. Los resultados indicaron que hay una relación positiva con la presión social, pero escasa o moderada (TE entre 0,2 y 0,5) con el ejercicio. Sin embargo, cuatro variables presentaron una relación moderada o extensa (0,5 a 0,8): “el apoyo familiar y las actitudes hacia el ejercicio, cohesión de las actividades y adhesión, habilidades para el ejercicio de terceras personas, y apoyo y complicidad familiares” (pág. 1).

Utilizamos este ejemplo para demostrar la importancia del metaanálisis. En el metaanálisis perceptivo-motor, el tema de discrepancia sobre el valor de este tipo de entrenamiento parece haberse resuelto: no es beneficioso. En el metaanálisis de las diferencias de género en el  $\dot{V}O_2$  máx, las mayores diferencias parecen deberse a diferencias en la composición corporal más que a diferencias en otros mecanismos. También parece que el ejercicio tiene un efecto positivo sobre el colesterol y sus componentes. Parece que el ejercicio beneficia a los niños, pero la ganancia es relativamente pequeña en relación con los adultos. El ejercicio mental es mejor que no practicarlo, pero no mucho mejor. Las diferencias de género en la actividad motora antes de la pubertad parece que dependen principalmente del entorno, pero el lanzamiento parece un ejercicio en el que la biología tiene mayor importancia antes de la adolescencia. Para terminar, la presión social tiene una relación positiva con el ejercicio, pero sus efectos son pequeños o moderados.

El metaanálisis, cuando se realiza de forma adecuada y se interpreta con precisión, ofrece la posibilidad de reducir una gran cantidad de estudios a unos principios básicos. Estos principios pueden ser las bases para el desarrollo de programas, investigaciones futuras y demostraciones teóricas, así como aplicaciones prácticas, como actividades y entrenamiento.

## Consideraciones metodológicas

Thomas y French (1986) publicaron un trabajo tutelar y un ejemplo de metaanálisis, parte del cual se ha adaptado aquí para presentar los pasos y temas importantes del metaanálisis.

Agradecemos a Karen French y a *Research Quarterly for Exercise and Sport* el permitirnos transcribir parte de este artículo.

Se puede dar una explicación como la siguiente:

Su respuesta será estrictamente confidencial. El cuestionario tiene un número de identificación a efectos exclusivos de envío. Con este sistema de enumeración puedo buscar su nombre fuera de las listas de envío cuando se reenvía el cuestionario. Nunca figurará su nombre en el cuestionario.

Otra estrategia que funciona bien es enviar una carta (con sello y dirección de reenvío) con el cuestionario. La carta contiene el número de identificación y se envía por separado cuando la persona devuelve el cuestionario. La carta sólo dice que se ha enviado el cuestionario y que no es necesario que el investigador envíe más datos. Este método asegura el anonimato e informa al investigador de que se ha respondido el cuestionario.

Es fundamental una buena tasa de respuesta. Con frecuencia, encontramos artículos y comunicaciones en prensa en las que se envían unas 2.000 encuestas, y sólo se reenvían 600. Esto es un número respetable, pero estas 600 personas sólo representan el 30% de reenvíos. No es una muestra aleatoria. Están autoseleccionadas y sus respuestas pueden ser algo distintas de las de las 1.400 que no han respondido por varios motivos.

## Análisis de los resultados y confección del informe

Estos dos últimos pasos se explican en el capítulo 19, sobre la presentación de los apartados de resultados y la discusión del informe del trabajo. La consideración principal es que debe elegirse el método de análisis en la fase de planificación del estudio. Muchos cuestionarios sólo se analizan por recuento de las respuestas a varios ítems y referencias del porcentaje de los encuestados que han respondido de una manera y el porcentaje de los que han respondido de otra manera. Normalmente, no se puede obtener una interpretación detallada de la significación si sólo se realiza un recuento simple. Por ejemplo, cuando el investigador sólo determina que el 18% de los que han respondido están muy de acuerdo con algún criterio, el 29% de acuerdo, el 26% en desacuerdo, el 17% muy en desacuerdo y el 10% no contestan, la reacción del lector puede ser “¿y qué?”. Los cuestionarios, como las encuestas, deben diseñarse y analizarse con el mismo cuidado y razonamiento científico que los estudios experimentales.

## El método Delphi

El **método de encuesta Delphi** utiliza cuestionarios aunque de manera distinta que la encuesta clásica. La técnica Delphi utiliza una serie de cuestionarios de manera que el encuestado alcance un consenso sobre el tema. Es un método que utiliza la opinión de expertos para ayudar a tomar decisiones sobre prácticas, requerimientos y objetivos.

El método incluye la selección de expertos o de personas con conocimientos que responderán a una serie de cuestionarios. Se preparan para su consideración un conjunto de ideas o cuestiones. Cada estadio de la técnica Delphi recibe el nombre de **ronda**. La primera ronda es exploratoria. Se pregunta a los encuestados por sus opiniones en distintos temas, objetivos, y se incluyen preguntas abiertas que permiten a los participantes expresar sus puntos de vista y opiniones.

Luego, se revisa el cuestionario y como resultado de la primera ronda se solicita a los encuestados que reconsideren sus respuestas en base al análisis de todos los encuestados en el primer cuestionario. Se realizan rondas posteriores, y los encuestados disponen de los resúmenes de los resultados previos y se les solicita que revisen la adecuación de sus respuestas. Finalmente, se consigue el consenso mediante rondas sucesivas de análisis y posterior valoración. Es importante el anonimato en el método Delphi; además el consenso entre expertos reconocidos proporciona una vía de com-

### Método de encuesta Delphi

Técnica de encuesta que utiliza una serie de cuestionarios de tal forma que los encuestados (normalmente expertos) alcanzan un consenso sobre el tema.

### Ronda

Estadio del método de encuesta Delphi en el que a los encuestados se les solicitan sus opiniones y valoraciones sobre distintos temas, objetivos y otros aspectos.

primera media milla. Al final de la segunda milla, uno abandonó; al final de la cuarta milla, hubo dos abandonos más; en la quinta, un cuarto hombre se cayó; al completar la quinta milla, abandonó un quinto participante; durante la octava milla, Downes, uno de los más rápidos, y con seguridad el corredor más atlético, se lesionó un pie, y abandonó al final de esa milla, quedando tres competidores a distancia.

Lo siguiente es el orden en que cada participante llegó a la tribuna de los jueces y la posición en cada milla.

	Millas									
	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	7 <sup>a</sup>	8 <sup>a</sup>	9 <sup>a</sup>	10 <sup>a</sup>
Stannard	3	4	3	3	3	2	2	1	1	1
Glarer	2	2	1	1	2	3	3	3	2	2
Mahony	1	1	5	5	5	4	4	4	3	3
Downes	5	3	2	2	1	1	1	2	abandono	
McGargy	6	7	7	7	4	abandono				
Wall	4	5	4	4	abandono					
Sutton	8	8	6	6	abandono					
Mallard	9	9	8	8	caída y abandono					
Vermilyea	7	6	abandono							

Lo siguiente es el tiempo realizado por Stanrard, el ganador, en cada milla. Mahony, el irlandés, corrió la primera milla en cinco minutos y veinticuatro segundos. Las apuestas en el estadio antes y después de empezar fueron numerosas, y se aportaron grandes cantidades tanto a favor como contra el tiempo. El favorito fue Downes; era muy conocido en el vecindario; recorrió las ocho millas en cuarenta y ocho minutos y medio; estaba bien entrenado por su padre, quien a los treinta y nueve años recorrió 17 millas en una hora y cuarenta y cinco minutos, completando las primeras doce millas y media en una hora y quince minutos.

	Min	Seg
1 <sup>a</sup> milla	5	36
2 <sup>a</sup> milla	5	45
3 <sup>a</sup> milla	5	58
4 <sup>a</sup> milla	6	25
5 <sup>a</sup> milla	6	2
6 <sup>a</sup> milla	6	3
7 <sup>a</sup> milla	6	1
8 <sup>a</sup> milla	6	3
9 <sup>a</sup> milla	5	57
10 <sup>a</sup> milla	5	54
	59	44

Mallard tenía fama de ser un gran corredor; había recorrido 16 millas en una hora y cuarenta y nueve minutos, con una parada para cambiarse las zapatillas. No empezó sobrio y se cayó en la quinta milla.

El alemán había cubierto la distancia entre New York y Harlaem, ida y vuelta (doce millas) en setenta minutos; sus amigos estaban muy orgullosos de esto. Había apostado casi 300 \$ a que él conseguiría el premio. En la sexta milla no estaba fuera de tiempo, y recorrió las diez millas en una hora y veintisiete segundos. Estaba con cuatro segundos de retraso en la octava milla. Durante una parte de la carrera lle-

vaba un pañuelo en la boca. Mahony, el irlandés, no ha entrenado demasiado; dejó su equipo en la calle Water, vino a la competición, corrió la primera milla en menos de cinco minutos y medio; al concluir la sexta milla tenía un retraso de un minuto y cuarto; al final de la octava milla, tenía dos minutos de retraso; tres minutos de retraso en la novena, y concluyó las diez millas en setenta y un minutos y tres cuartos. En el 25 del mes pasado, este hombre corrió ocho millas en cuarenta y un minutos cuarenta y seis segundos. McGargy estaba en malas condiciones; pero corrió las cinco millas en treinta y dos minutos y medio. Vermilyea estaba muy delgado y con un pésimo estado de salud; se desplazó andando durante treinta y ocho minutos, el pasado martes, para llegar a tiempo, y al día siguiente recorrió ocho millas en cuarenta y seis minutos; es un corredor excelente, aunque abandonó al finalizar la segunda milla por un dolor en el costado; también recibió un

Continúa

culado en la ecuación 13.4; para más detalles, ver Thomas y French, 1986). Así, los TE más exactos tienen un mayor peso en cada análisis.

### **Pruebas de homogeneidad**

Al metaanálisis se le ha criticado por ser una ensaladilla o por mezclar estudios con distintas escalas de medida, diseños y métodos. No había una prueba adecuada para determinar en qué medida todos los TE estaban estimando el mismo efecto del tratamiento en una población hasta que Hedges (1982) presentó su prueba de homogeneidad. El dato estadístico de homogeneidad,  $H$ , está específicamente diseñado para testar la hipótesis nula,  $H_0$ :  $TE_1 = TE_2 = \dots = TE_j$ . Esto es equivalente a decir que todos los TE probados vienen de la misma población de TE.

El dato estadístico  $H$  es la suma ponderada del cuadrado de las desviaciones de los TE de la media ponderada general. La contribución de cada TE a la media total o general se pondera por el recíproco de su varianza (ecuación 13.4). Los tamaños del efecto con varianzas pequeñas tienen más peso en el cálculo de la media general. Bajo la hipótesis nula,  $H$  sigue una distribución ji cuadrado con  $N - 1$  grados de libertad, donde  $N$  equivale al número de TE.

Cuando no se rechaza la hipótesis nula, todos los TE son similares y representan una medida similar de la eficacia del tratamiento. En este caso, el investigador debe indicar que la estadística de homogeneidad significa que los TE son homogéneos y utilizan TE medio ponderado con intervalos de confianza (ver cap. 6) para su interpretación. Si se rechaza la hipótesis nula, los TE no son homogéneos y no representan una medida similar de eficacia de tratamiento, o de agrupamiento. Se han propuesto dos métodos por Hedges (1982b) y Hedges y Olkin (1985) para buscar modelos explicativos de TE. A continuación presentamos estos modelos.

### **Análisis de la varianza y regresión ponderada**

El primer método para plantear un modelo explicativo de los datos del TE es análogo al ANOVA en que la suma de cuadrados del cálculo total se divide en suma de cuadrados entre dos (o dentro de grupos de) TE ( $H_B$ ) y suma de cuadrados intragrupos de TE ( $H_W$ ). Cada suma de cuadrados puede probarse como una ji cuadrado con  $k - 1$  gl para  $H_B$  y  $N - k - 1$  gl para  $H_W$  (donde  $N$  es el número de TE y  $k$  es el número de grupos). Así, una prueba puede orientarse hacia las diferencias entre –o dentro de– grupos ( $H_B$ ), y puede dirigirse una prueba para determinar la homogeneidad de todos los TE dentro de un grupo ( $H_W$ ). Una exposición más amplia del modelo categórico se encuentra en Hedges (1982b) y Hedges y Olkin (1985).

El segundo método propuesto por Hedges y Olkin (1980) para determinar un modelo explicativo del valor del TE es la técnica de regresión ponderada. Hemos preferido presentar de forma más detallada las técnicas de regresión por varios motivos. Primero, uno de los metaanálisis expuestos anteriormente sobre ejercicio y deporte utiliza técnicas de regresión para analizar el TE (Thomas y French, 1985). Segundo, es frecuente que una variable continua modifique el TE. Tercero, a menudo más de una característica del estudio puede modificar el TE. Esto es especialmente cierto cuando hay varios TE. El método de regresión puede incorporar un gran número de variables en el análisis sin sobredimensionar el nivel de alfa de la prueba estadística. Realizar muchas pruebas utilizando el método categórico u otros parecidos al ANOVA implica una sobrevaloración de alfa, por lo que el investigador debe comunicar el error experimental o ajustar el nivel de alfa mediante la técnica de Bonferroni (ver cap. 8). Cuarto, las variables categóricas pueden ser codificadas con facilidad de manera ficticia o por el efecto e incluirse en los métodos de regresión. Por ejemplo, artículos publicados frente a no publicados pueden tener códigos indicadores (1 y 0) e incluirlos en la regresión. Así, no es necesario realizar un análisis por separado de las variables categóricas.

En la técnica de regresión ponderada, cada TE pondera por el recíproco de su varianza. Muchos paquetes estadísticos comerciales (p. ej., SAS, SPSS, BIMED) ofrecen la opción de calcular la regresión ponderada. Así, el cálculo puede hacerse fácilmente en la mayoría de ordenadores.

- Sage, G.H. (1989). A commentary on qualitative research as a form of inquiry in sport and physical education. *Research Quarterly for Exercise and Sport*, 60, 25-29.
- Sammons, J.T. (1994, winter). Race and sport: A critical, historical examination. *Journal of Sport History*, 21, 203-278.
- Scheffé, H. (1953). A method for judging all contrasts in analysis of variance. *Biometrika*, 40, 87-104.
- Schein, E.H. (1987). *The clinical perspective in fieldwork*. Newbury Park, CA: Sage.
- Scherr, G.H. (Ed.). (1983). *The best of The Journal of Irreproducible Results* (p. 152). New York: Workman Press.
- Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82, 225-260.
- Schmidt, R.A. (1988). *Motor control and learning*. Champaign, IL: Human Kinetics.
- Schmidt, R.A., & Lee, T.D. (1999). *Motor control and learning* (3rd ed). Champaign, IL: Human Kinetics.
- Schutz, R.W. (1989). Qualitative research: Comments and controversies. *Research Quarterly for Exercise and Sport*, 60, 30-35.
- Schutz, R.W., & Gessaroli, M.E. (1987). The analysis of repeated measures designs involving multiple dependent variables. *Research Quarterly for Exercise and Sport*, 58, 132-149.
- Serlin, R.C. (1987). Hypothesis testing, theory building, and the philosophy of science. *Journal of Counseling Psychology*, 34, 365-371.
- Shafer, R.J. (1980). *A guide to historical method* (3rd ed.). Homewood, IL: Dorsey Press.
- Sheets-Johnstone, M. (1999). *The primacy of movement*. Amsterdam/Philadelphia: John Benjamins.
- Shore, E.G. (1991, February). Analysis of a multi-institutional series of completed cases. Paper presented at Scientific Integrity Symposium. Harvard Medical School, Boston.
- Siedentop, D. (1980). Two cheers for Rainer. *Journal of Sport Psychology*, 2, 2-4.
- Siedentop, D. (1989). Do the lockers really smell? *Research Quarterly for Exercise and Sport*, 60, 36-41.
- Siedentop, D., Birdwell, D., & Metzler, M. (1979, March). A process approach to measuring teaching effectiveness in physical education. Paper presented at the American Alliance for Health, Physical Education, Recreation and Dance national convention, New Orleans.
- Siedentop, D., Trousignant, M., & Parker, M. (1982). *Academic learning time-Physical education: 1982 coding manual*. Columbus: Ohio State University, School of Health, Physical Education, and Recreation.
- Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Simon, R.L. (1991). *Fair play: Sports, values, and society*. Englewood Cliffs, NJ: Prentice Hall.
- Singer, R.N., Hausenblas, H.A., & Janelle, C.M. (Eds.). (2001). *Handbook of sport psychology* (2nd ed.). New York: John Wiley and Sons.
- Slavin, R.E. (1984a). Meta-analysis in education: How it has been used. *Educational Researcher*, 13(4), 6-15.
- Slavin, R.E. (1984b). A rejoinder to Carlberg et al. *Educational Researcher*, 13(4), 24-27.
- Smith, M.L. (1980). Sex bias in counseling and psychotherapy. *Psychological Bulletin*, 87, 392-407.
- Snyder, C.W., Jr., & Abernethy, B., (Eds.). (1992). *The creative side of experimentation*. Champaign, IL: Human Kinetics.
- Sparling, P.B. (1980). A meta-analysis of studies comparing maximal oxygen uptake in men and women. *Research Quarterly for Exercise and Sport*, 51, 542-552.
- Spiriduso, W.W. (1987). Graduate program ranking in physical education. *Quest*, 39(2), 103-112.



No todas las distribuciones permiten suponer la normalidad.

las pruebas paramétricas (Harwell, 1990; Thomas, Nelson y Thomas, 1999). Otro inconveniente de las pruebas no paramétricas es que no hay programas informáticos específicos para el cálculo de las pruebas más complejas, como los casos multivariantes.

En el capítulo 6 aportamos un conjunto de métodos para valorar la normalidad de los datos de tal forma que puede determinarse cuándo hay que usar aproximaciones paramétricas o no paramétricas. Esto incluye la observación de la distribución de los datos (p. ej., utilizando un gráfico de tallo y hojas, un histograma de frecuencias para valorar si los datos siguen una curva normal) y evaluar el sesgo y la curtosis de la distribución. Es difícil decidir qué datos permiten suponer o no normalidad (Thomas, Nelson y Thomas, 1999). Por ejemplo, ¿qué grado de normalidad han de tener los datos para poder usar técnicas paramétricas? Hay numerosas ocasiones en que los datos del investigador no permiten la suposición de normalidad. Micceri (1989), por ejemplo, afirma que en pedagogía y psicología muchos de los

datos son no normales en grado moderado o elevado, y deben contemplarse las pruebas no paramétricas. Además, algunas veces, los únicos valores disponibles son frecuencias de observación o rangos (que no suelen tener una distribución normal), y el investigador deberá usar pruebas no paramétricas.

En este capítulo se presentan dos categorías de análisis no paramétricos. El primero es la ji cuadrado, que se utiliza para analizar la frecuencia de respuesta de distintas categorías; por ejemplo:

- Los niños más altos ¿son niñas o niños?
- ¿Cuántos deportistas de más de 30 años participan en deportes de aficionados?
- ¿Cuántas mujeres altas participan con regularidad en natación, carrera o ciclismo?

Segundo, se presenta una aproximación estándar para analizar datos ordenados donde los rangos no siguen una distribución normal. Para poder utilizar las técnicas del modelo lineal general (MLG) descritas en los capítulos 7 y 8, los datos deben seguir una distribución normal, y su gráfica de puntos debe seguir una línea recta. Son ejemplos de datos que no permiten el supuesto de normalidad:

- Respuestas de cuestionarios del tipo muy de acuerdo, de acuerdo, indiferente, en desacuerdo y muy en desacuerdo pueden considerarse datos ordenados; las respuestas se ordenan del 1 (muy de acuerdo) al 5 (muy en desacuerdo).
- Los datos de tiempo, velocidad, aceleración y recuento (p. ej., cuántas elevaciones de tronco mediante extensiones del codo *{push-ups}* pueden hacerse) a veces no están distribuidos normalmente.

En estos casos, no es adecuado el uso de métodos paramétricos correlacionales ni ANOVA (pruebas  $t$  incluidas) debido a que no pueden establecerse los supuestos para el uso de tablas  $r$ ,  $t$  y  $F$ .

En este capítulo proporcionamos un conjunto de métodos para datos ordenados según un rango que son equivalentes a los métodos paramétricos de correlación y ANOVA (pruebas  $t$  incluidas). El enfoque de estos métodos es el mismo que el presentado en los capítulos 7 y 8. El método utiliza los mismos programas informáticos (p. ej., SPSS, SAS, BIMED) para analizar los datos ordenados, pero en vez de utilizar las tablas  $r$ ,  $t$  y  $F$ , se realiza un cálculo (del dato estadístico  $L$ ) para compararlo en una tabla de ji cuadrado en la que no es necesario el supuesto de normalidad.

cometer el otro. El investigador debe decidir el nivel de significación (alfa) que hay que establecer, en otras palabras, el grado de predisposición a estar equivocado.

En estadística, el aspecto más importante es la utilización de un buen criterio. Los hallazgos estadísticos deben interpretarse de acuerdo con el contexto teórico utilizado y los resultados de trabajos previos. Esto conduce al concepto de potencia: estimación de los parámetros –combinación de alfa, tamaño del efecto y tamaño de la muestra– necesarios para determinar en qué medida tienen validez los resultados de un estudio.

## Evaluación de lo aprendido

Los valores siguientes indican el número de flexiones realizadas por 15 estudiantes:

8	4	13
11	6	2
6	7	5
5	0	4
8	6	5

1. Utiliza la ecuación 6.1 para calcular la media.
2. Utiliza la ecuación 6.2 para calcular la desviación estándar.
3. Utiliza la ecuación 6.3 para calcular la desviación estándar, y comprobar que el resultado es el mismo que el que se obtiene con la ecuación 6.2.
4. Reordena los datos de menor a mayor. Utiliza la fórmula  $(N + 1)/2$  para encontrar el punto de la mediana. Cuenta de principio a final el número de los valores para localizar la mediana.
5. Haz una lista con 50 nombres. Utiliza la tabla de números aleatorios (tabla A.1 del apéndice A), selecciona aleatoriamente a 24 individuos para luego distribuirlos aleatoriamente entre dos grupos de 12 cada uno.
6. Construye un diagrama de tallo y hojas utilizando intervalos de 10 valores (p. ej., 10-19, 20-29, etc.) con los 30 datos siguientes:

50	42	64	18	41	30	48	68	21	48
43	27	51	42	62	53	45	31	13	58
60	35	28	46	36	56	39	46	25	49

7. Según la ecuación 6.5, construye un intervalo de confianza para la media de una muestra ( $M = 50$ ,  $s = 6$ ,  $n = 100$ ). Establece un nivel de confianza del 95%. Comenta los resultados.

Por ejemplo, en un estudio de psicología del deporte en que se esté interesado en saber si puede influir en una lectura sobre el uso de esteroides un grupo de jugadores de rugby de una universidad. Además, se cree que las actitudes de los jugadores pueden modificarse por determinados rasgos de personalidad. Entonces, se seleccionan tres pruebas —una prueba de conocimientos sobre los esteroides, un inventario de actitudes sobre el uso responsable de fármacos y una medida de rasgos de personalidad— y se administran a todos los participantes. Las pruebas de conocimientos y de actitudes podrán realizarse antes y después de la lectura, y la prueba de rasgos de personalidad sólo antes de la lectura (los rasgos no deben cambiar y esta prueba sirve para estratificar de alguna manera a los jugadores). En el apartado de metodología hay que describir las tres pruebas y probablemente haya que incluir copias completas de cada una en el apéndice (ver cap. 5 sobre la ética del uso de pruebas estandarizadas). También hay que describir la fiabilidad (coherencia) y la validez (lo que mide la prueba), información que está disponible para cada una de las pruebas en las referencias adecuadas. Luego, hay que explicar las tablas de valoración (incluir una muestra en el apéndice) y los métodos de valoración (pero no hay que utilizar conversiones de las medidas inadecuadas como las del recuadro de la derecha).

Otro ejemplo es un estudio de respuesta motora en la que los tiempos de reacción y de movimiento de los participantes se miden bajo condiciones variables. En el apartado de metodología hay que describir la instrumentación para la prueba y aportar un

esquema o una fotografía. Si la salida del aparato es a través de un ordenador para controlar la situación y la recogida de datos, hay que describir el ordenador (marca y modelo) y cómo se ha establecido la interfaz. Hay que incluir, por lo menos en el apéndice, una descripción del funcionamiento del programa informático (a veces, una copia del programa completo). Hay que explicar las variables dependientes generadas por el tiempo de reacción y de movimiento, y expresar la fiabilidad calculada para estas características. Debe aportarse toda la información necesaria para el uso correcto de los instrumentos (o aparato) en el apartado de metodología y en el apéndice, y procurar que el apartado de metodología sea fluido.

### Conversiones métricas incorrectas

- 1 trillón de micrófonos = 1 megáfono
- 1 millón de bicicletas = 2 megaciclos
- 2.000 burlas = 2 kiloburlas
- 435,6 galletas = 1 libra de pastel
- 10 raciones = 1 decoración
- 10 miriápodos = 1 cienpiés
- 10 monólogos = 5 diálogos

## Descripción de los métodos

En el apartado de métodos hay que describir la forma de obtener los datos, incluyendo los métodos de comprobación en la recogida de los valores de las variables de interés. Es importante describir cómo se han realizado las pruebas y quién las ha realizado. Hay que detallar la planificación del entorno de las pruebas y las instrucciones dadas a los participantes (aunque se pueda poner parte de esta información en el apéndice). Si se trata de un estudio experimental, hay que describir los tratamientos aplicados a los distintos grupos de participantes. Cuando se planifiquen los métodos, hay que tener en cuenta los puntos siguientes:

- Recogida de datos
  - ¿Cuándo? ¿Dónde? ¿Cuánto tiempo será necesario?
  - ¿Tienes datos piloto que demuestren tu capacidad y conocimiento del uso de las pruebas, los equipos y de la manera en que responderán los participantes?
  - ¿Has realizado un esquema de la adquisición, registro y valoración de los datos? (Esto a veces está controlado por un ordenador, para ver si tu ordenador es femenino o masculino mira más abajo.)
- Planificación de los tratamientos (en estudios experimentales y cuasi experimentales).

- Spiriduso, W.W., & Lovett, D.J. (1987). Current status of graduate education in physical education: Program demography. *Quest*, 39(2), 129-141.
- Spray, J.A. (1987). Recent developments in measurement and possible applications to the measurement of psychomotor behavior. *Research Quarterly for Exercise and Sport*, 58, 203-209.
- Spray, J.A. (1989). New approaches to solving measurement problems. In M.J. Safrit & T.M. Wood (Eds.), *Measurement concepts in physical education and exercise science* (pp. 229-248). Champaign, IL: Human Kinetics.
- Stamm, C.L., & Safrit, M.J. (1975). Comparison of significance tests for repeated measures ANOVA design. *Research Quarterly*, 46, 403-409.
- Starkes, J.L., & Allard, F. (Eds.). (1993). *Cognitive issues in motor expertise*. Amsterdam: North Holland.
- Stock, W.A. (1994). Systematic coding for research synthesis. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis*. New York: Sage Foundation.
- Struna, N.L. (1985, June). In "glorious disarray": The literature of American sport history. *Research Quarterly for Exercise and Sport*, 56, 151-160.
- Struna, N.L. (1986, spring). E.P. Thompson's notion of "context" and the writing of physical education and sport history. *Quest*, 38, 24-27.
- Struna, N.L. (1988, December). Sport and society in early America. *International Journal of the History of Sport*, 5, 292-311.
- Struna, N.L. (1996a). *People of prowess: Sport, leisure, and labor in early Anglo-America*. Urbana: University of Illinois Press.
- Struna, N.L. (1996b). Sport history. In J. Massengale & R. Swanson (Eds.), *History of exercise and sport science*, Chapter 9. Urbana: University of Illinois Press.
- Struna, N.L. (2000). Social history and sport. In J.J. Oakley & E.G. Dunning (Eds.), *Handbook of sport studies*, Chapter 12. Beverly Hills, CA: Sage.
- Stull, G.A., Christina, R.W., & Quinn, S.A. (1991). Accuracy of references in the *Research Quarterly for Exercise and Sport*. *Research Quarterly for Exercise and Sport*, 62, 245-248.
- Suits, B. (1978). *The grasshopper. Games, life, and utopia*. Toronto: University of Toronto Press.
- Sundgot-Borgen, J. (1994). Risk and trigger factors for the development of eating disorders in female athletes. *Medicine and Science in Sports and Exercise*, 26, 414-419.
- Taylor, S.J., & Bogdan, R. (1984). *Introduction to qualitative research methods* (2nd ed.). New York: Wiley.
- Tew, J. (1988). Construction of a sport specific mental imagery assessment instrument using item response and classical test theory methodology. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Tew, J., & Wood, M. (1980). *Proposed model for predicting probable success in football players*. Houston, TX: Rice University Press.
- Thomas, J.R. (1977). A note concerning analysis of error scores from motor-memory research. *Journal of Motor Behavior*, 9, 251-253.
- Thomas, J.R. (1980). Half a cheer for Rainer and Daryl. *Journal of Sport Psychology*, 2, 266-267.
- Thomas, J.R., (Ed.). (1983). Publication guidelines. *Research Quarterly for Exercise and Sport*, 54, 219-221.
- Thomas, J.R., (Ed.). (1984). *Motor development during childhood and adolescence*. Minneapolis: Burgess.
- Thomas, J.R. (Ed.). (1986). Editor's viewpoint: Research notes. *Research Quarterly for Exercise and Sport*, 57, iv-v.
- Thomas, J.R. (1987). Are we already in pieces, or just falling apart? *Quest*, 39(2), 114-121.
- Thomas, J.R. (1989). An abstract for all seasons. *NASPSPA Newsletter*, 14(2), 4-5.
- Thomas, J.R., & French, K.E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin*, 98, 260-282.

mos, ya que si el lector puede entender el problema, crecerá su interés por la solución. Por esto, una norma importante es ésta: no hay que ser demasiado técnico. Para fines de comunicación, es más eficaz utilizar un vocabulario decidido, simple y directo que una jerga científica cargada de palabras polisílabas. Day (1983, 147-148) explica una historia clásica de los errores de la jerga científica:

Esto me recuerda al fontanero que escribió a la Agencia de Patentes diciendo que había descubierto que el ácido clorhídrico iba bien para limpiar los desagües atascados. Parafraseando a la Agencia, “Es indiscutible la eficacia del ácido clorhídrico, aunque su efecto corrosivo es incompatible con el mantenimiento del metal”. El fontanero contesta que está muy orgulloso de que la Agencia esté de acuerdo. La Agencia vuelve a intentarlo escribiendo: “No podemos asumir la responsabilidad de la producción de los residuos tóxicos y nocivos del ácido clorhídrico y recomendamos la utilización de un método alternativo”. El fontanero vuelve a responder que está muy orgulloso de que la Agencia esté de acuerdo con él. Al final, la Agencia escribió al fontanero: “No utilice ácido clorhídrico. Destroza las tuberías”.

Es importante tener en cuenta a la audiencia. A veces, puede suponerse que los lectores están bastante informados sobre el tema (o puede que esto sea lo primero que leen). Sin embargo, incluso un lector bien informado necesita refrescar alguna información general para comprender la naturaleza del problema, poder interesarse lo suficiente y apreciar tu razonamiento al plantear el problema. Debes recordar que tu audiencia no está completamente ni recientemente inmersa en el área concreta de investigación en la que tú te encuentras.

Los párrafos introductorios deben generar interés por el estudio; por ello, tu redacción habilidosa y sabia sobre el tema es especialmente valiosa en la introducción. El texto debe presentar la información general necesaria y explicar de forma rápida la motivación del estudio. [Una introducción agradable, unificada y bien escrita sirve para presentar el estado del problema con la claridad suficiente para que el lector pueda comprender el objetivo del estudio antes de leerlo con detalle.](#)

Hemos seleccionado las introducciones siguientes de revistas de investigación por su brevedad en la presentación y por su claridad. Esto no quiere decir que la brevedad por sí misma constituya un criterio ya que algunos temas requieren introducciones más extensas que otros. Por ejemplo, los estudios de desarrollo o validación de un tema teórico suelen requerir introducciones más largas que los de un tema de investigación aplicada. Además, las tesis y tesinas (en formato tradicional) pueden tener introducciones más o menos largas sólo por el coste de las páginas en el futuro.

## Ejemplos de buenas introducciones

Los ejemplos siguientes muestran algunas características deseables en una introducción que incluyen una introducción general, información de fondo, una referencia de los vacíos en la literatura y áreas en que es necesaria la investigación y una progresión lógica que conduzca a la definición del problema. Después de haberlas leído, comprueba si puedes escribir el objetivo de cada estudio.



A veces es fácil identificar el problema.

tintamente. La narración hace referencia a la forma de razonamiento histórico, sobre todo la crónica de algo del pasado. Entonces, la narración se sitúa en la creación de una historia descriptiva. Tanto las formas narrativas como las explicativas de razonamiento pueden encontrarse en la historia analítica. Ver Mandelbaum, *Anatomy*; Robert F. Atkinson, *Knowledge and Explanation in History* (Ithaca, NY, Cornell University Press, 1978); Dale H. Porter, *The Emergence of the Past: A Theory of Historical Explanation* (Chicago, University of Chicago Press, 1981); Allan Mefill, "Recounting the Past; 'Description' Explanation, and Narrative in Historiography", *American Historical Review* 94 (junio 1989): 627-650. Algunos historiadores también describen la historia de síntesis, que se estructura sobre historias existentes (fuentes secundarias) para obtener generalizaciones. Ver Thomas Beuder "Wholes and Parts: The Need for Synthetics in American History", *Journal of American History* 73 (junio 1986): 120-136.

<sup>12</sup> E.P. Thompson, *The Poverty of Theory and Other Essays* (New York, Monthly Review Press, 1978), 27-28. Parfraseando a Thompson: "Un historiador tiene derecho a... Hacer supuestos provisionales... que la prueba que ha obtenido tiene una existencia 'real' determinante e independiente de su existencia dentro de las formas de pensamiento, y esta prueba actúa como testigo de un proceso histórico real, y este proceso (o algún conocimiento aproximado de éste) es el objeto del conocimiento histórico". Y prosigue, "La prueba histórica es la que... no descarta su propio significado, pero lo cuestiona..." (págs. 28-29).

<sup>13</sup> Oscar Handlin, *Truth in History* (Cambridge, MA, Belknap Press, 1979), 120. Ver también Stephen R. Humphrey, "The Historian, His Documents, and Elementary Modes of Historical Thought", *History and Theory* 19 (1980): 1-20.

<sup>14</sup> Cuantitativa, o numérica, la prueba es lenta, pero es un resultado preciso en investigación histórica en kinesiología, ciencias del ejercicio y educación física. Es necesario utilizar muchas más cosas, sobre todo en los casos en que los historiadores quieran establecer condiciones, motivos de cambios y categorías como las clases sociales. La revista más útil de técnicas cuantitativas en historia es *Historical Methods*. Un texto práctico metodológico es Konrad H. Jarausch y Kenneth A. Hardy, *Quantitative Methods for Historians: A Guide to Research, Data, and Statistics* (Chapel Hill, University of North Carolina Press, 1991).

<sup>15</sup> Un ejemplo clásico de documento de autenticidad cuestionable que causó un gran impacto en la historia del deporte fue el de un esquema de un campo de béisbol atribuido a Abner Doubleday cuando era un niño de edad escolar en 1839. Abner Braves, que dijo haber sido compañero de clase de Doubleday y presenciar el episodio del dibujo, incluyó el croquis en una carta a la Comisión Mills, que estableció el origen del béisbol en 1907. Esta comisión aceptó el croquis y el testimonio de Braves como "prueba" de que Doubleday había inventado el juego (ver Shafer, *Historical Methods*, 127-147; Handlin, *Truth in History*, 111-124). Para una exposición maravillosa y simpática de un documento sospechoso y los problemas que genera, ver John D. Milligan, "The Treatment of an Historical Source", *History and Theory* 18 (1979): 177-196.

<sup>16</sup> Shafer, *Historical Methods*, 149-170; Handlin, *Truth in History*, 124-144.

<sup>17</sup> Shafer, *Historical Methods*, 150-158.

<sup>18</sup> Handlin, *Truth in History*, 165-226.

<sup>19</sup> Adrian Wilson, "Inferring Attitudes From Behavior" *Historical Methods* 14 (verano 1981): 140-144.

<sup>20</sup> E.P. Thompson, "Anthropology and the Discipline of Historical Context", *Midland History* 3 (primavera 1972): 41-55.

<sup>21</sup> "The Great Foot Race", *American Turf Register and Sporting Magazine* 6 (junio 1835): 518-520.

<sup>22</sup> Henry W. Fowler y F.G. Fowler, *The King's English* (Oxford, Oxford University Press, 1954); Savoie Lottinville, *The Rhetoric of History* (Norman, University of Oklahoma Press, 1976).

En EE.UU., la formación de posgrado, en particular la del doctorado, se basa en la formación de posgrado alemana. El espíritu de la universidad alemana de la búsqueda del conocimiento y la investigación con un objetivo de producción se ha extrapolado en gran medida a América (Rudy, 1962). A pesar de todo, a lo largo del tiempo ha habido algunos cambios en los requisitos para alcanzar el grado de doctor, aunque los objetivos básicos y las expectativas han permanecido inalteradas. El grado de doctor se concede como reconocimiento a un aspirante por su contribución académica y a la investigación en un ámbito específico del conocimiento mediante la aportación de conocimiento y teoría inéditas (Boyer, 1970).

Dicho de manera más sencilla, la investigación es la base de todo programa de doctorado, y el ensayo es el aspecto principal del grado de doctor. Se ha dicho que escribir la tesis representa una media del 39% del tiempo dedicado a la obtención del grado de doctor en los campos de bioquímica, ingeniería electrónica, psicología, física, sociología y zoología (Porter, Chubin, Rossini, Boeckmann y Connally, 1982).

## Un argumento a favor del formato de artículo

El objetivo principal de una tesis o tesina es aportar conocimientos nuevos con nivel científico. (*American Psychological Association*, 1959; Berelson, 1960; Porter y Wolfe, 1975). Un concepto muy repetido en las revistas universitarias es que las tesis y las tesinas proporcionan pruebas de competencia en la planificación, dirección y redacción de trabajos científicos. En términos de objetivo de programa, el estudio es una vivencia de aprendizaje valiosa en la cual las tesis y las tesinas son ejercicios prácticos en la ejecución de los distintos pasos del método científico para la resolución de problemas. Incluso el “ABD” (*all but done* [casi doctores], o quien ha acabado todo el trabajo para el doctorado excepto la exposición) agradece la contribución que la lectura de tesis aporta a la ciencia y al método científico (Jacks, Chubin, Porter y Connally, 1983).

Es un hecho la importancia de la comunicación de los resultados de investigación como una parte integrante del proceso de investigación. Por este motivo, uno de los objetivos de las tesis y tesinas es servir de vehículo de los resultados de la investigación inédita llevada a cabo por el estudiante de doctorado. Así, la tesina y tesis se vuelven parte del proceso de comunicación. A.L. Porter *et al.*, (1982) afirman que las tesinas representan una contribución importante a la base del conocimiento y que los autores de las tesinas publicadas se citan con más frecuencia que otras publicaciones que hayan escrito.

Además de la aportación potencial que puede hacer una tesis o tesina en un ámbito de estudio, en realidad sólo entre un tercio y la mitad de las tesis (incluidas algunas tesinas) están disponibles en publicaciones para los profesionales (McPhie, 1960; Porter *et al.*, 1982). Por supuesto, hay muchos motivos para no publicar nunca una tesis o una tesina. Uno de ellos es que, a pesar del interés de las instituciones por la investigación, muchos estudiantes no consideran la investigación como un objetivo importante. A. L. Porter *et al.*, (1982) afirman que el 24% de los estudiantes de doctorado encuestados manifestaban esta idea. Arlin (1977) va más lejos afirmando que la mayoría de los educadores nunca realizan otra publicación tras acabar la tesis o tesina. Además, el mundo laboral establece distintos grados de importancia a la investigación y a las publicaciones. También es un hecho innegable que no todas las tesis y tesinas merecen ser publicadas.

Es posible que otro factor que contribuye a la baja tasa de publicación es el estilo y el formato clásicos utilizados a menudo en las tesis y tesinas. Los becarios y los doctorandos más motivados dedicarán tiempo y esfuerzos a reescribir la tesina en un formato propio de artículo de revisión, pero los doctorandos menos moti-



En alguna ocasión, a los editores de revistas no les gustará tu artículo de investigación.

que respetar el límite de tiempo, y que resulta imposible presentar un trabajo completo dentro de este límite ¿qué se puede hacer? Proponemos presentar los hallazgos más importantes del artículo dividiendo el tiempo de presentación en fracciones de los 15 minutos de la presentación. Son útiles los soportes visuales como diapositivas, programas de presentación por ordenador o con transparencias.

- Introducción donde se citen algunos estudios importantes: 3 min
- Estado del problema: 1 min
- Método: 3 min
- Resultados (presentación de tablas y figuras; normalmente son mejores las figuras ya que las tablas son arduas de leer): 3 min
- Discusión (puntos principales): 2 min
- Preguntas y debate: 3 min
- Total: 15 min

En una presentación oral, los errores más frecuentes son la utilización de un tiempo excesivo en la metodología y una presentación pobre de los resultados (en el recuadro siguiente se muestra una presentación oral pobre de resultados). Un uso adecuado del soporte visual es la clave de una buena presentación oral, sobre todo de los resul-

### No utilices estas citas de comentaristas deportivos ingleses en tu presentación oral

“Moses Kiptanui, el keniatá de 19 años que cumplió 20 hace unas semanas.”

“Nos encontramos exactamente en la misma situación que al principio de la carrera, sólo que justamente a la inversa.”

“Nunca le han practicado una cirugía de rodilla en ninguna otra parte del cuerpo.”

“Ella no es Ben Johnson, pero entonces, ¿quién es?”

“Les debo mucho a mis padres, especialmente a mi madre y a mi padre.”

“El estadio Port Elizabeth es más circular que ovalado. Es largo y cuadrangular.”

“La cancha es tan llana como una bola de billar.”

“Mira la hora, te dará una idea de lo rápido que están corriendo.”

“Distas varias pulgadas de ser un milímetro perfecto.”

“Si la historia se repite a sí misma, creo que podemos esperar lo mismo de nuevo.”

tados. Inserta una breve reseña del problema en una diapositiva y muéstrala mientras realizas la exposición. Una diapositiva de los métodos experimentales reduce mucho la verborrea al explicar la metodología. Los esquemas de los resultados (sobre todo figuras y gráficas) son más eficaces que las tablas o la presentación oral. Hay que hacer las figuras y las gráficas simples y concisas. **Las tablas y las figuras que se han preparado para la redacción de un artículo raramente son útiles como soportes visuales de las presentaciones orales. Normalmente, contienen mucha información y no son lo suficientemente grandes.** Hay que tener un puntero a mano para señalar los aspectos más importantes. Para terminar, recuerda las cinco leyes de Thomas y Nelson de las presentaciones



Utilizar soportes visuales ayuda a mantener el interés de la audiencia durante las presentaciones científicas.

# Informática en investigación y ciencias



Ordenadores personales en análisis estadístico y redacción científica

Ordenadores en las medidas del movimiento y análisis de grupos de datos

deben presentarse como un acto de voluntad. Deben residir en lo que los conductistas llaman “la caja negra”. Esta caja metafórica se coloreó de negro ya que no puede verse a su través. El psicólogo conductista B.F. Skinner afirma que puede obviarse el sistema nervioso central (SNC), o conceptual, debido a que nunca podrá conocerse lo suficientemente bien para utilizar este conocimiento. Muchos psicólogos contemporáneos creen en la importancia de la cognición aunque intentan limitar el estudio a la física, química, fisiología o teoría computacional (Hamlyn, 1990; Dennett, 1991). Aunque se atreven a entrar dentro de la caja negra, a menudo rechazan interpretar lo que significa. Al igual que la mayoría de sus predecesores empíricos, rechazan tomarse en serio sus experiencias vitales y sus ideas.

Entre los que se toman las ideas en serio, se encuentran los investigadores que utilizan metodologías cualitativas (ver cap. 18). Sin embargo, los filósofos y los investigadores cualitativos utilizan la subjetividad de maneras muy distintas. Normalmente, los investigadores cualitativos trabajan con alguna forma de datos recogidos de manera sistemática: observación, entrevistas, conversaciones, interpretaciones de objetos de la cultura. Mientras que los filósofos, aunque no tienen restringido el uso de información, normalmente no recogen los datos de esta manera, no lo hacen si no encuentran un motivo para hacerlo. Pueden empezar su investigación con una idea, por ejemplo, un problema teórico abstraído de la experiencia.

Tanto los investigadores cualitativos como los filósofos emiten juicios, evaluaciones o interpretaciones que no pueden demostrarse, probarse, ni siquiera reducirse a un nivel de probabilidad estadística. Pero el trabajo cualitativo invita a especular sobre si un grupo de personas, por ejemplo, realmente piensan o sueñan. Los filósofos están menos interesados en establecer estas atribuciones, sobre todo para decir que saben lo que pasa por la cabeza de cualquiera.

Los filósofos de la actividad física discuten sobre qué *debe conocerse* acerca del ejercicio, sobre el valor de la salud, qué *tiene sentido* al comparar ejercicio con deporte, qué *significado pueden tener* experiencias estéticas de “flotar” en la ingravidez, qué *motivos subyacen* en el comportamiento ético en relación con la ingesta de drogas en el deporte.

Puede concluirse que la misión principal de la filosofía es observar la realidad mediante técnicas de reflexión, normalmente sin recogida sistemática de datos empíricos ni interés por hacer afirmaciones sobre lo que determinadas personas piensan, creen o sienten en determinadas circunstancias. Debido a que ésta es una definición muy general de la investigación filosófica, es útil identificar subdivisiones conocidas o ramas del pensamiento reflexivo:

- Metafísica
- Axiología
- Epistemología
- Poesía

En **metafísica**, los filósofos intentan analizar la *naturaleza de las cosas*. El enfoque de este trabajo comprende un espectro que va desde las simples distinciones (p. ej., separar deporte de danza) hasta las tesis más complejas y controvertidas (p. ej., describir la naturaleza de la perfección en competiciones).

El trabajo en una segunda rama de la filosofía llamada **axiología** se centra en el *valor de las cosas*, en logros, adquisiciones o estado de las cosas como la forma física, salud, conocimientos y excelencia (teorías sobre valores amorales), en el comportamiento humano, como saltarse o acatar las reglas del juego (teoría acerca de la ética), y finalmente en el arte y la belleza, como las cualidades de determinadas rutinas en gimnasia (investigación estética). Los debates de las páginas 251 a 256 profundizan en el segundo subdominio de la axiología: la ética.

Una tercera rama de la filosofía, la **epistemología**, trata de la *obtención del conocimiento* (¿qué hay que hacer para saber cosas?) y la *comprensión de este estado* (¿qué importancia tiene el conocimiento y con qué criterios se puede determinar su naturaleza?). Se realiza una investigación lógica (p. ej., determinar la fuerza de cualquier relación aparentemente estable entre cooperación y competición en el deporte), de

---

### Metafísica

Naturaleza de las cosas.

---

### Axiología

Valor de las cosas.

---

### Epistemología

Adquirir y comprender conceptos.

significativamente mayor de no atletas que están de acuerdo en que el jugador debe informar al árbitro sobre la falta al atrapar la pelota, mientras que una proporción mayor de atletas creen que no debería hacerlo.

## Restricciones de uso de la ji cuadrado

Aunque se ha dicho que la estadística no paramétrica no necesita los mismos supuestos acerca de la población que la estadística paramétrica, hay algunas restricciones para el uso de esta técnica. Las observaciones deben ser independientes, y las categorías, mutuamente excluyentes. Esto indica que una observación de cualquier categoría no debe estar relacionada con otra observación de otra categoría ni debe depender de ella. Por ejemplo, se puede preguntar a 50 personas sobre sus actividades preferidas. Si cada una refiere tres preferencias, no está justificado utilizar un total ( $N$ ) de 150, debido a que las preferencias de cada individuo pueden estar relacionadas y la ji cuadrado, sobredimensionada. Una observación debe ponerse en una única categoría. Las frecuencias observadas son exactamente esto: el número de eventos. Los cocientes y porcentaje no son adecuados. Otro aspecto relacionado es que el total de las frecuencias esperadas y el total de las frecuencias observadas de cualquier clasificación deben ser iguales. En el ejemplo 9.3 observamos, por ejemplo, que el total de las frecuencias esperadas de los atletas es el mismo que el total de sus frecuencias observadas. Lo mismo ocurre en los no atletas y en los totales de las columnas.

Normalmente, no se puede aplicar ji cuadrado en muestras pequeñas. La frecuencia esperada para cada celda no puede ser menor de 1,0. Además, algunos estadistas refieren que no más de un 20% de las celdas pueden tener valores esperados menores de 5. Al respecto, hay distintas opiniones. Algunos afirman que ninguna celda puede tener menos de 5, mientras que otros estadistas permiten que un 40% de las celdas tengan una frecuencia menor de 5. Un truco frecuente en los casos con muchas celdas con frecuencias esperadas menores de 5 es combinar categorías adyacentes, incrementando así los valores esperados.

Normalmente, los investigadores aceptan que una tabla de contingencia de  $2 \times 2$  debe tener la llamada corrección de continuidad. Esta corrección, habitualmente llamada **corrección de continuidad de Yates**, consiste en restarle 0,5 a la diferencia entre las frecuencias observadas y las esperadas de cada celda antes de elevarla al cuadrado:

$$\chi^2 \text{ corregida} = \sum [(O - E - 0,5)^2 / E] \quad (9.2)$$

Otra limitación de la tabla de contingencia  $2 \times 2$  es que el número total ( $N$ ) debe ser de 20 como mínimo.

Para terminar, la distribución esperada debe ser lógica y establecida antes de la recogida de datos. En otras palabras, la hipótesis (probabilidad, eventos, datos censales, etc.) ha de preceder al análisis. A los investigadores no se les permite buscar entre las distribuciones posibles la que se ajuste a sus hipótesis.

## Coefficiente de contingencia

Muchas técnicas correlacionales pueden utilizarse en estudios con datos discretos (p. ej., no continuos). **Se puede calcular la relación entre variables dicotómicas, como sexo y raza, utilizando el coeficiente de contingencia.** La prueba de significación es ji cuadrado. Recordemos que la tabla de contingencia se utiliza para detectar diferencias entre grupos de datos, tal como se ha descrito en el apartado anterior. También puede usarse una tabla de contingencia para determinar relaciones. En una tabla de contingencia puede haber cualquier número de filas o de columnas. Después de calcular la ji cuadrado, se puede calcular el coeficiente de contingencia  $C$ :

$$C = \sqrt{\chi^2 / (N + \chi^2)} \quad (9.3)$$

### Corrección de continuidad de Yates

Método de corrección de una tabla de contingencia  $2 \times 2$  mediante la resta en cada celda de 0,5 de la diferencia entre la frecuencia observada y la esperada antes de elevarla al cuadrado.

### Coefficiente de contingencia

Método de cálculo de la relación entre dos variables dicotómicas, como género y raza.

el MANCOVA y no lo utilizan adecuadamente. Lo mismo ocurre con ANCOVA; para una buena exposición del tema, ver A.C. Porter y Raudenbush (1987).

## Medidas repetidas con variables dependientes múltiples

Es frecuente que los experimentos tengan variables dependientes múltiples que se miden más de una vez (en el tiempo). Por ejemplo, en un estudio de adherencia, se miden las variables fisiológicas y las psicológicas una vez a la semana durante las 15 semanas del programa de entrenamiento. Si hay dos grupos de entrenamiento (distintos niveles de entrenamiento) y un grupo control, cada grupo con 15 individuos, y todos se valoran semanalmente (p. ej., 15 veces) para dos medidas psicológicas y tres fisiológicas, obtendremos un diseño de 3 niveles de ejercicio  $\times$  15 intentos ( $3 \times 15$ ) de 5 variables dependientes. Este diseño puede analizarse de distintas formas.

Se pueden hacer 5 ANOVA de  $3 \times 15$  con medidas repetidas en el segundo factor. Aquí se pueden seguir los métodos de medidas repetidas descritos anteriormente en este capítulo. Sin embargo, se puede sobrevalorar alfa al realizar análisis múltiples de los mismos individuos. Por supuesto que alfa puede ajustarse con la técnica de Bonferroni ( $\alpha = 0,05/5 = 0,01$ ), pero esto impide valorar la relación entre las variables dependientes, que puede ser importante y de interés. Un manual excelente sobre cómo abordar este problema es el de Schutz y Gessaroli (1987). Los conceptos siguientes están tomados de este trabajo, pero para el uso de este tipo de diseños hay que leer el texto completo y los ejemplos.

Hay dos opciones para este análisis: el modelo mixto de análisis multivariante (MMM) y el análisis multivariante doble (MD). Se utilizan en función de los supuestos que se realizan sobre los datos. En el estudio expuesto previamente (3 niveles de ejercicio  $\times$  15 réplicas de 5 variables dependientes), el análisis MMM contempla la variable independiente (niveles de ejercicio) como un caso multivariante verdadero al construir combinaciones lineales con las cinco variables dependientes y discriminar entre los niveles de la variable independiente. Si esta combinación es significativa, puede seguirse con una técnica  $F$  de reducción (para un método alternativo, ver Schutz y Gessaroli, 1987). Para el factor de medidas repetidas (y la interacción), se elabora una combinación lineal para cada intento, y se realiza un análisis regular de medidas repetidas para cada combinación lineal de los intentos. Esto significa que debe realizarse el supuesto de esfericidad tal como ya se ha descrito y puede utilizarse  $\epsilon$  para evaluar este supuesto con los mismos estándares descritos. La interpretación del cociente  $F$  calculado para los efectos principales de los grupos, intentos y la interacción grupo  $\times$  intento es igual que en los otros diseños. La cuestión que se plantea es ¿pueden cambiar los grupos a diferentes tasas a través de los intentos en la combinación lineal de las variables dependientes? Éste es el análisis recomendado si se puede hacer el supuesto de esfericidad ya que muchos autores creen que es más sólido. Sin embargo, éste es un supuesto difícil; sobre todo si hay más de dos o tres variables dependientes ponderadas en más de tres a cinco intentos.

El análisis MD no necesita la suposición de esfericidad. El análisis es igual que para la variable independiente ejercicio. En el apartado del análisis de medidas repetidas, se forma una composición lineal no sólo de las variables dependientes de cada intento sino también de los 15 intentos (y ellas mismas forman una composición lineal), de ahí el nombre de *multivariante doble*. Sigue siendo casi igual la interpretación de  $F$  de los dos efectos principales y de la interacción mientras que los seguimientos son más complejos.

McCullag y Meyer (1997) compararon cuatro métodos de proporcionar la información (ejercicio con retroalimentación, modelo de aprendizaje con modelo de retroalimentación, modelo de corrección con modelo de retroalimentación y modelo de aprendizaje sin modelo de retroalimentación) en el aprendizaje de la forma correcta de levantarse desde cuclillas sin cargar peso. Hay dos variables dependientes (resul-

## Muestra de carta de portada #2

### *Justificación del tema*

Apreciado \_\_\_\_\_:

Las responsabilidades adquiridas como director de educación física de un amplio distrito escolar han sido la causa de que me replantee el enfoque necesario para alcanzar unos objetivos aceptables en educación física en los programas de secundaria.

Con la cooperación del Distrito Unificado de San Diego, la Escuela de Magisterio y el Departamento Escolar de Educación Física de la Universidad de California, Los Ángeles, y el Gabinete de Educación de la Salud, Educación Física y Ocio del Ministerio de Educación del Estado de California, he iniciado un estudio en el área de educación física. Representantes del Consejo de la Presidencia en Educación Física han dicho que esta información también será valiosa para ellos. El objetivo del estudio es establecer criterios que puedan utilizarse en la planificación de actividades para las clases de educación física en secundaria (niños de 12-18 años) para desarrollar y mantener determinadas características de su forma física. Las características seleccionadas hacen referencia a la resistencia cardiorrespiratoria, fuerza y resistencia musculares, y elasticidad. La información obtenida se utilizará en la planificación de programas, ayudando a responder algunas cuestiones como: ¿Qué partes de las clases de educación física pueden utilizarse para determinadas actividades de mantenimiento de la forma física (como estiramientos), intervalos de carrera, pesas, etc.)? y ¿qué parte de las clases de educación física puede aplicarse a deportes y juegos de nuestra cultura (como baloncesto, voleibol, carreras, tenis, gimnasia, danza, etc.)? ¿Qué guías hay que establecer para planificar determinadas actividades de educación física? ¿Qué actividades de educación física deben utilizarse?

Un comité de fisiólogos del ejercicio le han nominado como persona calificada para opinar sobre este tema debido a su reconocido prestigio como director e investigador en el área de educación física. Sería de agradecer que aceptara responder a las preguntas del cuestionario y devolverlas en el sobre ya preparado para su mayor comodidad. Se le remitirá un resumen de las respuestas.

Afectuosamente:

Especialista en educación física.

© Escuela del Distrito Unificado de San Diego. Reimpresa con permiso.

La carta de seguimiento debe ser delicada. La persona no debe sentirse amonestado por no responder. Lo mejor es escribir dando a entender que la persona podría haber respondido, pero no lo ha hecho por algún descuido o error del investigador (ver en la pág. 294 una muestra de carta de seguimiento).

La no respuesta es especialmente frecuente cuando el cuestionario hace referencia a alguna área sensible. Por ejemplo, las encuestas sobre programas ofertados y de prácticas no suelen contestarse por escuelas con programas inadecuados y pocas prácticas. Así, es probable que las respuestas obtenidas estén sesgadas en favor de los mejores programas y los mejores profesores. Al margen de la naturaleza de la

# Relaciones entre variables

*El tabaco es una de las principales razones de ser de la estadística. Fletcher Knebel*



**E**n el capítulo 6 prometimos que tras exponer alguna información básica que ayude a comprender las técnicas estadísticas, empezaríamos a explicar con detalle algunas técnicas específicas. Empezaremos por la correlación.

La correlación es una técnica estadística utilizada para determinar la relación entre dos o más variables. En este capítulo presentamos varios tipos de correlación, la significación y la fiabilidad de los coeficientes de correlación, así como el uso de la correlación en las predicciones, además de las correlaciones parcial y semiparcial, y las ecuaciones de regresión múltiple. Para acabar, repasaremos brevemente las formas de correlación multivariante: correlación canónica, análisis factorial y modelización estructural.

## Lo que explora la investigación correlacional

A menudo, el investigador se interesa por el grado de relación entre, o correlación de, resultados, como la relación entre los resultados entre una carrera de longitud y una prueba del escalón como medidas de la función cardiovascular. A veces, un investigador quiere establecer la relación entre rasgos y comportamiento, como en qué manera los rasgos de personalidad están relacionados con la práctica de deportes de alto riesgo. Aún quedan más problemas de investigación correlacional que pueden implicar las relaciones entre medidas antropométricas como la medida de pliegues cutáneos y el porcentaje de masa grasa determinado mediante pesada hidrostática. En este caso, el investigador podría querer predecir el porcentaje de masa grasa a partir de la medida de pliegues cutáneos.

La correlación puede implicar dos variables, como en la relación entre el peso y la altura. También puede implicar tres o más variables, como cuando se estudia la relación entre un estado (variable dependiente), como el estado del sistema cardiovascular, y dos o más variables predictivas (variables independientes), como el peso corporal, el porcentaje de masa grasa, velocidad, fuerza muscular y otras. Esto es una correlación múltiple. Otra técnica, la correlación canónica, establece las relaciones entre dos o más variables dependientes y dos o más variables independientes. El análisis factorial utiliza las correlaciones entre un número de variables para identificar las relaciones subyacentes o factores. Para terminar, la modelización estructural aporta pruebas de cómo las variables pueden influir sobre otras variables de manera directa o indirecta.

### Correlación

Técnica estadística utilizada para calcular la relación entre dos o más variables.

### Prueba del escalón

Prueba utilizada para valorar el estado del sistema cardiorrespiratorio realizada mediante medidas de la frecuencia cardíaca después de subir y bajar un escalón varias veces.

## Una excelente tabla mejorada: de la mejor manera posible

Tabla 2. Datos de artículos del volumen 59, 1988

Primer autor	Información del TE	N <sup>c</sup>	TE* principal (de los efectos más importantes del estudio)		
Era	Sí	5-6	0,50*	0,10	1,42*
Simard	Sí	7	-2,71*	-1,59*	0,52
Kamen <sup>a</sup>	Sí	9	0,14	0,64	0,72
Kamen	Sí <sup>b</sup>	10	0,81	0,90	1,14
Kokhonen	Sí	9-12	-2,64*	-1,97*	-1,78*
Nelson <sup>b</sup>	Sí	13	0,73	0,85	1,76
Houst	Sí	20	-2,11*	-0,53*	0,25*
Hutcheson	Sí	34	-0,30*	-0,06	0,63*
Alexander	Sí	26-48	0,33	0,39	0,73*
Farrell	Sí	45-368	-0,51*	0,37*	0,77*
Abernethy	No				
Berger	No				
Doody	No				
Etnyre	No				
Heinert	No				
Ober	No				
Stewart	No				
Wesson	No				

\* TE es significativo,  $p < 0,05$ .

<sup>a</sup> Por comparación de grupo.

<sup>b</sup> Comparación de  $M_S$  determinantes del TE.

<sup>c</sup> El efecto principal es significativo, pero no hay información de la significación de la comparación posterior.

<sup>d</sup> Sin efectos principales significativos.

Otros ejemplos de uso irracional de las cifras se encuentra en las referencias a las estadísticas calculadas. Aunque las interfaces de salida de los ordenadores presentan las estadísticas (p. ej.,  $F$ ,  $r$ ) y probabilidades ( $p$ ) con cinco o más dígitos decimales, esto no significa que los números deban expresarse a este nivel. Lo idóneo es que tengas dos o (como máximo) tres decimales. Sin embargo, esto puede ser erróneo en los cálculos de probabilidades:  $t(22) = 14,73$ ,  $p < 0,000$ . Que  $p < 0,000$  significa que no es posible el error, y esto no puede ser debido a que debe haber probabilidad de error, si no ¿cómo podría ser una probabilidad? Lo que ha ocurrido es que la probabilidad exacta es algo parecido a  $p < 0,0003$  y el investigador la ha redondeado a  $p < 0,000$ . Esto no puede hacerse. Tal como se ha dicho antes, creemos que es más adecuado expresar la probabilidad exacta (p. ej.,  $p = 0,025$ ) y si esta probabilidad sobrepasa el límite alfa del experimento (p. ej.,  $p < 0,05$ ). Sin embargo, el último dígito de la probabilidad debe ser siempre 1 o mayor. El ejemplo anterior ( $p < 0,00023$ ) si se expresa con tres decimales deberá ser  $p < 0,001$ .

Hay otras cifras que suelen utilizarse de forma equívoca como las siguientes. En la revisión de una revista científica encontramos un artículo en el que se sometía a niños a un tratamiento de 12 semanas. El autor incluye la media de edad y la desviación estándar de los niños antes y después del tratamiento de 12 semanas. No es sorprendente que todos los niños tengan 12 semanas más de edad. El autor también ha cal-

**Columna 62. Puntuación media de la acción del grupo**

**Columna 63. Desviación estándar del grupo para la acción**

**Paso 9. Proporcionar información sobre cualquier otro grupo de participantes incluido en el estudio**

**Columna 64. Número de código del grupo, y los grupos pueden ser uno de los siguientes:**

- (6) De competición a nivel internacional u Olímpico
- (5) De competición a nivel nacional
- (4) De competición a nivel regional
- (3) Aficionados al deporte a nivel local/nivel recreativo.
- (2) De competición en deportes distintos del estudiado
- (1) No deportistas

**Columna 65. Calidad de diferenciación del grupo (expertos, elite, hábiles, experimentados, etc.)**

- (2) Bien definido
- (1) No hay suficiente información
- (0) No definido

**Columna 66. Grado de confianza en la selección de la acción**

- (2) Alto
- (1) Promedio
- (0) Bajo

**Columna 67. ¿Cómo se ha determinado esta categoría?**

- (3) Grabaciones de acciones previas
- (2) Criterio del entrenador
- (1) Otros criterios
- (0) No se menciona

**Columna 68. Número total de participantes del grupo**

**Columna 69. Sexo de los participantes**

- (3) Todos hombres
- (2) Todas mujeres
- (1) Mixto
- (0) No establecido

**Columna 70. Edad promedio de la muestra del grupo**

- (5) > 30
- (4) 20-29
- (3) 16-19
- (2) 13-15
- (1) < 12
- (0) No establecida

**Columna 71. Media de años de experiencia en el deporte**

- (5) > 10
- (4) 7-10

Tabla 16.1. *Odds ratio* de cáncer de mama en relación con el ejercicio

Tiempo total de ejercicio <sup>a</sup>	<i>n</i> casos <i>n</i> controles	OR univariante (nivel de confianza del 95 %)	OR multivariante (nivel de confianza del 95 %) <sup>b</sup>
Ninguna	195/154	1,0	1,0
0,1–0,7	103/90	0,92 (0,65–1,32)	0,95 (0,64–1,41)
0,8–1,6	84/102	0,66 (0,46–0,95)	0,65 (0,45–0,96)
1,7–3,7	103/100	0,92 (0,58–1,17)	0,80 (0,54–1,17)
≥ 3,8	61/99	0,50 (0,34–0,73)	0,42 (0,27–0,64)

a Horas por semana, desde la menarquía hasta unos 12 meses antes del estudio.

b Ajustado por edad de la menarquía, edad del primer embarazo a término, número de embarazos a término, meses de lactancia, historia familiar de cáncer de mama, índice de masa corporal en la fecha de referencia y uso de anticonceptivos orales.

Reimpresa, con permiso, de Bernstein *et al.*, 1994, "Physical exercise and reduced risk of breast cancer in young women", *Journal of the National Cancer Institute* 86: 1404-1408. Licencia de adaptación de Oxford University Press.

Manson, Hankinson y Colditz, 1999; Thune, Brenn, Lund y Gaard, 1997) apoyan los resultados presentados anteriormente. Esto es, altos niveles de actividad física acumulada durante la vida de las mujeres reduce el riesgo de padecer cáncer de mama.

### *Ventajas de los estudios caso-control*

A menudo, el diseño del estudio de caso-control es el primer paso de la investigación analítica ya que puede proporcionar estimadores válidos de la relación exposición-enfermedad en menos tiempo y con menos gasto monetario que el diseño de cohortes. Los diseños de caso-control:

1. Son particularmente eficaces para estudiar enfermedades poco comunes debido a que no necesitan poblaciones muy grandes ni períodos de seguimiento largos para acumular el número necesario de casos para el análisis.
2. Permiten contrastar hipótesis de exposiciones múltiples para un único resultado de enfermedad.
3. Proporcionan más recursos para recoger con más detalle la información sobre la exposición debido a su menor escala.

### *Inconvenientes de los estudios caso-control*

Los inconvenientes principales de los estudios caso-control se deben al hecho de que la información sobre la exposición se obtiene una vez diagnosticada la enfermedad, y al desafío de reclutar un grupo control adecuado. Los datos de la exposición y los *odds ratio* resultantes pueden verse alterados por problemas potenciales en cada uno de los aspectos siguientes. Primero, los casos pueden referir sus exposiciones de manera distinta a los participantes control sólo por el hecho de que su diagnóstico reciente de enfermedad a menudo pone en peligro la vida. Los sesgos de este tipo reciben el nombre de **sesgos de respuesta**. Por ejemplo, en un estudio caso-control en el que se estudia la relación entre la actividad física y la cardiopatía, los participantes enfermos (casos) pueden referir que eran menos activos de lo que realmente eran desde que saben que el ejercicio protege contra la cardiopatía. Esto puede incrementar la diferencia aparente entre los niveles de actividad física y los controles que refieren sus exposiciones sin sesgos de respuesta. En este ejemplo, los sesgos de respuesta pueden sesgar los *odds ratio* más allá del valor nulo de 1,0 y afirmar que hay una relación más fuerte de la que hay en realidad.

Segundo, para comparar las historias de exposición entre casos y controles para poder obtener *odds ratio* válidos, los participantes del grupo control necesitan ser representativos de la población de la que se han obtenido los casos. Esto significa que el grupo control no es extrapolable a la población general, sólo a la población de la

### Sesgos de respuesta

Errores sistemáticos introducidos por diferencias en la precisión de la respuesta al hacer comparaciones entre grupos (p. ej., entre casos y controles).