



Algunos aspectos teóricos y prácticos en la elaboración de encuestas en el ámbito escolar

Antonio Tinajas Ruiz

*Agregado de Educación Física
IES Joan Ramon Benaprés (Sitges)*

Palabras clave

encuesta, población, muestra, fiabilidad, error, muestreo aleatorio simple, muestreo aleatorio sistemático, cuestionario

Abstract

The survey is one of the most instruments of research in social sciences and is frequently used in the physical education scene. With the object of boosting its use by physical education professionals, the theoretical and practical principles that have to be followed in making a survey are explained in an understandable way. Special attention is dedicated to the determination of the example and to the selection and make up of the questionnaire. Lastly, the principles explained are applied to a practical case consisting in the realisation of a survey between physical education students of an institute.

Resumen

La encuesta constituye uno de los instrumentos más importantes de investigación en las ciencias sociales y es utilizada frecuentemente en el ámbito de la educación física. Con el objetivo de potenciar su uso por los profesionales de la educación física, se explican de una manera fácil de entender los principios teóricos y prácticos básicos que hay que seguir para elaborar una encuesta. Se dedica una especial atención a la determinación del tamaño de la muestra y su selección y a la elaboración del cuestionario.

Por último, se aplican los principios explicados a un caso práctico consistente en la realización de una encuesta entre el alumnado de educación física de un instituto.

Introducción

La encuesta es uno de los mejores instrumentos de investigación que tenemos a nuestro alcance en el ámbito de las ciencias sociales. Nos permite alcanzar el mejor conocimiento posible de la realidad con el menor coste económico y con un ahorro considerable de esfuerzo y de tiempo.

La encuesta ha sido muy utilizada en diferentes ámbitos de la educación física: desde el ámbito de la educación propiamente dicha (Bone y Contreras, 1985; Tinajas, 1996; Escudero, 1984), pasando por el conocimiento de la extensión social de la práctica físico-deportiva (García, 1991; Cañellas y Rovira, 1995; Serra, 1995), el uso de las instalaciones deportivas (Martínez-Tur, Tordera y Ramos, 1996) o la realidad laboral de los profesionales de la educación física y el deporte (Martínez del Castillo y otros, 1991; Martínez del Castillo, 1993).

Sin embargo, el esfuerzo que comporta la realización de una encuesta puede echarse a perder por la falta de rigor durante la ejecución de las diferentes etapas de la investigación. Quizá la etapa determinante es la elección del tamaño de la muestra y su selección. Una muestra insuficiente o mal escogida hace imposible la extrapolación posterior de los datos obtenidos de la muestra al conjunto de la población. Otra etapa importante es la elaboración del cuestionario.

El objetivo de este artículo no es hacer una descripción extensa y rigurosa del procedimiento que hay que seguir a la hora de llevar a cabo una investigación mediante una encuesta. Para eso ya existen buenos libros de referencia (Colton, 1986; Moser y Kalton, 1985;

AMPLITUD DE LA POBLACIÓN	AMPLITUD DE LA MUESTRA PARA MÁRGENES DE ERROR ABAJO INDICADOS					
	±1%	±2%	±3%	±4%	±5%	±10%
500					222	83
1.000				385	286	91
1.500			638	441	316	94
2.000			714	476	333	95
2.500		1.250	769	500	345	96
3.000		1.364	811	517	353	97
3.500		1.458	843	530	359	97
4.000		1.538	870	541	364	98
4.500		1.607	891	549	367	98
5.000		1.667	909	556	370	98
6.000		1.765	938	566	375	98
7.000		1.842	949	574	378	99
8.000		1.905	976	580	381	99
9.000		1.957	989	584	383	99
10.000	5.000	2.000	1.000	588	385	99
15.000	6.000	2.143	1.034	600	390	99
20.000	6.667	2.222	1.053	606	392	100
25.000	7.143	2.273	1.064	610	394	100
50.000	8.333	2.381	1.087	617	397	100
100.000	9.091	2.439	1.099	621	398	100
∞	10.000	2.500	1.111	625	400	100

p = proporción (en porcentaje) de los elementos portadores del carácter considerado. Si p es <50% la muestra necesaria es más pequeña.

Tabla 1. Tamaño de la muestra en función del tamaño de la población finita y del error para una $p = q = 50\%$ y un nivel de fiabilidad del 95,5% (Mateo, 1989, p. 292).

Rodríguez, 1991). El propósito es mucho menos ambicioso, pero a la vez más interesante: explicar de una manera comprensible los principios teóricos y prácticos que hay que seguir cuando elaboramos una encuesta y allanar el camino para que, quien lo desee, se adentre en el mundo de la estadística descriptiva y sus aplicaciones en el ámbito de la educación física.

El trabajo está explicado como si quisiéramos hacer una encuesta entre el alumnado de ESO de un gran instituto. El procedimiento sería exactamente el mismo si la encuesta fuese dirigida al alumnado de secundaria del Baix Llobregat, del Tarragonés o a todo el profesorado de educación física de Cataluña.

¿Qué queremos saber?

Lo primero que hay que hacer cuando ultimamos el reto de encuestar a un grupo de personas es un esfuerzo de reflexión para concretar aquello que queremos averiguar

mediante la encuesta. Se pueden preguntar muchas cosas, pero no todas tienen el mismo interés cuando el objetivo principal de la encuesta es, por ejemplo, obtener información para mejorar la actuación pedagógica del departamento de educación física, introducir modificaciones en el diseño curricular del área o averiguar los intereses de nuestro alumnado.

El tamaño de la muestra

Una vez concretados los objetivos, el primer problema que hemos de resolver cuando hacemos una encuesta es la determinación del tamaño de la muestra. Este dependerá, en primer lugar, del tamaño de la población. Si la población es muy pequeña, como cuando queremos conocer la opinión del grupo-clase sobre aspectos del comportamiento de su profesor de educación física, la muestra ha de coincidir necesariamente con toda la población. Dicho de otra manera: hay que encuestar a todos los alumnos de la clase.

Por otra parte, cuando la población de la que queremos conocer una opinión empieza a ser un poco grande (1.000 personas) podemos planteamos el escoger una muestra más pequeña que la población. El tamaño de la muestra en relación con el total de la población será tanto más pequeño cuanto más grande sea esta población. Por ejemplo, fijadas tres condiciones que más adelante explicaremos ($p = q = 0,5$; error = 5%; fiabilidad = 95,5%), la muestra sería de 222 individuos para una población de 500 (44,4%), de 286 para una población de 1.000 personas (28,6%) y de 385 para una población de 10.000 (3,8%) (tabla 1).

Veamos qué significan estas tres condiciones que determinan el tamaño de la muestra en función del tamaño de la población.

1. Primera condición: $p = q = 0,5$ (50%). Las letras p y q hacen referencia a las preguntas dicotómicas, es decir, aquellas que sólo aceptan dos respuestas excluyentes, por ejemplo:

- Indica tu sexo (hombre-mujer).
- ¿Pertenece a un club deportivo? (sí-no).

Sin embargo, una pregunta que acepta varias respuestas se puede considerar como un conjunto de preguntas dicotómicas. Tomemos como ejemplo esta pregunta y unos hipotéticos resultados en tanto por ciento:

- En general, consideras que la calificación que recibes en la asignatura de educación física es:

1. Superior a tus méritos (21%).
2. Coincidente con tus méritos (40%).
3. Inferior a tus méritos (39%).

Esta pregunta se puede descomponer en tres subpreguntas dicotómicas:

- a) ¿la calificación que recibes en la asignatura de educación física es superior a tus méritos? (sí, 21% - no, 79%).
- b) ¿la calificación que recibes en la asignatura de educación física es coincidente con tus méritos? (sí, 40% - no, 60%).
- c) ¿la calificación que recibes en la asignatura de educación física es inferior a tus méritos? (sí, 39% - no, 61%).

Si el porcentaje de personas que contestan sí a la subpregunta a es x , entonces el porcentaje de personas que contestan no a la misma subpregunta es $100-x$, es decir, la suma de porcentajes de las personas que contestan sí a las respuestas 1 y 2. En las preguntas dicotómicas la proporción (en tanto por uno) de una de las respuestas se denomina con la letra p y la proporción de la otra respuesta posible se denomina q . Obviamente, $p + q = 1$ y, por tanto, $q = 1 - p$ y $p = 1 - q$.

Pues bien, el tamaño de la muestra depende de la proporción de personas que se prevé escogerán una de las respuestas. El tamaño de la muestra es máximo cuando las proporciones p y q son iguales ($p = q = 0,5$) y es mínimo cuando p (o q) se aproxima a 0 y, por tanto, q (o p) se aproxima a 1 (tabla 2). Como que a menudo no se conoce cuál es la distribución de las respuestas, lo que se hace es coger una muestra con un tamaño adecuado a la peor de las posibles distribuciones, es decir, $p = q = 0,5$. Si después resultase que p y q son diferentes de 0,5 signifi-



ficaría que, de haberlo sabido, podríamos haber cogido una muestra más pequeña. Ahora bien, si hubiésemos cogido una muestra más pequeña suponiendo erróneamente una distribución de p y q diferente de 0,5 y después resultase que la distribución obtenida se aproxima más a 0,5 (50% para cada respuesta), significaría que los datos obtenidos no tendrían la fiabilidad necesaria para poder extrapolarlos al conjunto de la población.

Siguiendo con el ejemplo de la pregunta sobre las calificaciones en la asignatura de educación física, si la distribución de respuestas fuese el resultado de una encuesta anterior y se tratase de volverla a pasar, no sería necesario hacer la suposición de que la proporción de una de las tres respuestas podría ser del 50%. Por los datos de una encuesta anterior ya sabemos que la peor distribución de las tres respuestas se mueve alrededor de una $p = 0,40$ y una $q = 0,60$ (es la distribución que se acerca más a 0,5/0,5). El tamaño de la muestra para una distribución como ésta es menor que el tamaño de una muestra para una distribución 0,5/0,5. Cuando el número de encuestas

es muy grande y su coste considerable, vale la pena aprovechar la información previa de que disponemos para hacer este tipo de correcciones, sobre todo cuando las distribuciones son muy diferentes (por ejemplo, 0,2/0,8). Sin embargo, este aprovechamiento sólo es posible cuando todas las preguntas (y subpreguntas) cumplen la condición citada. Si hubiese una pregunta en el cuestionario para la que no conociésemos ningún dato y tuviéramos que considerar $p = q = 0,5$, la muestra, al menos para esta pregunta, debería ser superior. Esto supone tal nivel de complicación práctica que casi siempre se considera que la proporción de personas que pueden escoger una respuesta y la contraria es la misma ($p = q = 0,5$).

En cuanto al error, nos indica el intervalo en el que podemos esperar, con una cierta probabilidad, que se encuentre la verdadera media (u otro estadístico) de la población que estamos estudiando. Si la media de las respuestas a una pregunta es 4,56 y el error previsto es del 5%, querrá decir que la verdadera media estará comprendida (con una probabilidad de la que a continuación ha-

blaremos) entre $4,56 + 4,56 \times 0,05$ y $4,56 - 4,56 \times 0,05$, es decir, $4,56 \pm 0,23$ (4,33 – 4,79).

Por último, ¿qué significa una fiabilidad del 95%? Pues, siguiendo con el ejemplo anterior, es la probabilidad que hay de que la verdadera media se encuentre comprendida entre 4,79 y 4,33, dicho de otra manera, de cada 100 mediciones que hacemos de la media hay que esperar que 95 caigan dentro del intervalo 4,79-4,33.

Como es fácil de imaginar, el tamaño de una muestra crecerá a medida que disminuya el error máximo que estemos dispuestos a aceptar. También lo hará si aumenta la fiabilidad del estadístico que estemos calculando (una proporción, una media...). En investigación social es habitual aceptar errores entre el 3 y el 5% con una fiabilidad del 95% o del 95,5%.

La elección de la muestra

Una vez resuelto el problema de cómo determinar el tamaño de la muestra, hay que enfrentarse con otro: cómo escoger la

LÍMITES DE ERROR $\pm 2 \sigma$ en % (E)	VALORES PRESUMIBLES DE p Y DE q EN % ($p + q = 100$)													
	1/99	2/98	3/97	4/96	5/95	10/90	15/85	20/80	25/75	30/70	35/65	40/60	45/55	50/50
0,1	39.600	78.400	116.400	153.600	190.000	360.000	510.000	640.000	750.000	840.000	910.000	960.000	990.000	1.000.000
0,2	9.900	19.600	29.100	38.400	47.500	90.000	127.500	160.000	187.500	210.000	227.500	240.000	247.500	250.000
0,3	4.400	8.711	12.933	17.067	21.111	40.000	56.667	71.111	83.333	93.333	101.111	106.667	110.000	111.111
0,4	2.475	4.900	7.275	9.600	11.875	22.500	31.875	40.000	46.875	52.500	56.875	60.000	61.875	62.500
0,5	1.584	3.136	4.656	6.144	7.600	14.400	20.400	25.600	30.000	33.600	36.400	38.400	39.600	40.000
0,6	1.100	2.178	3.233	4.267	5.278	10.000	14.167	17.778	20.833	23.333	25.278	26.667	27.500	27.778
0,7	808	1.600	2.376	3.135	3.878	7.347	10.408	13.061	15.306	17.143	18.577	19.592	20.204	20.408
0,8	619	1.225	1.819	2.400	2.969	5.625	7.969	10.000	11.719	13.125	14.219	15.000	15.469	15.625
0,9	489	968	1.437	1.896	2.346	4.444	6.296	7.901	9.259	10.370	11.235	11.852	12.222	12.346
1,0	396	784	1.164	1.536	1.900	3.600	5.100	6.400	7.500	8.400	9.100	9.600	9.900	10.000
1,5	176	348	517	683	844	1.600	2.267	2.844	3.333	3.733	4.044	4.267	4.400	4.444
2,0	99	196	291	384	475	900	1.275	1.600	1.875	2.100	2.275	2.400	2.475	2.500
2,5	63	125	186	246	304	576	816	1.024	1.200	1.344	1.456	1.536	1.584	1.600
3,0	44	87	129	171	211	400	567	711	833	933	1.011	1.067	1.100	1.111
3,5	32	64	95	125	155	294	416	522	612	686	743	784	808	816
4,0	25	49	73	96	119	225	319	400	469	525	569	600	619	625
4,5	20	39	57	76	94	178	252	316	370	415	449	474	489	494
5,0	16	31	47	61	76	144	204	256	300	336	364	384	396	400
6,0	11	22	32	43	53	100	142	178	208	233	253	267	275	278
7,0	8	16	24	31	39	73	104	131	153	171	186	196	202	204
8,0	6	12	18	24	30	56	80	100	117	131	142	150	155	156
9,0	5	10	14	19	23	44	63	79	93	104	112	119	122	123
10,0	4	8	12	15	19	36	51	64	75	83	91	96	99	100
15,0	2	3	5	7	8	16	23	28	33	37	40	43	44	45

Tabla 2. Tamaño de la muestra de una población infinita en función del error y de la distribución de p y de q para un nivel de fiabilidad del 95,5% (Mateo, 1989, p. 290).

54463	22662	65905	70639	79365	67382	29085	69831	47058	08186	59391	58030	52098	82718
15389	85205	18850	39226	42249	90669	96325	23248	60933	26927	99567	76364	77204	04615
85941	40756	82414	02015	13858	78030	16269	65978	01385	15345	10363	97518	51400	25670
61149	69440	11286	88218	58925	03638	52862	62733	33451	77455	86859	19558	64432	16706
05219	81619	10651	67079	92511	59888	84502	72095	83463	75577	11258	24591	36863	55368
41417	98326	87719	92294	46614	50908	64886	20002	97365	30976	95068	88628	35911	14530
28357	94070	20652	35774	16249	75019	21145	05217	47286	76305	54463	47237	73800	01017
17783	00015	10806	83091	91530	36466	39981	62481	49177	75779	16874	62677	57412	13215
40950	84820	29881	85966	62800	70326	84740	62660	77379	90279	92494	63157	76593	91316
32995	64157	66164	41180	10089	41757	78258	96488	88629	37231	15669	56689	35682	40844
96754	17676	55659	44105	47361	34833	86679	23930	52249	27083	99116	75486	84989	23476
34357	86040	53364	71726	45690	66334	60332	22554	90600	71113	15696	10703	65178	90637
06318	37403	49927	57715	50423	67372	63116	48888	21505	80182	97720	15369	51269	69620
62111	52820	07243	79931	89292	84767	85693	73947	22278	11551	11666	13841	71861	98000
47534	09243	67879	00544	23410	12740	02540	54440	32949	13491	71628	73130	78783	75691
98614	75993	84460	62846	59844	14922	48730	73443	48167	34770	40501	51089	99943	91843
24856	03648	44898	09351	98795	18644	39765	71058	90368	44104	22518	55576	98215	82068
96887	12479	80621	66223	86085	78285	02432	53342	42846	94771	75112	30485	62173	02132
90801	21472	42815	77408	37390	76766	52615	32141	30268	18106	80327	02671	98191	84342
55165	77312	83666	36028	28420	70219	81369	41943	47366	41067	60251	45548	02146	05597
75884	12952	84318	95108	72305	64620	91318	89872	45375	85436	57430	82270	10421	05540
16777	37116	58550	42958	21460	43910	01175	87894	81378	10620	73528	39559	34434	88596
46230	43877	80207	88877	89380	32992	91380	03164	98656	59337	25991	65959	70769	64721
42902	66892	46134	01432	94710	23474	20423	60137	60609	13119	78388	16638	09134	59880
81007	00333	39693	28039	10154	95425	39220	19774	31782	49037	12477	09965	96657	57994
68089	01122	51111	72373	06902	74373	96199	97017	41273	21546	83266	32883	42451	15579
20411	67081	89950	16944	93054	87687	96693	87236	77054	33848	76970	80876	10237	39515
58212	13160	06168	15718	82627	76999	05999	58680	96739	63700	37074	65198	44785	68624
70577	42866	24969	61210	76046	67699	42054	12696	93758	03283	83712	06514	30101	78295
94522	74358	71659	62038	79643	79169	44741	05437	39038	13163	20287	56862	69727	94443
42626	86819	85651	88678	17401	03252	99547	32404	17918	62880	74261	32592	86538	27041
16051	33763	57194	16752	54450	19031	58580	47629	54132	60631	64081	49863	08478	96001
08244	27647	33851	44705	94211	46716	11738	55784	95374	72655	05617	75818	47750	67814
59497	04392	09419	89964	51211	04894	72882	17805	21896	83864	26793	74951	95446	74307
87155	13428	40293	09985	58434	01412	69124	82171	59058	82859	65988	72850	48737	54719
98409	66162	95763	47420	20792	61527	20441	39435	11859	41567	27366	42271	44300	73399
45476	84882	65109	96597	25930	66790	65706	61203	53634	22557	56760	10909	98147	34736
89300	69700	50741	30329	11658	23166	05400	66669	48708	03887	72880	43338	93643	58904
50051	93137	91631	66315	91428	12275	24816	68091	71710	33258	77888	38100	03062	58103
31753	85178	31310	89642	98364	02306	24617	09609	83942	22716	28440	07819	21580	51459

Tabla 3. Tabla de números aleatorios (Mateo, 1989, p. 287).

muestra. La elección más fiable de la muestra es la que utiliza el método aleatorio simple. Es muy sencillo. Se trata de numerar toda la población sometida a estudio (1, 2, 3, 4, ... n) y escoger aleatoriamente tantos números como personas tenga la muestra. Para escoger los números se utiliza una tabla de números aleatorios (tabla 3). Si, por ejemplo, el número de alumnos de un instituto (la población objeto de estudio) es

1.258, deberemos coger números de 4 cifras de la tabla, ya sea por columnas, ya sea por filas. Aquellos números que sean superiores a 1.258 no los tendremos en cuenta. Tampoco tendremos en cuenta aquellos que salgan repetidos. Veamos dos ejemplos.

Si leemos cada columna de cifras de arriba a abajo (comenzando por la columna de la izquierda y yendo después a la derecha), ob-

tendremos los siguientes números de cuatro cifras: 5186, 0421, 4393, 0649... Pues bien, formarían parte de nuestra muestra los alumnos con los números de lista 421 y 649, etc. Si leemos por filas, también de izquierda a derecha, los números que obtenemos son: 5446, 3226, 6265, 9057, 0639... En este caso, el primer alumno o alumna de la muestra será el que tenga el número de lista 639.



Este método es muy lento cuando la muestra es muy grande. Entonces se utiliza el método llamado muestreo aleatorio sistemático. Si tenemos una población de 1.000 personas y hemos de escoger una muestra de 100, lo que hacemos es coger una de cada diez personas ($1000/100 = 10$) empezando por una persona escogida al azar entre las diez primeras.

Si utilizamos la tabla 3 por filas encontramos que el primer número de dos cifras comprendidos entre el 01 y el 10 es el 06 (cuarta columna de la primera fila). Las personas de la muestra tendrían los siguientes números de orden de lista: 6, 16, 26, 36, 46, etc. Este método, sin embargo, no lo podemos utilizar si tenemos la población ordenada según un criterio que varía sistemáticamente con el orden de lista de la población. Si tuviésemos un listado de la población ordenado de manera que se alternase un chico con una chica, la aplicación de este método haría que toda la muestra estuviese formada exclusivamente por chicos o exclusivamente por chicas. Una forma de utilizar un muestreo aleatorio sistemático minimizando el riesgo de que se produzca este tipo de problemas es utilizar listados ordenados alfabéticamente.

La representatividad de la muestra en relación con el conjunto de la población puede mejorar mucho si conocemos la existencia de subgrupos dentro de la población que pueden tener una opinión diferente. Un caso típico es el de los subgrupos resultantes de la clasificación de la población en función de la edad, del sexo o de ambas variables. Si creemos que la opinión de nuestro alumnado no depende de la edad en el intervalo estudiado (12-16 años), pero que sí depende del sexo, deberemos hacer un muestreo que represente proporcionalmente el número de chicos y de chicas de la población del centro (a esto se le llama estratificar la muestra). Confeccionaremos dos listados, uno de los chicos y otro de las chicas, y escogeremos una muestra en que los chicos y las chicas estén en la misma proporción que en la población. Para la elección de los chicos y de las chicas de cada listado utilizaremos un muestreo aleatorio simple o sistemático.

El cuestionario

Una vez aclarado qué queremos saber hemos de decidir de quién queremos obtener esta información. En un centro con alumnado entre los 12 y casi los 20 años es normal que encontremos un amplio abanico de intereses y de opiniones y, lo que es más importante, una gran diferencia en cuanto al espíritu crítico hacia la realidad que los rodea. Por tanto, habrá que plantearse a menudo la conveniencia de elaborar una sola encuesta (la muestra de la que puede ser estratificada o no) para todo el centro o confeccionar más de una.

A continuación viene la labor de redacción de las preguntas. Las preguntas han de tener estas características:

1. Han de ser cortas y con un vocabulario sencillo que las haga comprensibles.
2. Han de ser precisas, de manera que todas las personas encuestadas las entiendan de la misma manera.
3. Es preferible que las respuestas sean cerradas (todos han de escoger de entre las que ofrece el cuestionario, sin poder dar la suya).
4. No ha de ser posible que una persona no pueda encontrar entre las respuestas cerradas una que incluya su opinión.
5. Las respuestas han de ser excluyentes entre sí.

En cuanto al diseño de las respuestas, estas pueden ser cualitativas o cuantitativas. Si escogemos respuestas de tipo cualitativo hemos de tener muy en cuenta a la hora de escoger los adjetivos de manera que exista una simetría entre las expresiones utilizadas, preferiblemente con un término intermedio: muy bueno - bueno - normal - malo - muy malo. Esto no siempre es posible. Con las respuestas cuantitativas o numéricas hay que seguir los mismos criterios, pero es mucho más fácil. Se acostumbra a utilizar 3 escalas numéricas: del 1 al 5, del 1 al 7 y del 0 al 10. Esta última escala permite una representación mental mejor de los datos que se obtengan.

El trabajo de campo

Por último, hay que diseñar el procedimiento que utilizaremos para llevar a cabo la encuesta y en qué condiciones se producirá. Si hemos de utilizar personas que nos ayuden a pasar la encuesta, conviene explicarles muy bien la finalidad de la encuesta y la solución a las posibles dudas que pueda plantear la persona encuestada a la hora de responder a las preguntas, teniendo mucho cuidado de no determinar las respuestas con sus comentarios. En este sentido, conviene prever el tipo de ayuda que darán los encuestadores e incluso el texto de las aclaraciones que previsiblemente deberán dar. Es muy importante que cuando las preguntas se contesten en grupo no se produzcan comentarios en voz alta que puedan influir en los demás. Los encuestadores han de vigilar también que no queden preguntas sin contestar o que se escojan respuestas incongruentes.

Es muy recomendable hacer una prueba piloto antes de hacer la encuesta definitiva. Consiste en pasar la encuesta a un grupo pequeño de personas que sea representativo, por sus características, del conjunto de la muestra. De esta manera podemos detectar errores no percibidos y situaciones imprevistas y solucionarlos: redacción incorrecta de las preguntas o de las respuestas, errores tipográficos, dificultades de comprensión o de interpretación de los textos, etc. A veces también es interesante conocer cuánto tiempo se tarda en contestar la encuesta.

El tratamiento de los datos

Antes de trasladar las encuestas a una base de datos hay que darles un vistazo para detectar aquellos errores que puedan ser resueltos por deducción (respuestas incongruentes, dobles respuestas, etc.). Una vez trasladadas las encuestas a la base de datos las deberemos volver a repasar porque podemos haber cometido errores de transcripción.

En cuanto a los programas estadísticos que permiten el tratamiento de los datos que

proporciona una encuesta, hay varios en el mercado. Los más conocidos son el Stat Graphics, el SPSS (Peña, 1993) y el SyStat (Peña y Vidal, 1993). Son programas caros que exigen una cierta dedicación para familiarizarse con su utilización. El programa más popular es el SPSS y existen en el mercado editorial algunos manuales para introducirse en su uso (Sánchez, 1990). A pesar del precio y las dificultades a la hora de utilizarlos son indispensables en estudios epidemiológicos y otras investigaciones de cierta complejidad.

Si los objetivos son mucho menos ambiciosos es muy probable que no estén justificados los esfuerzos necesarios para la comprensión del funcionamiento de estos programas y su uso (mucho menos aún la adquisición de los programas). Hay soluciones más baratas y sencillas como las aplicaciones estadísticas de algunos programas facilitados por el Departament d'Ensenyament (Bujosa y Manrique, 1991).

Un procedimiento mucho más rudimentario basado en el diseño de una buena plantilla donde recoger los datos, el uso de una calculadora científica y una buena dosis de paciencia son suficientes para traducir los datos en porcentajes y medias.

Un caso práctico

Qué queremos saber

Imaginemos que queremos hacer una encuesta entre los 960 alumnos de ESO de un instituto, el 45% de los cuales son chicos y el 55% restante son chicas. Nos hemos propuesto conocer sus hábitos de práctica físico-deportiva fuera del horario escolar (actividades practicadas y frecuencia). También queremos saber qué actividades físico-deportivas les gustaría practicar dentro de la asignatura de educación física y, de las actividades que ya forman parte de su currículum, cuáles son las que les gustan más y cuáles las que menos. Por último, aprovechamos la ocasión para hacer dos preguntas sobre el profesorado de educación física: una sobre su talante como docente y otra, más concreta, sobre las calificaciones que reciben en la asignatura de educación física. La encuesta la haríamos a final de curso.

La elección de la muestra

Como el número de alumnos nos parece excesivo para encuestarlos a todos, nos decidimos por encuestar sólo a una parte de todo el alumnado. Queremos tener en cuenta que el 55% del alumnado son chicas y el 45% restante son chicos, ya que estamos convencidos de que las respuestas a las preguntas que haremos pueden estar condicionadas por el sexo del alumnado.

Por otra parte, queremos cometer un error máximo del 5% con una fiabilidad del 95,5%. Para el cálculo exacto del tamaño de la muestra utilizamos la siguiente ecuación:

$$n_{\text{muestra}} = \frac{2^2 \cdot N_{\text{pob}} \cdot p \cdot q}{e^2 \cdot (N_{\text{pob}} - 1) + Z^2 \cdot p \cdot q}$$

En esta fórmula e = error (en tanto por uno), p y q (en tanto por uno), z = variable tipificada que tiene el valor 1,98 para una fiabilidad del 95% y de 2 para una fiabilidad del 95,5%. Si sustituimos las variables por sus valores correspondientes tenemos:

$$n_{\text{muestra}} = \frac{2^2 \cdot 960 \cdot 0,5 \cdot 0,5}{0,05^2 \cdot (960 - 1) + 2^2 \cdot 0,5 \cdot 0,5} = 285 \text{ personas}$$

Si el 45% de la población son chicos y el 55% chicas, la muestra deberá mantener esta proporción. En consecuencia, deberemos escoger aleatoriamente 128 chicos ($0,45 \times 285$) y 157 chicas ($0,55 \times 285$) utilizando cualquiera de los métodos explicados: el muestreo aleatorio simple o el sistemático. Si podemos elaborar dos listas, una para cada sexo, mejor. Si no podemos, primero escogemos los chicos (o las chicas), sin tener en cuenta los del otro sexo y después al revés.

Si utilizamos un muestreo aleatorio sistemático, escogeremos uno de cada 3 chicos y una de cada 3 chicas ($960/285 = 3,4$) empezando por el alumno que aparezca en tercer lugar en cada lista (empezando por la izquierda de la primera fila de la tabla de números aleatorios aparecen los números 5, 4, 4, 6, 3, 2, etc. El primer número igual o menor que 3 es el 3). Dentro de cada lista, escogeremos los chicos y las chicas que tengan los siguientes órdenes de lista: 3, 6, 9, 12, etc. Así hasta conseguir el número de 128 chicos y de 157 chicas.

El cuestionario

Aunque las características del instituto donde hacemos nuestra investigación son imaginarias, no tienen por qué serlo las preguntas de la encuesta. Incluso, las preguntas que propongo serán contestadas durante el mes de mayo de 1998 por el alumnado de ESO de mi instituto. Son las expresadas en la siguiente página (anexo).

El trabajo de campo

Cuando son alumnos los que han de contestar la encuesta, es muy importante garantizar que cada uno responda con total libertad, sin que pueda sentirse influido por las opiniones o comportamientos de los demás. En este sentido, se ha de transmitir a los encuestados la idea de que la encuesta es como un examen, que se ha de contestar sin hacer comentarios en voz alta y sin observar lo que hacen los compañeros vecinos.

El tratamiento de datos

Aparte de las puntuaciones medias y de los porcentajes, puede ser interesante buscar posibles diferencias entre las respuestas dadas por los chicos y las dadas por las chicas. También podemos estudiar el grado de coincidencia entre las actividades físico-deportivas que practican fuera de la escuela y las que desearían practicar en la escuela. Otro foco de atención es el que desean practicar aquellas personas que fuera de la escuela no hacen ningún tipo de actividad.



ANEXO.

Encuesta del área de educación física. Curso 97/98

Curso de ESO: 1 2 3 4

Grupo: A B C D E F H I J K L Q R S T Z

Edad: ____ años

Sexo: 1 Hombre 2 Mujer

1. En general, ¿cuáles de estos comportamientos crees que se dan entre el profesorado de educación física que has tenido durante este curso? (no te contradigas: no puedes escoger a la vez 1 y 2 o 5 y 6, por ejemplo).

- | | |
|--|---|
| 1 Es amable, simpático. | 9 Enseña mal. |
| 2 Es antipático. | 10 Es puntual. |
| 3 Se interesa por el alumnado que tiene dificultades. | 11 No es nada puntual. |
| 4 No se interesa por el alumnado que tiene dificultades. | 12 Acostumbra a dejarte en ridículo. |
| 5 Tiene sentido del humor. | 13 Mantiene el orden sin autoritarismo. |
| 6 No tiene sentido del humor. | 14 No es capaz de mantener el orden. |
| 7 Reconoce públicamente lo que haces bien o tu esfuerzo. | 15 Es justo a la hora de evaluar. |
| 8 Enseña bien. | 16 Es injusto a la hora de evaluar. |

2. En general, ¿cómo valorarías la calificación que has recibido en los créditos *comunes* y en los créditos *variables* del área de educación física? Marca con un círculo los números correspondientes.

CRÉDITOS	SUPERIORES A TUS CONOCIMIENTOS	COINCIDENTES CON TUS MÉRITOS	INFERIORES A TUS MÉRITOS
COMUNES	1	2	3
VARIABLES	4	5	6

3. ¿Cuántas veces realizas actividades físico-deportivas *fuera de la escuela* (como mínimo de 20 minutos de duración) que comporten un esfuerzo o dificultad respiratoria y que te hagan sudar? Sólo puedes escoger *una* respuesta.

- | | |
|--------------------------------|--|
| 1 Diariamente. | 4 Dos o tres veces al mes. |
| 2 Dos o tres veces por semana. | 5 Alguna vez al año o menos. |
| 3 Una vez por semana. | 6 No puedo por incapacidad o enfermedad. |

4. ¿Practicas *habitualmente* alguna actividad físico-deportiva *fuera del horario escolar*?

- 1 Sí 2 No (pasa a la pregunta 6)

5. Señala con un círculo el número que se corresponde con las actividades físico-deportivas que practicas fuera del horario escolar *una o más de una vez por semana*: (Puedes escoger más de una respuesta)

- | | | |
|-----------------------------------|--------------------------------|----------------------|
| 1 Bádminton | 9 Excursionismo | 17 Patinaje |
| 2 Baloncesto | 10 Fútbol | 18 Squash/raquetball |
| 3 Béisbol | 11 Fútbol Sala | 19 Tenis |
| 4 Bolos/Bowling | 12 Frontón (pelota mano, pala) | 20 Tenis de mesa |
| 5 Caminar (como ejercicio físico) | 13 Frontenis | 21 Voleibol |
| 6 Ciclismo (montar en bici) | 14 Gimnasia mantenimiento | 22 Balonmano |
| 7 Culturismo/musculación | 15 Jogging/Footing (correr) | 23 Otro: _____ |
| 8 Danza | 16 Natación | |

6. ¿Cuál o cuáles de estas actividades físico-deportivas te gustaría practicar en la escuela? Señala con un círculo el número (o los números) correspondientes. Puedes escoger hasta tres respuestas.

- | | | |
|-----------------------------------|--------------------------------|----------------------|
| 1 Bádminton | 9 Excursionismo | 17 Patinaje |
| 2 Baloncesto | 10 Fútbol | 18 Squash/raquetball |
| 3 Béisbol | 11 Fútbol Sala | 19 Tenis |
| 4 Bolos/Bowling | 12 Frontón (pelota mano, pala) | 20 Tenis de mesa |
| 5 Caminar (como ejercicio físico) | 13 Frontenis | 21 Voleibol |
| 6 Ciclismo (montar en bici) | 14 Gimnasia mantenimiento | 22 Balonmano |
| 7 Culturismo/musculación | 15 Jogging/Footing (correr) | 23 Otro: _____ |
| 8 Danza | 16 Natación | |

Bibliografía

- BONE, A.; CONTRERAS, A. (1985). "Necesidades e intereses de los alumnos de Bachillerato y de BUP". *Educación Abierta*, n.º 47, pp. 45-60.
- BUJOSA, J.M.; MANRIQUE, S. (1991). *Aplicacions estadístiques amb Framework*. Departament d'Ensenyament de la Generalitat de Catalunya. Programa d'Informàtica Educativa. Barcelona.
- CAÑELLAS, A.; ROVIRA, J. (1995). "Opinión: Los hábitos deportivos de la población adulta barcelonesa". *Apunts: Educación Física y Deportes*, n.º 42, pp. 75-79.
- COLTON, T. (1986). *Estadística en medicina*. Salvat Editores S.A. Barcelona.
- GARCÍA FERRANDO, M. (1991). *Los españoles y el deporte (1980-1990). Un análisis sociológico*. Ministerio de Educación y Ciencia. Consejo Superior de Deportes. Madrid.
- MARTÍNEZ DEL CASTILLO, J.; PUIG, N.; FRAILE, A.; BOIXEDA, A. (1991). *Estructura ocupacional del deporte en España. Encuesta en los sectores de entrenamiento, animación, docencia y dirección*. Ministerio de Educación y Ciencia. Consejo Superior de Deportes. Madrid.
- MARTÍNEZ DEL CASTILLO, J. (1993). "Los licenciados de los INEF: posiciones actuales y perspectivas de futuro". *Apunts: Educación Física y Deportes*, n.º 32, pp. 102-120.
- MARTÍNEZ-TUR, V.; TORDERA, N.; RAMOS, J. (1996). "Tipología de instalaciones deportivas en función de su carácter público o privado: diferencias en la gestión y el uso". *Apunts: Educación Física y Deportes*, n.º 43, pp. 91-102.
- MATEO, J.M. (1989). *Estadística en investigación social*. Editorial Paraninfo, S.A. Madrid.
- MOSER, C., KALTON, G. (1985). *Survey Methods in Social Investigation*. Gower. Hants.
- PEÑA, J. (1993). "SPSS for Windows: facilidad para el usuario". *PC World*, febrero, pp. 205-210.
- PEÑA, J.; VIDAL, M.C. (1993). "SyStat 5.0 para Windows: Programa estadístico". *PC World*, marzo, pp. 171-174.
- RODRÍGUEZ, J. (1991). *Métodos de muestreo. Cuadernos metodológicos*. Centro de Investigaciones Científicas. Madrid.
- SÁNCHEZ, J.J. (1990). *Introducción al análisis de datos con SPSS/PC+ versiones 2.0 y 3.0*. Alianza Editorial, S.A. Madrid.
- SERRA, A. (1995). *Enquesta sobre la pràctica d'activitats físicoesportives a Catalunya*. Generalitat de Catalunya. Direcció General de l'Esport. Barcelona.
- TINAJAS, A. (1996). "Educació Física: l'enquesta com a instrument per a l'elecció de continguts a l'ESO". *Guix*, n.º 224, pp. 67-71.