FF 17 21

Pedagogía experimental, fiabilidad de las medidas en pedagogía. Autor Piedad Granados García Tenorio. Fecha de emisión 24 de febrero.

La medición ocupa un lugar importante en la investigación experimental y trata de ofrecer los datos básicos necesarios para que el investigador pueda poner a prueba su hipótesis. Dos son al menos las condiciones esenciales que prestarán consistencia a tales datos. Por una parte, la fiabilidad, es decir, la confianza que podemos poner en los datos.

Y por otra, la validez o concordancia entre lo medido por los instrumentos que se utilizan y lo que en realidad se desea medir. De estos dos conceptos, fiabilidad y validez, vamos a hablar en este espacio radiofónico. FIABILIDAD Para introducirnos en el tema de la fiabilidad, pensemos en algo tan habitual en el mundo del estudiante como son los exámenes.

¿Podríamos asegurar que un examen produce resultados consecuentes y, por tanto, satisfactorios en el sentido de la finalidad que persigue? La realidad es que sí se puede asegurar siempre que el examen sea fiable, naturalmente. Pero, ¿con esto queremos decir que sea válido? No. La fiabilidad y la validez de un examen no son dos atributos recíprocos.

Un examen es válido cuando, de un modo demostrable, mide lo que se pretendía medir. Entonces, para que un examen sea válido, ¿ha de ser fiable? Sí, pero puede ser fiable y no por ello ha de ser válido. Imaginemos que estamos midiendo distancias con un metro defectuoso.

¿Tendría sentido registrar los resultados? Es evidente que no, ya que las medidas que hiciéramos carecerían de fiabilidad y, por tanto, sería imposible obtener con ellas una valoración válida. Supongamos que, en lugar de tener un metro inútil, tuviéramos una báscula perfecta de gran precisión. ¿Tendríamos con ella un instrumento fiable? Sí, pero con este instrumento no podríamos medir la distancia que nos proponíamos.

Efectivamente, aquí tenemos un caso en que existe fiabilidad, pero no hay validez. Son dos conceptos que hay que considerar por separado. Volviendo al ejemplo de los exámenes, hay que asegurarse, en primer lugar, de que produzcan resultados consecuentes.

Pero no por ello podremos asegurar, en principio, que los resultados vayan a ser correctos, en el sentido de que el examen permita valorar aquellas características en las que estamos interesados. Es muy importante. Un profesor debe tener en cuenta los factores que afectan a la fiabilidad, asegurarse de que las pruebas que emplea son lo más satisfactorias posibles a este respecto y asegurarse, asimismo, de que pueda valorarse la fiabilidad de esas pruebas.

¿Cuáles son los factores que pueden afectar a la fiabilidad de un examen? Pues, por ejemplo, el contenido y la misma estructura del examen. Si se formula una pregunta imprecisa o ambigua que deje en duda a los alumnos respecto a la clase de respuesta que se espera de ellos, se reducirá la exactitud de las evaluaciones. Pero entonces, por razones análogas, el examen en

conjunto no puede ser fiable si no muestra de forma adecuada las capacidades que se trata de comprobar.

En efecto, de todo ello podemos deducir que la fiabilidad aumenta en función del número de preguntas que contiene un examen. Si ponemos un examen con una sola pregunta, deja, en principio, demasiadas probabilidades a la suerte e incluso podría darse el caso de que esa pregunta fuera la única que no sabe un alumno o la única que sabe un alumno. Por tanto, cuanto mayor número de preguntas contenga un examen, tanto más fiables llegarán a ser los resultados, sobre todo si esas preguntas han sido escogidas de manera que nos permitan asegurar que se refieren a cada aspecto relevante de la materia de que se trata.

Sí, hay muchas preguntas que demuestran que las calificaciones otorgadas en base a una impresión subjetiva varían de un examinador a otro e incluso de una ocasión a otra para un mismo examinador. ¿Entonces podríamos pensar que las preguntas de ensayo disminuyen el grado de fiabilidad de un examen? Efectivamente, a menos que se establezca expresamente un plan detallado y objetivo de calificaciones en tales pruebas. ¿Existe algún sistema ideal de calificaciones? No podemos contestar de un modo concreto, pero en principio lo sería cualquier sistema que produzca resultados similares independientemente de quién sea el profesor que califique los exámenes.

¿Hay algún otro factor que influya en la fiabilidad de un examen? Sí, hay otra serie de factores que escapan a nuestro control. En la medición educativa tratamos de valorar el conocimiento y la capacidad que los alumnos han adquirido, basándonos en su ejercicio en una ocasión particular. Evidentemente, no tenemos garantías de que esta realización, aun cuando podamos valorarla con exactitud, haga plena justicia a las capacidades de sus respectivos ejecutantes.

El grado en que un alumno demuestra efectivamente el conocimiento o la capacidad que posee depende de una gran variedad de factores personales, como son el estado de ánimo, la salud, la motivación. Nada o casi nada es lo que puede hacer un profesor para evitar que la fiabilidad de un examen resulte afectada por tales factores. Y lo que es peor, nunca podrá controlar los resultados de sus acciones en este sentido.

Pero todo esto hace poco fiables las mediciones de los exámenes. En algunos casos sí, aunque a veces es posible salvar esta dificultad o al menos amortiguar sus efectos, repitiendo esas pruebas tras un corto intervalo y comprobar si los resultados son consecuentes. Las diferencias que surgen en estos casos no son imputables a defectos en el examen o errores de calificación, y muchas de ellas no serán sino el reflejo de cambios en las características que estamos valorando.

Porque si el intervalo de repetición de la prueba es breve, los alumnos no pueden haber tenido ocasión de incrementar notoriamente sus conocimientos, de tal modo que justifiquen esas diferencias. A este procedimiento, para determinar el coeficiente de fiabilidad, se le denomina repetitivo. El procedimiento repetitivo exige la preparación de dos test o exámenes que sean estrictamente paralelos y que midan los mismos aspectos del conocimiento.

En efecto, y como eso no es nada sencillo, se suelen buscar criterios de fiabilidad que puedan aplicarse a un solo test o examen. En esto consiste, por ejemplo, el procedimiento de mitades, constituido por un test dividido en dos partes que están en correlación. Lo primero que hay que hacer en este tipo de métodos es asegurarse de que las dos mitades sean comparables.

Además, hemos de tener en cuenta el hecho de que la correlación que se obtiene de esta comparación se basa en test que son más cortos que el original, y como antes decíamos, la fiabilidad de un test depende de su longitud. Cuanto más largo sea el test, mayor fiabilidad. Esto ocurrirá así a menos que empleemos una fórmula que compense el hecho de que estamos examinando un test que es exactamente la mitad de largo que el original.

La fórmula a aplicar sería Fiabilidad de un test es igual a 2R dividido por 1 más R, donde R es la correlación medida entre las dos mitades del test. Repetimos. Fiabilidad de un test es igual a 2R dividido por 1 más R, donde R es la correlación medida entre las dos mitades del test.

Si la fiabilidad de una parte de un examen no corresponde a la fiabilidad de todo el conjunto, tal vez un examen pueda ser fiable considerado todo entero, y no por ello podremos tratar como fiables a partes de ese examen. Podemos tener justificación para comparar las notas obtenidas por algunos alumnos en un examen o en un test, como un todo, pero hay que ser más cautos a la hora de centrar tales comparaciones sobre las relaciones de ciertas partes de ese test o examen, porque en este último caso estaríamos tratando con un nivel más bajo de fiabilidad. ¿Se puede valorar la fiabilidad de alguna otra manera? Existen otros métodos de valorar la fiabilidad de un examen, en general más elaborados, como el método de Kidder-Richardson, que basan la estimación sobre la clase de evidencia que brinda el análisis de preguntas.

Pero estos ya son otros métodos que merecen un capítulo aparte. Nuestra preocupación de cara a la fiabilidad es y debe ser hacer hincapié en la necesidad de asegurar que cualquier prueba o examen que se aplique sea todo lo fiable posible, aunque humildemente reconozcamos, e incluso establezcamos como base inicial, que la fiabilidad perfecta es inalcanzable.