

Las difíciles relaciones entre teoría, práctica y metodología en la psicología del siglo XXI. (II) Propuesta metodológica*

Juan Delgado Sánchez-Mateos
Universidad de Salamanca

La metodología propuesta para el conjunto de las indagaciones especializadas y las investigaciones científicas en psicología ha sido típicamente la parte de la metodología experimental conocida como “diseño de experimentos”, y definida como “el método científico”. En este trabajo se procede de modo inverso: revisando las características de la investigación experimental en disciplinas científicas maduras, y basándonos en desarrollos teóricos y argumentales de algunas modernas filosofías de la ciencia, se propone una metodología abierta y rigurosa que probablemente se pueda generalizar a los ámbitos de la investigación científica y de la indagación especializada en las ciencias del comportamiento, en sus vertientes científico-técnica y práctica. Pero, para usarla, tenemos que abandonar unos hábitos y aprender otros. Y proceder lentamente: nadie resuelve nada en un único intento.

Palabras clave: metodología psicológica, teoría, práctica, cuantificación, experimentación.

The difficult relationship between theory, practice and methodology in twenty-first century psychology.(II) A methodological proposal

Specialist research in psychology has typically applied the experimental methodology known as “experiment design”, and defined as “the scientific method”. In this study we proceed in the opposite direction: reviewing the charac-

*Este trabajo es la continuación de “Las difíciles relaciones entre teoría, práctica y metodología en la psicología del siglo XXI. (I). Sobre algunas disyuntivas”, publicado en el vol. 39 (3) de *Anuario de Psicología*.

Correspondencia: Juan Delgado Sánchez-Mateos. Departamento de Psicología Básica, Psicobiología y Metodología de las Ciencias del Comportamiento. Facultad de Psicología. Avda. de la Merced, 109-131. 37005 Salamanca.
Correo electrónico: jdelgado@usal.es

teristics of experimental research in the most developed disciplines, and basing ourselves on theoretical and argumental developments of certain modern philosophies of science, we propose an open and rigorous methodology which can in all probability be extended to other areas of scientific research and specialized investigation in behavioural sciences, in both its scientific-technical and practical aspects. But in order to use this methodology, we must give up certain habits and learn others, and above all we must proceed slowly but surely, because immediate solutions are rare.

Key words: Psychological methodology, theory, practice, quantification, experiments.

En un trabajo previo (Delgado, 2008) discutí las razones por las que parece que quepa defender la necesidad de un replanteamiento metodológico en las ciencias del comportamiento, basándome en la necesidad de hacerse cargo de las dificultades que plantean objetos de estudio peculiarmente definidos, en ámbitos no claramente deslindados ni categorizados y abordados metodológicamente desde los restos de un enfoque parcial, tardopositivista y sesgado del que se presume como único enfoque científico posible en estas ciencias.

Defendí que los objetos de estudio estaban peculiarmente definidos basándome en un análisis del tipo de teoría más frecuente en las ciencias del comportamiento, teoría y conceptualización de origen funcional, ontológicamente irrelevante, de restos positivistas y operacionalistas. Los conceptos implicados en ese modo de entender la teoría se dan por medidos mediante el expediente de ordenar sus propiedades en escalas ordinales que después se analizan como cuantitativas sin demostrar, ni siquiera intentarlo, que las dimensiones de los objetos medidos posean estructura cuantitativa. Lo cual debía llevarnos a la cuestión de la naturaleza y propiedades de esos “objetos”. Debía llevarnos, pero, a lo que se ve, no nos lleva.

Defendí también que había que deslindar las intenciones y los intereses implicados en la investigación, para lo cual me serví de la disyuntiva entre razón técnica y razón práctica, de amplias resonancias filosóficas en nuestra cultura, pero cuyas consecuencias se ignoran (o parecen ignorarse) en la indagación especializada en las ciencias del comportamiento. Así, no será lo mismo indagar, como ejemplo sensible, acerca de los mejores modos de apoyo y ayuda a las víctimas de catástrofes que investigar los componentes expresivos, o los fisiológicos, de las emociones de las personas.

Finalmente, defendí que no era desde una metodología única como podían abordarse los estudios de indagación especializada o de investigación científica, y menos desde una metodología mimética con un tipo de propuesta metodológica concreta: la derivada del concepto de diseño experimental fisheriano, con su jerga de contraste de hipótesis, del propio concepto de hipótesis, y del concepto epistemológico de conjetura-refutación (muchas veces, conjetura-verificación) implicado en esta propuesta metodológica. En otro lugar (Delgado, 2006a, 2006b) comencé este análisis que ahora prolongo.

1. Mirando a nuestro alrededor

Mi primera propuesta será que abandonemos algunos de los “tics” aprendidos acerca de cómo se lleva a cabo una investigación experimental en las ciencias del comportamiento, “tics” que se han generalizado al resto de modalidades metodológicas o de investigación en estas disciplinas, abandonando por el camino conceptos tales como “variable independiente”, “dependiente”, hipótesis, etc. Nos llevaremos la sorpresa de que también así podemos planear y diseñar un trabajo de indagación o de investigación.

Pero cabe preguntarse, ¿por qué hay que hacer tal cosa? ¿No será un intento de cambiar las palabras, el modo de referirnos a lo que hacemos, aunque todo siga igual finalmente? ¿Acaso no se sabe bien cómo es una investigación científica y qué la caracteriza? ¿Qué hay que cambiar, entonces?

La respuesta a estas preguntas no es sencilla. Para darla vamos, mirando fuera de nuestros dogmáticos y locales modos de pensar, a hacer dos cosas:

1. Mirar cómo se caracteriza y se ejecuta la investigación experimental en áreas desarrolladas, circunscribiéndonos específicamente al diseño de experimentos, y

2. Mirar cómo se conceptúa la metodología experimental en una parte de la filosofía de la ciencia actual.

Con dos intenciones así mismo:

1. Replantear nuestros locales y dogmáticos hábitos y nuestro discurso *sui generis* acerca de la investigación experimental circunscribiéndolo al sistema ciencia-técnica, y

2. Tratar de prolongar los nuevos hábitos y discurso a investigaciones científicas no experimentales así como a indagaciones especializadas en el sistema interacción-comunicación, en las áreas de la praxis.

Aunque mirar alrededor sea más incómodo que seguirnos mirando a nosotros mismos, a veces, por contraste, aprendemos algo.

¿Confirmar teorías o aprender de los errores?

En 1996, Deborah Mayo formula finalmente una propuesta epistemológica de consecuencias metodológicas directas en su libro *Error and the growth of experimental knowledge*. Su pensamiento toma forma alrededor de dos fuentes de influencia decisivas: por una parte, los trabajos de los nuevos experimentalistas (Goodison, Hacking), puede que también de los de Nancy Cartwright; por otra, su profundo conocimiento, compartido con Hacking (1965, 1975, 1990, 2001) de la probabilidad y la inferencia estadística, en concreto de la propuesta de Neyman y Pearson (Mayo, 1981, 1982, 1996, 2006), de la que hace una clara defensa.

Pero no son los estadísticos los aspectos que me interesa destacar en estas páginas, sino el cambio de visión que nos traen los nuevos experimentalistas, Ian Hacking (1983) como autor que juzgo más relevante, y la propia Deborah Mayo (1996). Basándose en cómo hacen los científicos su trabajo, no en cómo lo racionalizan los filósofos de la ciencia o los metodólogos hasta este mo-

mento, sino, insisto, en cómo lo hacen los científicos, intentan captar lo que caracteriza la obtención del conocimiento. De Popper se quedan con el concepto de “prueba severa” para mantener a raya el error. De Kuhn con su caracterización de la “ciencia normal”. Y se olvidan del resto. La ciencia ¿es esencialmente inductiva, como quieren los positivistas? ¿Es hipotético-deductiva como propone Popper? ¡Qué problema tan terrible! Tal vez merezca la pena cambiar esas preguntas por otras. Por ejemplo, ¿qué hacen los científicos cada día?

Esto tiene respuestas sencillas: no suelen estar comprobando teorías ni cambiando de paradigma cotidianamente. Más bien afinan sus instrumentos, controlan sus variaciones, se preocupan de medir con más precisión, toman datos, aplican modelos a esos datos, analizan cuidadosamente los resultados y los someten a crítica, propia y de sus pares, descubren errores, a veces los amplifican para saber cómo afectan, otras veces los eliminan... Puro trabajo.

Los cambios de paradigma y la comprobación de teorías, de los que tanto saben los aficionados a estas cosas, no suelen ser sus preocupaciones cotidianas. A la larga, tal vez alguna teoría cambie. Imre Lakatos nos demostró que en el fondo los teóricos se aferran al núcleo (*hardcore*, término de peculiares resonancias) de sus teorías, y cambian detalles *ad hoc* para no abandonarlas nunca del todo.

Hacking, sobre todo, y Mayo nos han permitido entender que no existe tal cosa como una metodología científica única, que no siempre se pretende comprobar hipótesis, que casi nunca se pretende rechazar o afirmar una teoría, que se puede trabajar científicamente sin unas y sin otras, que lo que importa es otra cosa: lo que el científico hace para mantener bajo control el proceso que está estudiando. Mayo cita específicamente a Laudan y su enfoque naturalista, ya presente en Kuhn, y mantiene que lo que hace el científico es lo que constituye un saber como más o menos fundado.

Es interesante la comparación que presentan, atribuida a Kuhn, entre astrónomos y astrólogos: la diferencia entre la astrología y la astronomía radica en que los astrónomos están en disposición de aprender de los fallos de sus predicciones, de sus errores, y los astrólogos no. Los astrónomos pueden refinar sus instrumentos, probar en busca de perturbaciones posibles, postular la existencia de planetas no detectados, etc., y llevar a cabo un trabajo detallado y cuidadoso para ver si los cambios que introduzcan eliminan el problema originado por una predicción no cumplida. Los astrólogos, en cambio, no tienen el recurso de aprender de los errores de la misma manera. Esto, y no si la astronomía es una ciencia y la astrología deja de serlo, es lo que hace que con la astronomía se obtenga conocimiento en el que poner confianza (y fondos públicos o privados para financiar investigaciones).

Actividades y actitud del investigador

Una de las propuestas de Mayo (1996) que juzgo más útil desde un punto de vista metodológico es su tratamiento de la idea de que aprendemos de los errores, y de cómo lo hacemos. En la ciencia normal, las anomalías proporcio-

nan oportunidades para aprender del error. El espíritu de la ciencia no es evitar la anomalía y el error, sino la capacidad de aprender de ellos. Así propone un conjunto de actividades para el investigador:

1. *Comprobación tras la prueba.* Se trata de someter a crítica severa nuestros resultados, revisar errores, predicciones no confirmadas, sorpresas,... Porque formarán la base de ulteriores intentos de control o amplificación del error.

2. *Planificación antes de la prueba.* Obviamente, conocidos los resultados previos, revisado el conocimiento conceptual y experimental previo, sólo la cuidadosa planificación permite mantener a raya el error. Planificación que integra razonamientos alrededor de cómo manejar el error.

3. *Repertorio de errores.* En la planificación, una exhaustiva lista de los factores que pueden hacer variar los resultados, y el modo en que se espera que lo hagan, es imprescindible

4. *Efecto de las equivocaciones.* Pero no siempre puede preverse todo, ni siempre nuestras previsiones son acertadas. Hay un espacio para la sorpresa. Hay un espacio para aprender. Podemos prever qué efecto tendrá el error previsto, pero ¿qué efecto han tenido los errores imprevistos? ¿Cuáles han sido?

5. *Simulación de los errores.* No sólo hay que evitar los errores. Hay que aprender de ellos. ¿Qué variaciones se obtendrían en los resultados en caso de que se variase tal o cual factor imprevisto anteriormente, o no convenientemente controlado, o conceptualizado?

6. *Amplificación y lectura de los patrones de error.* ¿Por qué no introducirlos en una situación de investigación y comprobar su efecto? Tarea, por otra parte, típica en la experimentación, convirtiendo en bloques o en factores las fuentes de variación pertinentes.

7. *Robustez.* ¿Qué información se obtiene de un proceso en el que la variación de factores y errores no provoca cambios en el proceso estudiado? ¿En qué medida éste varía o permanece inalterado? En esa medida, el fenómeno observado y estudiado será confiable, podremos suponer que no es un artificio experimental

8. *Exploración severa del error.* En resumen, pruebas severas. Si, tras darles todas las oportunidades de variación a los factores y los errores, los resultados se mantienen ¿no será plausible decir que se mantienen? ¿Cómo podrá afirmarse que es necesaria la falsación para que el conocimiento avance?

Si hay una teoría o una hipótesis (tanto como si no las hay), y tras una exploración severa del error aprendemos algo ¿no será razonable poner confianza en esa teoría, en esa hipótesis, o meramente en los hallazgos obtenidos? Tal vez éstos provoquen un cambio teórico o paradigmático, lo que sería en todo caso una consecuencia externa a la investigación. Tal vez no lo provoquen. En cualquier caso, se habrá aprendido algo.

Una demostración en clase

Benton J. Underwood (1976), en el capítulo 2 de su libro clásico *Psicología experimental* incluía un experimento con el que ejemplificaba la necesidad de

algunas técnicas de control. Propone una experiencia sencilla que se puede usar como demostración en clase y que puede servir para comunicar a los estudiantes algunas de las características de la investigación experimental. Se trata de un experimento de percepción del tiempo, de la duración.

El alumno ha sido adiestrado a pensar en los siguientes términos: hay una teoría que afirma que las personas estimamos la duración de los acontecimientos con cierta imprecisión, que será tanto mayor cuanto más largo sea el intervalo temporal a estimar. A esta conjetura la llama “hipótesis teórica”, o “hipótesis general”. Por vía deductiva (y a lo que se ve, por arte de magia) se concreta en una “hipótesis empírica”, a la que se llega “operacionalizando” los “constructos” implicados: “duración temporal” e “imprecisión”. Se llega así a la siguiente “hipótesis experimental”: *Si aumentamos el tiempo entre dos sonidos que se “presentan” a un conjunto de sujetos, entonces éstos aumentarán sus estimaciones en segundos (sobre-estimaciones) del tiempo transcurrido entre los sonidos.* Se realiza un experimento para probar esta hipótesis experimental y proporcionar así apoyo empírico a la teoría de la que ésta se deduce.

Además, sin profundizar en ello, se traduce la hipótesis empírica en una hipótesis estadística. Ésta se opone a otra hipótesis estadística llamada “nula” formulada como conjetura a rechazar. Si se rechaza la nula, se afirma la hipótesis estadística que traduce la empírica o experimental. Así, la empírica o experimental queda verificada, con lo que la hipótesis teórica recibe apoyo empírico. Con un pequeño saltito, afirmamos que la teoría queda confirmada.

El estudiante que así razone recibirá la aprobación de su profesor de metodología experimental, y se reafirmará en la admiración de tan elegante esquema de investigación. Hasta que lea más, mire fuera, y vea que, amén de la falacia implicada en el “rechazo de la nula”, los manuales de diseño de experimentos de otras disciplinas utilizan una jerga diferente. Bueno. Será eso: otra jerga para hablar de lo mismo.

¿Será sólo otra jerga? Cambiemos la pregunta. ¿Realmente la investigación científica se plantea de tan artificial modo? ¿Se formulan hipótesis “si... , entonces...”? ¿Hipótesis teóricas, empíricas, alternativas y nulas? ¿De dónde salen una y otras? En otros términos: todo eso ¿cómo se hace?

Veamos otra forma de hacer.

Construcción del diagrama o gráfica de causa y efecto

Kaoru Ishikawa propuso un procedimiento gráfico para organizar los factores que pueden influir en el resultado de un proceso. El diagrama resultante recibe el nombre de “diagrama de causa y efecto”, o “de Ishikawa” o, por aquello que recuerda su forma, “diagrama de raspa de pescado (*fishbone*)”. En la figura 1 se ha confeccionado un ejemplo de tal diagrama sobre el experimento de estimación de la duración que Underwood (1976) incluía como experimento práctico en su capítulo 2.

Les debemos decir a nuestros estudiantes para qué y por qué es bueno usar esta gráfica. Pero lo dejamos aquí para más adelante, y continuamos ahora con nuestro ejemplo. Para confeccionar la gráfica de Ishikawa para el experimento de

Underwood, nos preguntamos, en primer lugar, qué puede afectar al juicio de una persona que estima el tiempo que ha pasado entre dos sucesos. Nada de hipótesis generales ni empíricas, ni historias, ni racionalizaciones. Preguntas sencillas.

Para responder, hacemos una lista de los factores que, desde nuestro conocimiento como expertos, son relevantes (o sea, tras una exhaustiva revisión y un no menos exhaustivo estudio de la literatura científica sobre estimación de la duración, consulta a otros estudiosos o expertos, etc.). Agrupamos convenientemente los factores que hemos decidido como relevantes en categorías, y los representamos en el diagrama. Así, éste se constituye en un instrumento gráfico que nos permite racionalizar la primera parte del proceso de investigación.

Estamos con ello construyendo un repertorio de errores posibles (Mayo, 1996), del que se deriva una lista de las cosas que hay que hacer para controlarlos, para hacer mínimas las probabilidades de que lo que queremos estudiar resulte encubierto, confundido o distorsionado por los múltiples factores que causalmente puedan encubrirlo, confundirlo o distorsionarlo.

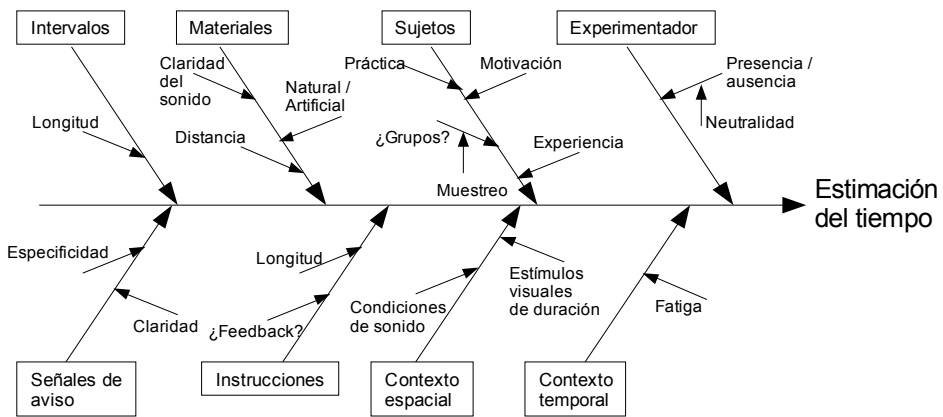


Figura 1. Ejemplo de gráfica causa-efecto para un experimento sobre estimación de la duración de un intervalo temporal. En ella, como ejemplo, aparece una lista, en este caso parcial, de los factores que se prevé que afecten al proceso o fenómeno bajo estudio.

Así, para la estimación de la duración, será importante el tipo de señal de inicio y de fin de intervalo a estimar, si es específica, en el sentido de si es una señal puntual, verbal etc., la duración de ese intervalo, el número de ensayos, la exposición previa del sujeto a este tipo de tareas, el lugar en que va a llevarse a cabo el estudio, si hay o no señales visuales de duración, la fatiga derivada del momento del día en que se realice el estudio, ...

Hacer explícitos estos factores y agruparlos es una manera directa de racionalizar el proceso de planeamiento de un estudio, que nos obliga de modo inmediato a hacer explícitas las técnicas, procedimientos o estrategias que utilizaremos

para evitar la confusión y el error, para que no afecte lo que no debe afectar al proceso estudiado, para que la variabilidad de ese proceso sea mínima, para mantener constante lo que deba mantenerse constante, o para variar, en su caso, aquello que queremos ver si se relaciona con la variación del fenómeno que pretendemos estudiar. De esta lista organizada (diagrama de Ishikawa) se deriva la necesidad de hacer explícito lo que va a constituir nuestra estrategia de investigación. Y abrirá, además, la posibilidad de que alguien replique nuestro estudio ampliando o cambiando nuestra lista de factores, organizándolos de otra forma, o manifestando su acuerdo o desacuerdo con cada uno de ellos.

Construcción de tablas de decisiones sobre los factores

Posteriormente se hace explícito el conjunto de acciones que, para cada factor identificado, van a ponerse de hecho en funcionamiento. Coleman y Montgomery (1993) propusieron una excelente guía. A su trabajo original remitimos, aunque valga la pena señalar aquí que existe un formato plausible, y en cierta medida estándar, de especificar qué haremos con cada uno de los factores listados anteriormente e incluidos en la gráfica de Ishikawa, formato que se ejemplifica en la tabla 1. En ella se muestra la estructura de valores a especificar para los factores (variables de control) que se decida variar.

TABLA 1. ESPECIFICACIÓN DE LAS CARACTERÍSTICAS CLAVE DE LOS FACTORES (VARIABLES DE CONTROL) DE UN ESTUDIO, Y SUS EFECTOS. SE DEBEN CORRESPONDER LAS VARIABLES CON LOS FACTORES QUE SE PROPONEN EN LA GRÁFICA CAUSA-EFECTO.

<i>Variables de control</i>	<i>Unidades</i>	<i>Niveles y rangos normales</i>	<i>Precisión y error</i>	<i>Valores propuestos</i>	<i>Efectos anticipados</i>
<i>Duración de los intervalos</i>	Segundos	1 – 50	± 0,1	8, 12, 20, 31	Sobreestimación mayor para duraciones mayores
<i>Distancia del sujeto</i>	Metros	1 – 10	± 0,5	2, 4, 8	Aumento de la sobreestimación para distancias mayores
<i>Etc.</i>					

Tablas análogas se construyen para aquellos que se decida mantener constantes o para los factores de ruido (*nuisance variables*) que, inevitables, haya que aleatorizar, bloquear, etc. No se olvide que la especificación de los valores que se espera que tomen, o de las consecuencias de ello, o de la falta de control sobre ello, sustituye con ventaja a las imprecisas, parciales, y superfluas hipótesis de muchos de nuestros escolásticos trabajos de investigación. Y, además, permite anticipar el efecto que esperamos que produzcan los errores, para así, amplificarlos, o reducirlos, o eliminarlos o estudiarlos, según decidamos en nuestra estrategia de investigación (Mayo, 1996).

Teniendo claro qué ocurre, ocurrirá o esperamos que ocurra con cada uno de los factores determinados, y qué decidimos hacer para controlar cada uno de ellos, haciéndonos conscientes así del repertorio de errores que somos capaces de controlar, estamos en disposición de especificar los pasos concretos que daremos para que se desarrolle el proceso global de investigación. En este punto, la gráfica IDEFO (www.idef.com) será de una inestimable ayuda.

Construcción de la gráfica IDEFO

Se trata de una gráfica en la que se muestra la subdivisión del experimento, la observación o la indagación, en tanto que proceso temporalmente ordenado, en un conjunto de tareas específicas a llevar a cabo, de los materiales necesarios para cada tarea, de los recursos imprescindibles, y del modo en que se va a ejecutar, de la información necesaria para finalizar cada paso del proceso, lo que hará posible iniciar el (o los) siguiente(s) paso(s) o tarea(a). Las figuras 2 y 3 representan respectivamente esquemas para una tarea aislada, y para un conjunto de tareas componentes de un proceso general.

Algunas tareas proporcionan resultados que se manifiestan externamente (damos instrucciones al sujeto, y le preguntamos si ha entendido la tarea, por lo tanto registramos su respuesta), otras simplemente posibilitan los materiales de entrada para una nueva fase del proceso general.

En la medida en que nos sirve obligándonos a hacer explícitas cada una de las decisiones que en el desarrollo de la investigación habremos de tomar, y hacerlo antes de ejecutar la investigación, la gráfica IDEFO ayuda (y obliga) a racionalizar el proceso de investigación, a que nuestras acciones como investigadores se correspondan con nuestros objetivos, ya obligatoriamente puestos de manifiesto en las tablas de control de los factores, a que la investigación se estructure como un proceso racional.

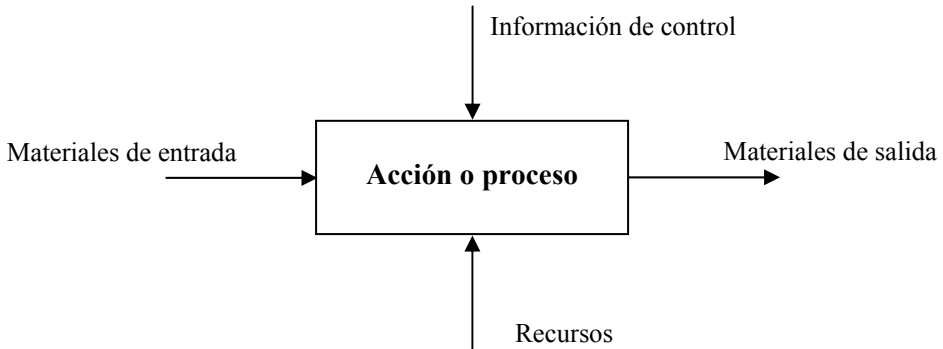


Figura 2. Elementos básicos de una gráfica IDEFO. Con un conjunto de acciones o procesos conectados se construye un modelo de proceso general, por ejemplo un proceso de investigación.

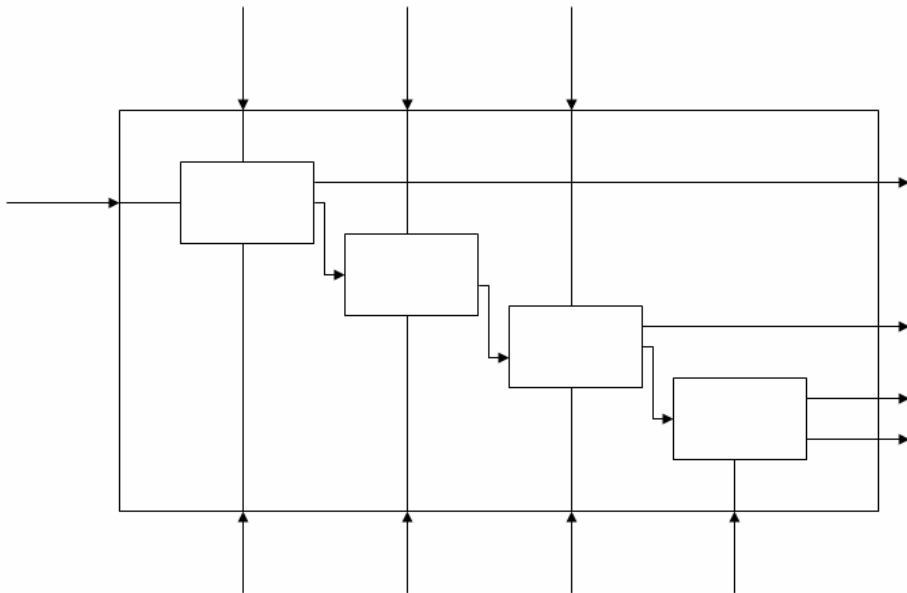


Figura 3. Ejemplo del esquema general de una gráfica IDEFO compuesta por un conjunto ordenado de acciones o procesos dependientes.

Así, para abandonar definitivamente nuestro ejemplo con el experimento de percepción de la duración, habrá que indicar en alguna parte del proceso modelado con la gráfica IDEFO mediante qué objetos físicos se emitirán los sonidos espaciados temporalmente, qué formato tendrán los objetos físicos que hayan de servir para registrar las respuestas de los participantes, mediante qué objetos físicos controlaremos la precisión de la medición de los intervalos entre los sonidos, etc. Y, en cada caso, qué resultados se obtendrán: sonidos, instrucciones, hojas de respuesta, por ejemplo. En otros términos, con la gráfica IDEFO nos obligamos a especificar qué vamos a hacer, cómo nos vamos a comportar como investigadores, si vamos a estar sentados, si vamos a usar altavoces, ordenador, qué software, en su caso, o en qué disposición espacial situaremos a los participantes ante la emisión de los sonidos.

Es obvio que en cada investigación surge lo inesperado. Y que hay áreas de indagación, o momentos en cada estudio, donde lo inesperado ha de surgir casi obligatoriamente. Esto no es óbice para la utilización de estrategias estructuradas de planeamiento de la investigación. Más bien todo lo contrario. Si lo inesperado nos sorprende en una indagación mal planeada, parece peor, ¿no?

Además, admitir la aparición de lo inesperado nos debe poner sobre aviso de que la investigación nunca es un proceso acabado, sino uno iterativo, de aproximaciones sucesivas al experimento, o a la investigación mejor posible dentro de nuestras limitaciones de tiempo, de recursos, de conocimientos, o de creatividad. Seguir manteniendo que se realiza un experimento, y con él se

prueba de una vez una hipótesis (incluso se cambia de paradigma y todo), es sencillamente una simplificación, calificable siempre como inconveniente, muchas veces como simplista y alguna vez como cínica. Todo aquél que ha investigado conoce bien lo provisional de cada decisión del investigador, y cómo la testaruda realidad se impone ante nuestras expectativas.

Entonces...

A la luz de un análisis de este proceso, que podría desarrollarse exhaustivamente, y será bueno hacerlo en clase para que los estudiantes sepan exactamente cómo diantres se ejecuta de hecho una investigación, parece evidente que la escolástica metodológica de la formulación de hipótesis, de la clasificación de variables en independientes, dependientes, extrañas y mediopensionistas es poco más que un ejercicio de reconstrucción *a posteriori* de lo que un proceso de investigación representa: un conjunto de tareas dirigidas específicamente a mantener a raya el error en la indagación y la investigación científica.

Y, lo que es más, éste esbozado aquí es un esquema válido para todos los tipos de indagación e investigación, cualitativa o cuantitativa, experimental u observacional, de sucesos o de procesos. Será más fácilmente realizable en unos ámbitos que en otros, será más o menos fácil especificar (e incluso no será posible alguna vez) el conjunto de valores que se espera o se propone para algún factor. Pero el proceso de racionalización implicado es imprescindible si queremos que la investigación sea algo más que un pretexto para que engorde nuestro currículo.

Es curioso, los diagramas, guiones y gráficas a los que acabo de aludir son práctica común en el diseño de investigaciones en la industria, en procesos de control de calidad, y en la ingeniería. Y los psicólogos (que, ya se sabe, leemos metodología aplicada a la psicología) sin enterarnos....

2. La última dicotomía simplista

Uno de los puntos más débiles de nuestra idea de metodología la constituye el hecho de que situamos el pensamiento metodológico, o bien meramente en el mundo de los datos, o, como mucho, en la vinculación que con una flecha hacia arriba (inducción) o una hacia abajo (deducción) establecemos entre dos mundos: el de la teoría y el de los datos. Como si la teoría fuese ese universo en el que Habermas decía que “se enajena el mensajero ante el sacro acontecer”. Como si los datos fuesen el lugar de los obreros, de los analistas, o, en un absoluto disparate, de los positivistas. Como si no pudiese hablarse de los mundos intermedios, que hay varios. Las cosas son más complejas.

Mayo (1996) recoge y organiza una idea de Patrick Suppes (1962) para quien aquellos esquemas simplistas de la relación entre teoría y datos habrían de sustituirse por esquemas que pusieran de manifiesto que entre ambos existe y se articula un conjunto de modelos intermedios que componen una jerarquía organizada.

Hay que señalar que para esta autora las teorías y las hipótesis tienen mucha mayor relevancia que para los nuevos experimentalistas, en concreto para Hacking, y, modestamente, para este autor. Pienso, más bien, que el juego de la relación entre teoría y datos es el de lo que decimos y lo que hacemos. Si lo que decimos no tiene un correlato en nuestros actos como investigadores, en nuestro comportamiento como indagadores, en nuestra participación activa en la generación del conocimiento, ya podemos construir edificios (o chozas) conceptuales elegantes y plausibles, enunciados verbales brillantes o redes de significados complejas y aplaudidas por el público. Pero, sobre esa idea, tal vez haya tiempo de avanzar en otra ocasión. Ahora estábamos con Mayo, y su manejo cuasi-ortodoxo de los conceptos de teoría e hipótesis. Incluso desde esa cuasi-ortodoxia se puede repensar cómo funciona realmente la jerarquía de modelos implicada en la relación entre los datos y las teorías, nuestra última dicotomía de esta serie.

Hay que indicar, obligatoriamente, que la propuesta de Mayo no puede resumirse en estas pocas páginas. Prácticamente todo su libro (Mayo, 1996) consiste en este desarrollo. Por eso sencillamente bosquejo su punto de vista, y remito a su excelente y sugerente trabajo¹. En breve y esquemáticamente, esa jerarquía, para Mayo (1996), estaría compuesta por:

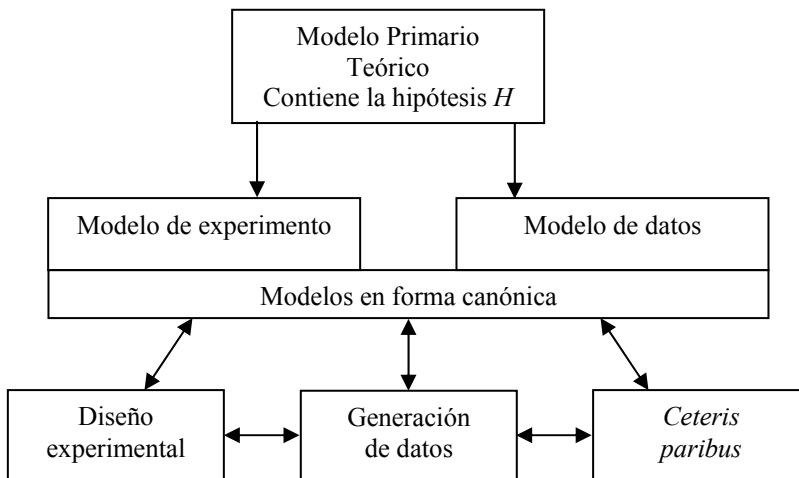


Figura 4. Esquema de la jerarquía de modelos de Mayo (1996). Respecto de la conveniencia y sentido de formular los modelos de experimento y datos en forma “canónica”, consúltese el trabajo original.

¹ Preveo poner de manifiesto mis modestos desacuerdos con Deborah Mayo en otra ocasión. Pero no puedo dejar de señalar ahora que, junto con el acuerdo de fondo sobre su trabajo, y específicamente en lo que hace a estos modelos jerárquicos en distintos niveles, tengo muy serias dudas de que la teoría de la inferencia de Neyman y Pearson sea tan plausible como modelo científico. Y esto tiene como consecuencia desacuerdos ulteriores. A pesar de todo, y aunque deba señalar estas tímidas, y no irrelevantes, discrepancias, considero que el trabajo de Mayo es una seria plataforma desde la que seguir trabajando.

Modelo primario

El nivel superior de la jerarquía es el de la teoría, donde se encuentra, en su caso, la hipótesis principal (H) de una investigación. Existen diferentes modelos de teoría. Cualquiera de ellos se consideraría el modelo primario.

Modelo de experimento

Por debajo de ese nivel se encuentra el modelo experimental. Se trata aún de un modelo ideal, del experimento mental, por lo tanto no realizado, supuesto, que se constituiría en un análogo de los aspectos más sobresalientes del modelo primario: qué se espera que ocurra en un experimento adecuado en el caso de que H sea verdadera. Este modelo experimental contiene el tipo de experimento que se podría realizar, el número de ensayos que se juzga adecuado, qué registros habrían de llevarse a cabo y de qué características se tomarían registros. Contiene los aspectos clave del experimento. Pero aún es un modelo ideal, no directamente ejecutable. Contiene, del mismo modo, una traducción de H a la que Mayo llama Hipótesis experimental (H_{exp}), que puede o no ser de naturaleza probabilística, estadística. Especifica qué técnicas vinculan los datos experimentales a las preguntas de investigación. A veces esos vínculos se modelan estadísticamente, a veces no. Lo que es independiente de la naturaleza (estadística o no) de H_{exp} .

Modelo de datos

El tercer tipo de modelo de la jerarquía es el modelo de datos. Los datos nunca son simples, nunca están ahí: hay que generarlos, trabajarlos. Una media o una mediana es un dato modelado, los datos brutos (“crudos” los llaman en México, con muy buen criterio) son la base de los datos modelados. Los datos modelados han de ser paralelos a los aspectos relevantes de H : deben dar información sobre H . Éstos son los que valen como evidencia, porque son más precisos que los datos individuales, ya que éstos “están crudos”.

Del modelo de datos se derivan nuestras posibilidades de utilizar éstos en una argumentación libre de error. Incluye los supuestos que han de cumplirse para ser aptos tanto para la evaluación de H o H_{exp} como para su tratamiento posterior. Y esos supuestos deben cumplirse, y ¡ha de evaluarse si se cumplen y en qué medida! Y se debe chequear que los instrumentos, en su caso, proporcionan datos aptos.

Ha de ser posible remodelar los datos en el caso de que no cumplan los supuestos: por ejemplo, a veces se usan transformaciones matemáticas (logarítmicas para tiempos de reacción, por ejemplo clásico) para conseguirlo. La relación de estos modelos de datos con los diseños experimentales a realizar y con los análisis de datos a llevar a cabo es muy íntima.

Nivel inferior

En el nivel inferior, llegamos a los modelos de diseño experimental, de generación de datos y de condiciones *ceteris paribus*.

Modelo de diseño experimental

Contiene los detalles concretos de asignación de sujetos a grupos, de asignación de tratamientos, de si se va a realizar un experimento doble-ciego, con placebo, etc. Por supuesto, de acuerdo con el modelo experimental y el de datos de los que depende jerárquicamente. Para ciertos modelos de datos, ciertos modelos de diseño. Es evidente que no se usan los mismos diseños para, por ejemplo, datos modelados sobre frecuencias en variables cualitativas que sobre datos modelados sobre medición continua y con relaciones presumiblemente no lineales.

Como es evidente que el modelo experimental condiciona el modelo de diseño experimental: el experimento mental permite y dirige el diseño del experimento.

Modelo de generación de datos

Incluye la especificación de los equipos, instrumentos y sus teorías asociadas. Como teoría asociada a la generación de datos, recordemos, como ejemplo concreto, que un acoplador de conductancia se basa en una teoría muy compleja que incluye, entre otros, un modelo eléctrico de la piel.

En este nivel se especifican y, en su caso, se realizan los ajustes necesarios de los instrumentos (ya sean aparatos o cuestionarios). La evaluación de los errores experimentales posibles es parte de este proceso.

Modelo ceteris paribus

Las condiciones *ceteris paribus*, “a igualdad de otras cosas”, son todas aquellas que no se incorporan a los modelos de datos, y que hay que mantener constantes, controlar o ignorar, en su caso.

También debe ser posible registrar las condiciones *ceteris paribus* más importantes para así poder ejercer sobre ellas un control a posteriori: por ejemplo, un control estadístico, como el análisis de covarianza, u otros tipos de control manipulativo o argumental.

¿Dicotomía teoría/datos?

La articulación de estos modelos hace evidente que entre teoría (vagamente, difusamente entendida como una construcción verbal) y datos (vaga, impropriadamente entendidos como la única fuente de inferencia) existen otros niveles. Y que esos niveles, aquí referidos a los experimentos, aunque puedan extender-

se a otros modos de investigación, son susceptibles de un análisis netamente, si no exclusivamente, metodológico.

Es en esos niveles en los se produce la interacción entre especialistas en la investigación: entre expertos en generación de datos, expertos en instrumentación, expertos en control estadístico, expertos en la generación de experimentos ideales a partir de modelos conceptuales, etc.

Pretender que la teoría es el mundo del teórico, que puede ahí hacer de su capa un sayo, y que el metodólogo es el que convierte las hipótesis teóricas en “operaciones”, o, peor aún, el que analiza los datos que nadie sabe de qué curso de razonamiento provienen, es no comprender los procesos correctos de investigación científica, con sus complejidades, sus complicaciones y su extrema dificultad.

La ciencia no es fácil ni es trivial. Es compleja, difícil e importante. Cuesta trabajo: exige rigor, cuidado, control, interacción de expertos, de especialistas, estudio, reflexión, creatividad, tesón, apertura y atención. Pero exige, sobre todo, pensamiento. Y una variedad de pensamiento muy útil en la ciencia (y en la vida): el pensamiento crítico.

3. Finalmente

Por entre las rendijas de esquemas y texto, espero que se haya podido entrever la vinculación del proceso de investigación propuesto en el apartado 1 de este trabajo y la jerarquía de modelos que Mayo articula y propone, y que he recogido en el apartado 2. Estimo que es una vinculación sobre la que hay que seguir trabajando, y, si se me permite una expresión coloquial, “en ello estamos”. Hay que prolongar, adaptar y concretar estos planteamientos desde el mundo de la investigación experimental al de la indagación sistemática especializada en las áreas prácticas. Pero estoy convencido de que es a partir de este tipo de propuestas como avanzaremos en la dirección que nos lleve a una investigación más rigurosa en psicología. Y una investigación más rigurosa será, probablemente, más difícil y más lenta. Pero eso debe preocuparnos menos que la posibilidad de que cada investigación nos haga aprender algo, a nosotros y a la sociedad a la que pertenecemos y que nos paga para que estudiemos, aprendamos y lo difundamos. No para que divulguemos muchas cosas que significan poco o nada. Aunque haciendo esto último ascendamos en la valoración de nuestros pares (los que hacen lo mismo), o de nuestros censores, de los que en estos tiempos dependemos cada vez más para que se nos considere alguien. Aunque seamos alguien, en cualquier caso, lo que siempre será mejor que no ser nadie y brillar.

No es discutiendo acerca de las dicotomías que más o menos ficticiamente nos separan como vamos a avanzar en disciplinas como la psicología. Creo que avanzaremos en la medida en que seamos capaces de hacer nuestro trabajo dejando bien claro, en primer lugar, cuáles son los intereses que nos guían: prácticos o técnicos. De ello se derivan los supuestos en que basamos nuestra indagación o nuestra investigación, y sus objetivos. Desde ahí, el análisis conceptual

imprescindible es diferente según nuestros intereses. Con los intereses, supuestos y conceptos bien definidos, que recojamos datos cualitativos o cuantitativos, experimentales, de observación natural, *ex post facto* o etnográficos, es menos relevante que el proceso de racionalización implicado. Y la recogida de datos es, en sí misma y en cualquier proceso de indagación o investigación, una de las tareas más complejas, ingratas y expuestas a la posibilidad de error. Desde el momento en que proponemos, discutimos y decidimos los instrumentos que nos permitirán obtener datos, estamos sometiéndonos a las restricciones que los modelos de instrumentación (vídeo, cuestionarios, taquistoscopios o equipos de potenciales evocados) imponen sobre lo que vamos a considerar dato.

Llevar a término un buen diseño de investigación es una tarea compleja para la que, afortunadamente, existen ayudas (gráficas, de planificación, etc.). Pero no existen recetas. No es defendible que haya una única forma. No puede continuar más tiempo un discurso académico incorrecto, y mucho menos dogmático e incorrecto.

Podemos enseñar a nuestros estudiantes cómo se ejecuta una investigación científica o una indagación especializada, paso a paso, punto por punto, siguiendo un plan, con ayudas gráficas, con modelos de datos, de experimento, de observación o cualquier otro modelo que sea necesario o que se proponga y argumente como tal, siendo cuidadosos en la elección del conocimiento disponible relevante, articulando los factores que presumiblemente influirán sobre lo estudiado, teniendo en cuenta cómo los vamos a mantener bajo control cuando ejecutemos la investigación, rectificando desde nuestros errores, como un proceso iterativo... Es decir, podemos intentar que aprendan cómo se investiga. Y que aprendan que el proceso de investigación será necesariamente diferente en cada investigación concreta, en cada ámbito de indagación, en psicología social o en la psicología de la motivación, o en la infantil, o del pensamiento, o en ergonomía.

Podemos hacerlo desde un desarrollo epistemológico moderno, naturalista o no (es una opción, obviamente), realista en los conceptos, no anclado a una teoría de la ciencia demasiado preocupada por los grandes sistemas (teoría, paradigma, programa de investigación) y desencarnada de la manera en que se obtiene conocimiento cada día, sino a una epistemología comprometida con las relaciones entre el discurso y la acción, con el vínculo indisoluble entre lo que hacemos y cómo comunicamos lo que hemos aprendido.

También podemos seguir por el trillado camino de la(s) variable(s) independiente(s) y la(s) variable(s) dependiente(s), de las extrañas controladas y la técnicas de control, de la formulación de hipótesis mediante deducción desde una teoría con la que crear una revolución científica, o no, y cambiar de paradigma, o seguir en el actual, penosamente; el del análisis estadístico de los datos, vengan de donde vengan y sean como sean, mediante SPSS, llamando al metodólogo de guardia; ... y pretender que éste es un esquema válido para cualquier investigación.

Me acabo de dar cuenta de que he escrito “esquema” “válido” e “investigación”. No debía haberlo hecho, porque pienso que esta dogmática e irrelevante forma de ver las cosas puede ser una reconstrucción a posteriori de un viejo ideal tardopositivista: el del “híbrido Popper-Fisher” (Delgado, 2006a, 2006b), pero no es un esquema, ni es válido, ni representa lo que es una investigación.

Aunque, eso sí, plegarse a ese dogma ha facilitado el progreso personal de un importante número de *científicos* de las *ciencias del comportamiento* que han publicado sus trabajos siguiendo aquel *esquema* que creían *válido* de lo que era una *investigación científica*. Que han publicado, para su bien, en cualquier caso, porque si no publicas, ya se sabe, pereces.

REFERENCIAS

- Coleman, D. E. & Montgomery D. C. (1993). A systematic approach to planning for a designed industrial experiment. *Technometrics* 35 (1)1 - 12.
- Delgado, J. (2006a). Psicología en crisis. Metodología dogmática. Encuentros y desencuentros. *Anuario de Psicología*, 37(1-2), 7-26.
- Delgado, J. (2006b). Publicar sobre crisis y dogmas provoca encuentros y desencuentros. *Anuario de Psicología*, 37(1-2), 99-120.
- Delgado, J. (2008). Las difíciles relaciones entre teoría, práctica y metodología en la psicología del siglo XXI. (I) Sobre algunas disyuntivas. *Anuario de Psicología*, 39 (3), 387-409.
- Hacking, I. (1965). *Logic of statistical inference*. Cambridge: Cambridge University Press.
- Hacking, I. (1975). *El surgimiento de la probabilidad: Un estudio filosófico de las ideas tempranas acerca de la probabilidad, la inducción y la inferencia*. Barcelona: Gedisa, 1995.
- Hacking, I. (1983). *Representar e intervenir*. México: Paidós, 1996
- Hacking, I. (1990). *La domesticación del azar. La erosión del determinismo y el nacimiento de las ciencias del caos*. Barcelona: Gedisa, 1991
- Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge: Cambridge University Press.
- Mayo, D. G. (1981). In defense of the Neyman-Pearson theory of confidence intervals. *Philosophy of Science*, 48, 269-280.
- Mayo, D. G. (1982). On after-trial criticisms of Neyman-Pearson theory of statistics. En P. D. Asquith & T. Nickles (Eds.), *PSA 1982. Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Vol 1. East Lansing, Michigan: Philosophy of Science Association.
- Mayo, D. G. (1996). *Error and the growth of experimental knowing*. Chicago: University of Chicago Press.
- Mayo, D. & Spanos, A (2006). Severe testing as a basic concept in a Neyman-Pearson philosophy of induction. *British Journal for the Philosophy of Science*, 57 (2), 323-357.
- Suppes, P. (1962). Models of data. En E. Nagel, P. Suppes, & A. Tarski (Eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*. Stanford: Stanford University Press, 252-261. Editado en castellano en Patrick Suppes: *Estudios de filosofía y metodología de la ciencia*. Madrid: Alianza, 1988.
- Underwood, B. J. (1976). *Psicología experimental*. México: Trillas. 2ª reimpresión de la 1ª edición en castellano (original en inglés de 1966, 2ª edición).