

**INTRODUCCIÓN A LA INVESTIGACIÓN EN
TERAPIA ASISTIDA POR PERROS**

(2ª edición)

ALFONSO BLANCO PICABIA

LUIS RODRÍGUEZ FRANCO

(coords.)

INTRODUCCIÓN A LA INVESTIGACIÓN EN TERAPIA ASISTIDA POR PERROS (2ª EDICIÓN)

Los gráficos y marcas registradas que aparecen en las capturas de pantalla realizadas para este volumen pertenecen a sus legítimos dueños. Los autores de esta obra muestran su compromiso para modificar o enmendar cualquier error que vulnere dicho *copyright*.

No está permitida la reproducción total o parcial de este libro, ni su tratamiento informático, ni su transmisión de ninguna forma o por cualquier medio, ya sea electrónico, mecánico, por fotocopia, por registro u otros métodos, sin el permiso previo y por escrito de los titulares del *copyright*.

© Los autores (2010, 2012)

Diseño de cubierta: Pablo Sánchez Borrego

ISBN (1ª ed.): 978-84-693-0387-0
(2ª ed.): En curso

Registro legal (1ª ed.): SE-377-2010
(2ª ed.): En curso

Impreso en España / Printed in Spain

ÍNDICE

| | |
|---|-----------|
| 1. INICIACIÓN A LA PLATAFORMA WEBCT..... | 1 |
| (Luis Rodríguez Franco, Javier López-Cepero Borrego, Nuria Blanco Piñero y M^a Jesús Lara Medina). | |
| 1. Iniciando el curso. | 1 |
| 2. Configuración previa del ordenador personal. | 2 |
| 2.1. Configurar el navegador de Internet. | 2 |
| 2.2. Conseguir otros programas de utilidad. | 8 |
| 3. Acceder al servicio WebCT de la Universidad de Sevilla. | 10 |
| 3.1. Acceso a la plataforma WebCT. | 10 |
| 3.2. Primer contacto con la plataforma WebCT. | 12 |
| 3.3. Control de novedades en la plataforma. | 17 |
| 4. Las opciones de uso y servicios de la WebCT del curso. | 19 |
| 4.1. Correo electrónico. | 19 |
| 4.2. Foros de discusión. | 25 |
| 4.3. Calendario. | 27 |
| 4.4. Evaluaciones. | 28 |
| 4.5. Mis calificaciones. | 30 |
| 2. BÚSQUEDA BIBLIOGRÁFICA..... | 31 |
| (Javier López-Cepero Borrego). | |
| 1. Buscadores generalistas. | 32 |
| 1.1. Uso básico de los buscadores generalistas. | 32 |
| 1.2. Uso avanzado de los buscadores generalistas. | 37 |
| 1.3. Uso de Google Académico. | 42 |
| 1.4. A modo de resumen. | 46 |
| 2. Biblioteca de la Universidad de Sevilla. | 48 |
| 2.1. Buscar información científica en la Web US. | 48 |
| 2.2. ¿Qué tipos de recursos existen? | 51 |
| 2.3. Búsquedas en FAMA. | 55 |
| 2.4. Búsquedas de Recursos-e. | 61 |

3. LA INVESTIGACIÓN: MÉTODO Y PROCEDIMIENTO..... 81

(Nuria Blanco Piñero).

| | |
|---|-----|
| 1. Conceptos básicos. | 83 |
| 1.1. Población. | 83 |
| 1.2. Contexto o marco teórico. | 86 |
| 1.3. Distribución normal de la muestra de estudio. | 91 |
| 1.4. Homocedasticidad o igualdad de varianzas. | 93 |
| 1.5. Objetivos de estudio. | 93 |
| 1.6. Hipótesis. | 95 |
| 1.7. Medir. | 98 |
| 1.8. Validez de la recogida de datos. | 99 |
| 1.9. Fiabilidad. | 100 |
| 2. Tipos de estudios en función de la metodología. | 101 |
| 2.1. Estudios Experimentales. | 102 |
| 2.2. Estudios cuasi-experimentales. | 103 |
| 2.3. Estudios no experimentales u observacionales. | 105 |
| 3. Tipos de estudio en función del diseño de medida | 106 |
| 3.1. Estudios transversales o diseños entregrupos. | 106 |
| 3.2. Estudios longitudinales o diseños intragrupos. | 107 |
| 4. Desarrollo del trabajo de investigación. | 108 |
| 4.1. Etapa conceptual (teórica). | 109 |
| 4.2. Etapa de planificación (diseño). | 113 |
| 4.3. Etapa empírica (práctica). | 122 |
| 4.4. Etapa interpretativa (significado). | 124 |
| 5. Referencias. | 127 |
| 6. Anexos. | 128 |
| Anexo 1. | 129 |
| Anexo 2. | 137 |
| Anexo 3. | 149 |

| | |
|---|------------|
| 4. EVALUACIÓN CON MÉTODOS OBSERVACIONALES..... | 159 |
| (Javier López-Cepero Borrego, Marta Ramos Carrera, Cristina Soria Poveda y Marta Blanco Piñero). | |
| 1. Plantear un problema de investigación. | 161 |
| 2. Recogida de datos. | 165 |
| 2.1. Nivel de participación. | 165 |
| 2.2. Grado de sistematización. | 167 |
| 2.3. Tipos de respuesta observable. | 171 |
| 2.4. Grado de inferencia. | 174 |
| 3. Técnicas de registro. | 176 |
| 3.1. Registro narrativo. | 176 |
| 3.2. Escalas de apreciación (Rating scales). | 178 |
| 3.3. Protocolos de observación conductual. | 180 |
| 3.4. A modo de conclusión. | 182 |
| 4. Diseño de un sistema de categorías. | 183 |
| 4.1. Determinar las unidades de análisis. | 184 |
| 4.2. Definir las unidades de medida. | 191 |
| 4.3. Diseñar el formato de registro conductual. | 196 |
| 4.4. Análisis de la calidad del dato. | 204 |
| 4.5. Revisar el sistema de categorías. | 212 |
| 5. Referencias. | 214 |
| 5. CREACIÓN DE TEXTOS CIENTÍFICOS..... | 217 |
| (Nuria Blanco Piñero y M^a Ángeles Perea Mediavilla). | |
| 1. Un paso previo. Revisar material. | 218 |
| 2. Publicar en una revista. | 222 |
| 2.1. Publicación de artículos científicos en revistas. | 227 |
| 2.2. Estructura básica de un artículo. | 228 |
| 2.3. En resumen. | 243 |

| | |
|---|-----|
| 3. Difusión en reuniones y congresos profesionales. | 243 |
| 3.1. Tipos de publicaciones. | 245 |
| 3.2. Pasos para realizar una aportación. | 248 |
| 3.3. Cómo preparar una comunicación oral corta. | 249 |
| 4. Referencias. | 252 |
| 5. Anexos. | 255 |
| Anexo 1. Normas Vancouver de citación. | 255 |
| Anexo 2. Normas APA de revisión bibliográfica. | 261 |

PRÓLOGO

Las Terapias Asistidas por Animales han ganado presencia en los medios de comunicación durante los últimos años, sirviendo de escaparate para que la población general conozca su existencia. En consecuencia, se ha generado una clara expectación en torno a las posibilidades que éstas ofrecen.

Sin embargo, la visibilidad de este recurso terapéutico continúa siendo bastante reducida tanto a nivel institucional como académico. A día de hoy, sigue resultando muy difícil introducir un animal dentro de determinados recursos sanitarios (como puedan ser los hospitales), a la vez que algunas autoridades en la materia (como Alan Beck, Aubrey Fine, Aaron Katcher, James Serpell y otros) insisten en señalar la escasa productividad científica registrada en torno al campo de las Terapias Asistidas por Animales. En otras palabras: a nivel profesional, queda mucho camino por recorrer.

Abrir un hueco para las Terapias Asistidas por Animales en la Universidad supuso (y aún supone) un reto importante,

tanto por su dificultad como por su alto valor estratégico. Instaurar estas intervenciones como alternativas válidas y confiables a otras aproximaciones terapéuticas exige de una solidez que sólo la investigación es capaz de aportar; y en este ámbito, la Universidad está llamada a ejercer como líder, como hogar, como andamio que guíe un crecimiento sólido y constante.

Por eso hoy, a principios de 2012, y con este manual recién salido de la imprenta, nos llena de orgullo poder inaugurar la III Edición del *Master en Aplicaciones del Perro a la Terapéutica Humana*, un nuevo curso en que esperamos seguir aportando nuestro particular granito de arena a las tareas de formación, difusión y desarrollo de las Terapias Asistidas por Animales en España.

Gracias a todos los que nos acompañáis en este esfuerzo.



CAPÍTULO 4. EVALUACIÓN CON MÉTODOS OBSERVACIONALES

Javier López-Cepero Borrego, Marta Ramos Carrera, Cristina Soria
Poveda y Marta Blanco Piñero

Observar y sus derivados son palabras de uso muy frecuente en nuestro idioma. El diccionario de la RAE define la raíz, de un modo muy sucinto, como *examinar atentamente*; es probable que todos podamos coincidir en aceptar esta propuesta pero, si lo pensamos detenidamente, una nueva pregunta emerge de la anterior. ¿Qué significa *atentamente*?

En las ciencias del comportamiento, la observación es el prototipo de herramienta para la recogida de datos (Fernández-Ballesteros, 2004). Sin embargo, cuando hablemos de observación como método científico, debemos tener en cuenta una serie de requisitos que ayudan a mejorar y objetivar nuestros resultados, a la vez que trazan una separación entre observación científica y no científica (Anguera, 1985):

1. La observación científica sirve para cubrir un determinado objetivo de investigación, busca cumplir una finalidad concreta, como puede ser comprobar una hipótesis.
2. Es sistemática, planificada. Algunos autores, como León y Montero (2003), señalan ésta como la característica más definitoria.
3. Dentro de su organización, establece mecanismos de control dirigidos a contrastar la calidad del proceso.
4. Debe comprobar la validez y fiabilidad de los resultados.

En la literatura pueden encontrarse muchas propuestas acerca de qué pasos debemos seguir para llevar a cabo una investigación mediante métodos observacionales. En el presente texto, se recupera el esquema que propone Anguera (1985) por su simplicidad y como guía sencilla del proceso en cuatro pasos:

1. Formulación de un problema
2. Recogida de datos
3. Análisis de datos
4. Comunicación de resultados

A lo largo de este capítulo, se van a explorar los dos primeros puntos y parte del tercero. Así, se conocerán los elementos básicos del planteamiento de un problema de investigación, la creación de herramientas para evaluar, las decisiones necesarias para administrar estas herramientas, y el análisis de la calidad de los datos (validez y fiabilidad). El análisis de los resultados (para corroborar nuestras hipótesis) y los medios de comunicación científica quedan fuera del presente capítulo.

1. PLANTEAR UN PROBLEMA DE INVESTIGACIÓN

Tal como describe Arnau (1995), existen dos grandes tradiciones dentro del campo de la investigación; estas son, la tradición experimental y la tradición asociativa. Cada una de ellas se diferencia en dos grandes aspectos:

1. El tipo de hipótesis que plantean.
2. El tipo de diseño que llevan a cabo.

A continuación se mencionan, brevemente, las características de cada una de estas tradiciones, para llegar a una mejor comprensión del tema tratado.

La tradición experimental, se ha centrado en plantear *hipótesis de tipo causal*, es decir, intenta explicar cómo una variable determinada (variable independiente) puede modificar o influir sobre la variable objeto de estudio (variable dependiente) (Blanco, 2010). Un ejemplo de hipótesis causal sería *comprobar si la presencia de un perro* (variable independiente) *disminuye los sentimientos de soledad que sufren muchas personas mayores* (variable dependiente). A partir de este tipo de hipótesis, el diseño que se lleva a cabo puede ser experimental o cuasi-experimental.

El *experimental* se caracteriza por ser el investigador quién lleva a cabo la elección de los individuos que van a formar parte de la investigación, además de la asignación de los grupos al azar. Una vez asignados al azar, uno de ellos es el que va a ser tratado, y el otro actúa como grupo control, es decir, no recibe tratamiento. En este tipo de diseño, el investigador puede manipular las variables independientes y controlar las variables extrañas, es decir, aquellas que no

forman parte del objeto de estudio, pero que pueden influir en los resultados.

Por otro lado, en el diseño *cuasi-experimental*, los grupos con los que se trabaja no pasan por la fase de aleatorización (como ocurre en el diseño experimental), y el investigador ejerce un menor control sobre las variables extrañas.

Con respecto a la *tradición asociativa*, esta se caracteriza por crear hipótesis a través de las cuales se intenta comprobar el grado de asociación entre las variables o la magnitud de cambio entre estas (Arnau, 1995). De modo que la obtención de los datos puede llevarse a cabo a través de observación indirecta (mediante encuestas, por ejemplo) o directa (por medio de la metodología observacional, sobre la que se está trabajando a lo largo de este capítulo).

Anguera, Blanco y Losada (2001), proponen varios criterios, cuya combinación da lugar a posibles diseños observacionales. Veamos cuáles son cada uno de estos por separado.

- En primer lugar, propusieron la dicotomía *idiográfico* vs. *nomotético*. Un estudio *idiográfico* evalúa unidades (por ejemplo, un único sujeto o un único grupo que

conforme una unidad). Por otro lado, un estudio nomotético hace referencia a colectivos de unidades, es decir, varios sujetos individuales o varios grupos.

- En segundo lugar, plantearon la dicotomía *puntual vs. seguimiento*. Un estudio puntual es aquel en el que sólo se lleva a cabo la evaluación u observación en un único momento; mientras que el seguimiento hace referencia a varias sesiones de evaluación a lo largo del tiempo.
- Por último, estos autores incorporaron la dicotomía *unidimensional vs. multidimensional*. La primera de ellas conlleva que sólo sea un único elemento el que se estudia u observa. Sin embargo, multidimensional implica que son varios los niveles observados.

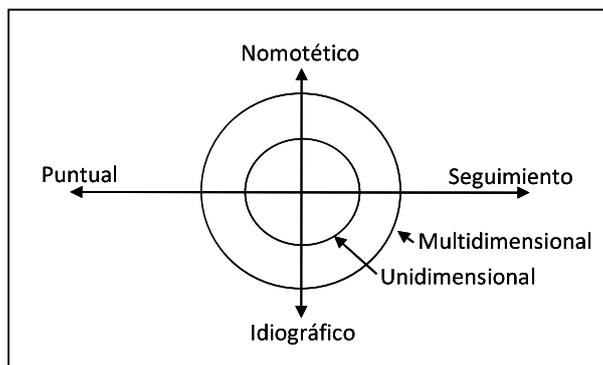


Gráfico 1. Propuesta de Anguera, Blanco y Losada (2001).

Una vez decididos aquellos aspectos que se van a tener en cuenta y que conformarán nuestro problema de investigación, estaremos en disposición de iniciar el segundo paso: la recogida de los datos.

2. RECOGIDA DE DATOS

Una vez decididos los objetivos exactos de nuestra investigación, es necesario elegir (y a menudo, confeccionar) las herramientas de evaluación para contrastar las hipótesis planteadas. A continuación se detallan algunas de las cuestiones a tener en cuenta para plantear el diseño de la recogida de datos.

2.1. Nivel de participación

Para Anguera (1985), el grado de participación del observador puede tener tres niveles: externa o no participante; interna o participante; y autoobservación.

La *observación no participante* se caracteriza por ser llevada a cabo por una persona ajena al grupo observado. Tiene la ventaja de que, al no existir interacción directa con los sujetos observados, todos los recursos del evaluador quedan libres para realizar el registro; sin embargo, no puede pasarse por alto que la presencia de un extraño contamina el contexto que se observa, y ello resta naturalidad a la situación.

La *observación participante* es realizada por una persona que interactúa con las personas observadas. Esto no es sinónimo de total espontaneidad ni de que el investigador olvide cuál es su labor, aunque debe atenderse a los posibles sesgos que la interacción pueda estar añadiendo a las conductas registradas.

Si el observador se incorpora al contexto para realizar la investigación, puede denominarse *observación-acción*, mientras que cuando la persona que observa pertenece al contexto antes de iniciar la investigación, puede hablarse de *participación-observación*. La mayor parte de las investigaciones observacionales con humanos utilizan una de estas modalidades.

Por último, encontramos la *autoobservación*, en la que el evaluador y el evaluado es una misma persona. Para Fernández-Ballesteros (2004), este tipo de observación no puede ser tratado de modo conjunto con el resto, ya que mientras que la observación de otros sujetos nos ofrece información sólo de las conductas evidentes (externalizadas, directamente observables), la autoobservación nos da acceso a información no evidenciable, como las pertenecientes al campo cognitivo del sujeto. Para esta autora, la observación debe ofrecer información acerca de los *actos*, pero no sobre los *objetos* observados en sí mismos. Si bien la autoobservación nos permite alcanzar una información más rica que las anteriores, las dificultades metodológicas que conlleva son múltiples (como la imposibilidad de comprobar la fiabilidad de los datos) y esto obliga a poner en cuarentena los hallazgos derivados de su uso.

2.2. Grado de sistematización

Recogida de datos es una etiqueta que designa un proceso completo, y no un momento puntual en el tiempo. Con ello se hace referencia al hecho de que la recogida no tiene por qué darse en un único paso; normalmente, cuando

se acomete una investigación partiendo desde cero, se necesitan varios pasos sucesivos para crear y perfeccionar el material de evaluación. Puede resultar útil presentar las aportaciones de Anguera (1985) acerca del grado de estructuración de las observaciones. Esta autora propone tres niveles de sistematización en la observación según el momento de la recogida de datos: observación no sistematizada, observación sistematizada o controlada, y observación muy sistematizada.

Por *observación no sistematizada* entendemos aquella que se emprende sin disponer de un marco teórico previo. Esto no significa que no se tengan ciertas expectativas acerca de qué puede ocurrir (el observador es un agente activo, no es neutro), pero esta es una fase de toma de contacto con el objeto de estudio. En esta primera fase no estructurada, el objetivo inmediato es detectar qué conductas nos resultan de interés y de qué modo se pueden evaluar. En este momento del proceso no se dispone de un conjunto de categorías que nos facilite diferenciar tipos de conducta de un modo claro, ya que aún no se ha decidido qué comportamientos son de interés y cuáles no.

Fernández-Ballesteros (2004) propone el *registro narrativo* como técnica de elección para este momento. De un modo básico, se puede describir como la anotación de frases completas en las que el observador registra los hechos que considera que ocurren durante la sesión. Esta técnica se detalla y ejemplifica en el apartado 3.1 de este capítulo.

Existen varias razones para considerar la observación no sistematizada sólo como un paso previo al registro estructurado. Dado que el observador anota lo que percibe sin hacer uso de un instrumento bien definido, existen muchas posibilidades de que distintos observadores registren datos muy divergentes. Además, dejarse llevar por la intuición facilita que los eventos más salientes (una persona que habla en voz muy alta, o una conducta llamativa en algún sentido) ensombrezcan a otros en la elección. Y a todo esto hay que sumar el hecho de que improvisar formas de expresar las propias percepciones supone un gasto de recursos cognitivos que puede competir con el hecho de atender adecuadamente a la situación evaluada.

La *observación sistematizada* (o *controlada*) llega en un segundo momento. Su nombre proviene del uso de catálogos de conducta, ya sea mediante categorías o con formatos de

campo (ver apartado 3.3), que permiten alcanzar un mayor grado de detalle en los registros. Como contrapartida, las observaciones quedan sesgadas por la herramienta usada, ya que se nos obliga a clasificar lo que percibimos en base a unas premisas determinadas.

Las categorías usadas pueden depurarse entre una sesión y la siguiente. A tal efecto, disponer de grabaciones en vídeo facilita tomar decisiones sobre la calidad del instrumento, comprobar su fiabilidad y su validez, etc. Disponer de este material puede ser muy relevante para comprobar los efectos que ciertas conductas tienen en un momento posterior (contingencia), y que por su latencia pueden pasar desapercibidas en el momento del registro.

Por último, encontramos la *observación muy sistematizada*. Anguera reserva esta categoría para las situaciones en las que el investigador tiene un alto control sobre las variables contextuales en que la conducta tiene lugar (1985). La teoría que sirve de marco a la investigación y los métodos de observación están bien establecidos, y se busca poder replicar la situación repetidas veces, de un modo muy similar a los diseños cuasi-experimentales. Como ejemplo paradigmático, podemos citar las investigaciones sobre

interacción social en grupos de Bales (1950). Este tipo de observación tiene un problema básico, que radica en la dificultad para generalizar las conclusiones desde un contexto tan específico hacia ambientes naturales.

2.3. Tipos de respuesta observable

Existe un abanico muy amplio de conductas que se pueden registrar, tanto que resulta imposible detallarlas todas. Sin embargo, todos los registros tienen un punto en común: han de referirse a conductas evidentes, que sean externalizadas. A continuación, se propone una clasificación de los tipos más frecuentemente analizados.

Las *conductas verbales* (o *digitales*; Jackson, Watzlawick y Beavin, 1967) son muy sencillas de reconocer. Aunque categorizarlas correctamente puede resultar complejo, existen varias cuestiones de utilidad para todos los casos. Por ejemplo, decidir si vamos a clasificar las conductas en base al nivel denotativo (literalmente, con un nivel de inferencia bajo) o connotativo (por el significado o función que, se interpreta, tiene la verbalización), o si se va a contemplar como verbalizaciones las articulaciones que no

conformen frases con sentido evidente. El cuadro 1 recoge algunos ejemplos.

| |
|---|
| <p><i>Me gustaría saber más del tema</i></p> <ul style="list-style-type: none">• Puede clasificarse como afirmación, o• Puede interpretarse como petición de información o pregunta <p>(Una persona recibe una reprimenda y gruñe) <i>Grrrrr</i></p> <ul style="list-style-type: none">• No ha emitido una conducta con contenido verbal objetivable• Puede inferirse que el sonido tiene la función de una queja |
|---|

Cuadro 1. Ambigüedad en la clasificación de conductas verbales.

Mayor dificultad supone describir las *conductas no verbales* (o *analógicas*; Jackson y cols., 1967). En muchas ocasiones, los mensajes verbalizados entran en conflicto con estos niveles analógicos, como en el caso de una persona que muestra estar de acuerdo acerca de algo y, por otra parte, arruga la frente y las aletas de la nariz, o utiliza un tono de voz que denota enfado; por esta razón, es importante contemplar ambos niveles de análisis y determinar cómo vamos a clasificar este tipo de circunstancias. Nuevamente, explorar todas las conductas posibles es inviable, por lo que la

experiencia directa y el análisis de la fiabilidad entre jueces será la que ayude a refinar los sistemas de registro que diseñemos.

De un modo esquemático, se establecen categorías de registro para:

- Prosodia (tono de voz, fluidez, velocidad...)
- Expresión facial (mirada, nariz, arrugas en la frente, boca... puede encontrarse un completo catálogo en Anguera, 1985: pp. 97-98)
- Movimiento de extremidades (retorcerse las manos, cruzar las piernas...)
- Movimientos por el contexto observado (por zonas ocupadas, por proximidad con otras personas...)

Al definir las conductas observables, también se debe tener en cuenta la distinción que aportan Bakeman y Quera (1996) entre estado y evento. Mientras que las conductas estado tienen una duración prolongada en el tiempo, como *mantener una conversación* o *permanecer sentado*, los eventos son momentáneos y su duración resulta menos relevante, como *parpadear* o *cambiar la mirada de un interlocutor a otro*. Por supuesto, una misma conducta puede

ser definida como evento o como estado, según sea necesario para cubrir nuestros objetivos de investigación; una conversación puede medirse por su duración como *estado*, o bien descomponerse y registrarse en función del número de veces que los interlocutores pasan de ser emisores a receptores como *eventos*.

2.4. Grado de inferencia

La observación permite registrar cosas que ocurren en el ámbito público, que se expresan de manera evidente. Las cogniciones (ideas, recuerdos...), por ejemplo, no pueden registrarse directamente; sólo podemos conocer lo que los sujetos nos comentan (verbalizando) o demuestran (mediante registros electrofisiológicos, etc.).

Las conductas evidentes y concretas (como levantar una mano, mirar un objeto, sentarse....) tienen un grado de inferencia muy bajo, y por ello, distintos observadores entrenados suelen tener una gran coincidencia en su registro. Sin embargo, en múltiples ocasiones este tipo de conducta no ofrece la información necesaria para cubrir nuestros objetivos de investigación; por ejemplo, si interesa registrar una conducta no evidente (como el uso del engaño con la

intención de conseguir un determinado final) no hay que conformarse con registrar lo evidente. A modo de ejemplo, se plantea la situación recogida en el cuadro 2.

| |
|--|
| Un niño (sujeto A) llega a la sala de observación y ve un caramelo sobre la mesa. Lo coge y se lo guarda en el bolsillo. A continuación, un investigador entra en la sala y busca el caramelo, preguntando al niño; el niño responde que no sabe dónde está. |
|--|

Cuadro 2. Situación hipotética de observación.

Si el objetivo es estudiar qué *dice* el niño, la transcripción ofrece una información concreta y suficiente; si se pregunta a varios observadores, el nivel de acuerdo sobre qué dice el niño será alto. Sin embargo, si lo que se quiere registrar es si efectivamente el niño *sabe* dónde está el caramelo, el registro puede dar lugar a disensiones: ¿Está engañando al investigador? ¿Ha olvidado que lo ha cogido? ¿Se refiere a otro caramelo distinto cuando responde que no lo sabe? Lo cierto es que no se tiene manera de objetivar las intenciones del sujeto observado, y que todos los registros que se realizan al respecto parten de inferencias.

Kerlinger y Lee (2001) advierten de la imposibilidad de eliminar la inferencia de la observación. Sin embargo, consideran que, sometida al control y sistematización exigidos para la investigación, esta interpretación del observador se convierte en una herramienta más. Por ello, el énfasis debe ponerse en mejorar las medidas de calidad, y no tanto en intentar eliminar todo tipo de (inevitables) interpretaciones.

3. TÉCNICAS DE REGISTRO

Aunque en la actualidad se pueden encontrar herramientas de evaluación de fiabilidad y validez contrastadas para un amplio abanico de objetos de estudio, es muy frecuente encontrarse ante la necesidad de confeccionar una herramienta de registro original y adaptada a nuestros objetivos. Por esta razón, se hace necesario conocer las opciones más comunes entre las que podemos elegir para llevar a cabo nuestra investigación.

3.1. Registro narrativo

Es una herramienta útil al inicio de la recogida de datos, sobre todo en el caso de no disponer de una

herramienta de evaluación capaz de cubrir los objetivos de investigación. El registro narrativo puede tomar muchas formas, desde la anotación mediante frases completas de las percepciones del observador hasta un conjunto de abreviaturas y palabras clave (cuadro 3).

| |
|--|
| <p>El paciente se sienta en una silla de la sala de espera y mira repetidas veces su reloj con impaciencia. Mira con frecuencia la puerta, mientras mueve la pierna con nerviosismo. Se levanta y se lleva la mano a la boca. Se mueve por la sala, mientras mira por la ventana y suspira. Vuelve a sentarse y da pequeños golpes con el pie en el suelo...</p> |
|--|

Cuadro 3. Registro narrativo de un paciente.

Su utilidad radica en que ofrece al investigador un primer contacto con el objeto de estudio, facilitando que éste realice, de manera sucesiva, abstracciones desde las conductas percibidas, reuniéndolas y clasificándolas paulatinamente. De este modo, pueden enlazarse conceptos para formar teorías o hipótesis.

Anguera (1985) apunta que, dentro de la relativa libertad de este tipo de registro, resulta útil recabar información sobre: participantes (número, sexo, edad,

relación entre ellos...), ambiente (disposición física, normas existentes...), objetivo de los participantes (qué los lleva a estar en ese contexto), comportamiento (desencadenantes de la conducta, objetivo aparente, interacciones, efectos en el futuro...) y características temporales (frecuencia, duración...).

Los registros narrativos presentan ciertas dificultades metodológicas. Cone y Foster (1982; tomado de Anguera, 1985) resaltan dos posibles fuentes de error, derivadas de la falta de categorías para clasificar la conducta observada; la agrupación de conductas diferentes como una misma cosa, y la descripción divergente de conductas muy cercanas entre sí. Cuesta decidir si, en el ejemplo expuesto, los golpes que el paciente da con el pie en el suelo son o no la misma cosa que el movimiento *nervioso* de la pierna.

Por todo lo expuesto, queda clara la necesidad de elaborar sistemas de categorías que permitan clasificar las observaciones de un modo válido y fiable.

3.2. Escalas de apreciación (Rating scales)

Las escalas de apreciación son una herramienta de sencilla aplicación y corrección. Se componen de un conjunto de conductas sobre las que se pregunta la frecuencia de

aparición, mediante el uso de una escala ordinal (con niveles ordenados de mayor a menor, como las de tipo Likert; ver tabla 1).

| | Nunca | Pocas veces | A veces | Con frecuencia | Casi siempre |
|---|-------|-------------|---------|----------------|--------------|
| Se muestra nervioso | | | | | |
| Se queja de sus síntomas | | | | | |
| Solicita ayuda al cuidador | | | | | |
| Participa en las actividades propuestas | | | | | |
| ... | | | | | |

Tabla 1. Ejemplo de escala de apreciación o *Rating scale*.

En este caso, el observador evalúa de un modo general la aparición de determinadas conductas. Las escalas de apreciación son temporalmente muy económicas, pero introducen el interrogante sobre la exactitud de la percepción del evaluador. Los juicios globales, en muchas ocasiones, no coinciden con los datos registrados mediante técnicas más detalladas (como los protocolos de observación conductual);

por ello, esta técnica se recomienda para casos en que sea inviable llevar a cabo otro tipo de registros.

3.3. Protocolos de observación conductual

Son la principal herramienta que se va a usar cuando realicemos observaciones. Dentro de esta denominación se pueden encontrar *sistemas de categorías, sistemas de signos y formatos de campo*.

Los *sistemas de categorías* y los *sistemas de signos* se conforman de una lista de elementos conductuales que pueden aparecer durante la sesión (Anguera, 1985). Estas conductas son mutuamente excluyentes entre sí, de modo que es imposible clasificar a un mismo evento usando dos categorías o signos a la vez. La diferencia entre categorías y signos es el nivel de detalle en la observación; así, las categorías suelen hacer referencia a conjuntos de conductas, y exigen un mayor grado de interpretación por parte del observador, mientras que los signos son elementos muy concretos, objetivables, y exigen una inferencia menor. Lo vemos en el cuadro 4.

| |
|---|
| <p>Categoría: (conjunto de conductas agrupadas)</p> <ol style="list-style-type: none"> 1. Conducta verbal del paciente ... <p>Signos: (conductas concretas y pormenorizadas)</p> <ol style="list-style-type: none"> 1. Realiza una pregunta 2. Realiza una afirmación 3. Realiza una negación 4. Emite sonidos sin verbalizar |
|---|

Cuadro 4. Ejemplos de categoría y signos.

Así pues, sistemas de signos y categorías tienen ciertos parecidos, si bien abogan por distintos niveles de detalle a la hora de explicitar las unidades a las que se debe atender.

Por otra parte, aparecen los *formatos de campo*. A diferencia de los anteriores, el formato de campo de conforma de subcategorías no excluyentes entre sí. De este modo, para un mismo evento observado, puede ser necesario realizar varias anotaciones, lo cual puede elevar el nivel de dificultad del registro (cuadro 5).

Si observamos la *conducta verbal del paciente*, podemos contemplar:

A. quién se dirige:

1. A otro paciente
2. A un profesional sanitario
3. A un acompañante...

B. Contenido:

1. Queja
2. Propuesta
3. Petición...

C. Lugar:

1. Sala de espera
2. Pasillo...

Así pues, si el paciente realiza una queja a una enfermera en un pasillo, el código a registrar sería [A2 B1 C2]

Cuadro 5. Ejemplo de formato de campo.

3.4. A modo de conclusión

Las estrategias de registro expuestas pueden combinarse entre sí a lo largo de la investigación. Según comenta Anguera (1985), un proceso ideal debe partir de una perspectiva abierta, sin comprometerse con ninguna teoría (como en un registro narrativo) y continuar con una abstracción progresiva sobre qué conductas se van a observar, y de qué son indicadoras (en función del grado de inferencia asumido).

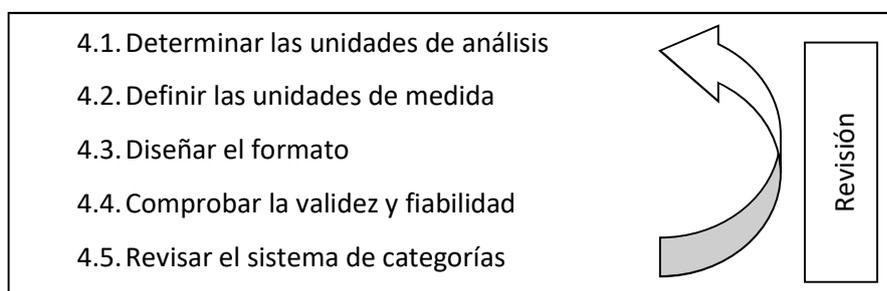
En realidad, la clasificación propuesta pretende servir sólo como una primera guía conceptual, pero se ha de ser conscientes de que detallar las virtudes y defectos de cada uno de los métodos de registro expuestos sobrepasa los límites del presente capítulo. Atendiendo a la popularidad de cada tipo de registro, se ha optado por dedicar especial atención a confeccionar un sistema de categorías, detallando los pasos necesarios para crear una herramienta de observación adecuada para cubrir nuestros objetivos de investigación.

4. DISEÑO DE UN SISTEMA DE CATEGORÍAS

Según hemos comprobado hasta el momento, investigar es un proceso que puede subdividirse en distintos momentos, y se suele ser necesario un tiempo de evolución para estar en disposición de pasar de un apartado al siguiente. Así, el grado de sistematización tiende a crecer a lo largo del proceso, a la vez que se transforman o cambian los instrumentos de evaluación.

Dado que este capítulo se dirige a alumnos de postgrado que se enfrentan, en la mayoría de casos, a su primera incursión en el campo de la investigación, consideramos de utilidad dedicar un apartado completo a recorrer desde principio a final el proceso de creación de un sistema de categorías. Para ello, propondremos una guía con los pasos necesarios para lograr un instrumento adecuado a nuestros objetivos.

De un modo esquemático, vamos a desarrollar nuestro sistema de categorías a través de los siguientes pasos recogidos en el cuadro 6.



Cuadro 6. Esquema del proceso de creación de un sistema de categorías.

4.1. Determinar las unidades de análisis

El primer reto a superar consiste en determinar qué cosas vamos a observar. Es probable que partamos de un

cierto conocimiento previo, ya sea por la experiencia propia o porque hayamos dedicado el inicio de la investigación a tomar contacto con el objeto de estudio; sin embargo, debemos tener en cuenta que la intuición personal no es suficiente garantía para realizar esta selección.

Para Fernández-Ballesteros (2004), las conductas seleccionadas deben cumplir con las tres “D”: Denominación, Delimitación y Definición. La *denominación* hace referencia al etiquetado, a disponer de un nombre que la haga reconocible; la *delimitación*, a que debemos considerar elementos que pueden transformar la conducta observada (momento del día, lugar de aparición...), y por lo tanto, plantear los límites de qué es y qué no es la conducta; y por *definición* hacemos referencia a plasmar por escrito y de manera operativa los contenidos de la conducta (esto es, que quede explicada de modo que pueda evaluarse de modo válido y fiable). El cuadro 7 muestra un ejemplo de las “3D”.

(ver página siguiente)

En una residencia de personas mayores, el equipo de investigación quiere intervenir sobre el “aislamiento” de algunos residentes. Para ello necesitan evaluar en cuáles de ellos se da esta situación. Sin embargo, los investigadores encuentran que los profesionales que atienden a estos mayores no alcanzan un acuerdo al determinar quienes se encuentran “aislados”.

Para salir de dudas, se procede a buscar una definición operativa.

1. En primer lugar, se propone cambiar la etiqueta que se usaba hasta el momento, ya que “aislamiento” puede referirse tanto a la persona que no sale nunca de la habitación, como a la que no se suma a ciertas actividades de grupo por preferir otras. Se consulta a los profesionales qué tipo de situación es la que quieren intervenir, y se baraja “retraimiento” como nueva etiqueta. Finalmente, se opta por “relación social”, ya que es más sencillo definir y registrar la aparición de conductas que su no aparición.
2. Se delimitan tiempo y contexto de estas conductas: “Las relaciones sociales se evalúan en las horas que van desde el desayuno a la cena”; “la observación tendrá lugar en sala de estar y habitaciones”...
3. Se realiza una primera definición del objeto de estudio: “se considera relación social a participar de actividades, programadas o no, con otros residentes para favorecer el ocio y el contacto social”

Cuadro 7. Denominación, delimitación y definición de relación social para su observación.

Para Anguera (1985), una definición puede hacerse más operativa añadiendo indicadores concretos, ejemplos y contraejemplos de las conductas a observar (ver tabla 2):

- Un indicador es la unidad más concreta posible al hablar de una conducta, y debe ser fácilmente medible. Para la temperatura, los grados marcados sobre un termómetro son el indicador. Algunos de los temas clásicos en ciencias del comportamiento, como estrés, violencia, ansiedad... han sido estudiados desde distintas perspectivas, registrando distintos indicadores, y eso puede llevarnos a conclusiones muy dispares a pesar de partir de la misma etiqueta. Los indicadores explicitan qué incluimos bajo las etiquetas que elegimos.
- Con los ejemplos y contraejemplos tratamos de ilustrar las situaciones a las que los observadores se van a enfrentar, intentando que su capacidad de discriminación entre situaciones aumente. Cuanto mayor es la claridad y exactitud del catálogo de conductas a observar, mayor será el acuerdo entre jueces, y con ello, la fiabilidad de los datos.

| Categoría | Indicadores | Ejemplos y contraejemplos |
|---|---|---|
| Participación en actividades de ocio con residentes | Participan de las actividades de ocio programadas por el centro | Baile de salón, campeonatos deportivos... convocados por la dirección del centro |
| | Realizan actividades de ocio organizadas por residentes | Jugar a las cartas, salir a pasear u otras actividades propuestas por los residentes |
| | * No se registra esta categoría cuando realice actividades de ocio en solitario, como la lectura, escritura, ver la televisión... aunque en la sala haya otras personas | |
| Relación con otros residentes | Visitas presenciales a otros residentes | Recibir visitas en su cuarto o visitar los cuartos de otros residentes |
| | Contacto a distancia con otros residentes | Telefonar, escribir correspondencia en papel o formato digital, o usar otros medios similares |
| | * No se registra si las visitas o llamadas son de profesionales del centro | |

Tabla 2. Indicadores y ejemplos y contraejemplos para facilitar la observación.

El catálogo ofrecido en el ejemplo tiene ciertos problemas que saldrían a la luz en cuanto tratemos de utilizar este sistema de categorías; sin embargo, tener en cuenta dos

principios básicos hace posible anticiparse y ahorrar tiempo y esfuerzo:

- *Exhaustividad:* las categorías propuestas deben contemplar todas las acciones de interés, ya que las que no queden descritas no podrán registrarse. El sistema de categorías contempla las actividades de ocio y de relación social entre los residentes, y excluye las realizadas con los profesionales del centro, pero no menciona qué hacer, por ejemplo, con las actividades realizadas con familiares. En el caso de que un familiar lleve al mayor de vacaciones o que acuda al centro para acompañarlo durante unas horas, nuestro registro no tiene prevista una categoría donde anotarlo, y con ello, perdemos información relevante. Antes de dar por cerrado el sistema de categorías, debemos pensar en situaciones que, si bien pueden quedar fuera de nuestro objetivo inmediato, puedan tener una influencia sobre el resultado del registro.
- *Exclusividad mutua:* las categorías propuestas deben describir conductas de manera excluyente, o lo que es lo mismo, cada una de las categorías debe servir para registrar de manera específica unas acciones, pero

evitar las que sean distintas. Supongamos que un residente visita el cuarto de un compañero entre las 16 y 17hs. para jugar a las cartas:

| Acción observada | Indicadores | Categorías |
|---|----------------------------------|--|
| Jugar a las cartas en el cuarto de un compañero | Visitar cuarto de otro residente | Relación con otros residentes |
| | Jugar a las cartas | Participar en actividades de ocio con residentes |
| En este caso, el observador puede verse en la duda de cuál de las dos categorías registrar para el horario 16-17hs. Esto facilita que distintos observadores registren de distinta manera una misma acción, con lo que la fiabilidad del registro descende. | | |

Tabla 3. Definición (no exclusiva) de conductas para la observación.

Cuando determinemos nuestras unidades de análisis, conseguir exhaustividad y mutua exclusividad puede necesitar de un refinamiento paulatino de nuestro sistema de categorías. Si creamos un sistema partiendo de conceptos teóricos (*estrategia racional*; Anguera, 1985), la probabilidad de que nuestra herramienta de registro presente fallos es superior a si utilizamos una *estrategia empírica*, creando las

categorías a través de una abstracción basada en la observación, aunque ésta no sea sistematizada al principio.

4.2. Definir las unidades de medida

Una vez seleccionado el abanico de conductas que van a ser registradas, debemos considerar qué tipos de datos necesitamos anotar en dicho registro. Álvarez (1975) propone un decálogo de características que pueden resultar de interés, de entre las cuales destacamos las recogidas por Fernández-Ballesteros (2004):

- *Frecuencia*: registrar si la conducta ocurre. Es un registro sencillo de obtener y que permite extraer varios datos de interés. Puede expresarse en forma de tasa, dividiendo el número de apariciones entre unidades de tiempo.
- *Duración*: tiempo en que aparece la conducta (unidades de tiempo transcurridas-segundos, minutos...-entre inicio y final). También puede referirse al tiempo en que la conducta ocurre, o al intervalo transcurrido entre distintas emisiones de la conducta, o entre la conducta y un determinado evento (ya sea precedente o consecuente).

- *Intensidad*: algunas conductas pueden medirse usando unidades continuas (metros recorridos, tiempo utilizado, masa de alimento ingerido, etc.), o bien pueden registrarse en una escala ordinal (sonrisa, risa, hilaridad... son distintos niveles de expresión de la alegría).
- *Orden*: sucesión en el tiempo entre distintas conductas, o entre ciertos estímulos y conductas (precedentes o consecuentes). Álvarez usa el término *ritmo* para hablar de los ciclos establecidos en estos intercambios. Cuando nuestro objeto de estudio necesita del estudio de la contingencia o relación entre conductas, estamos ante un estudio secuencial, lo que conlleva una serie de especificidades que serán discutidas en el apartado 4.3.

El tipo de muestreo temporal que vamos a realizar es una de las decisiones cruciales en el diseño de una investigación observacional. Quera (1989) describe que los registros que podemos llevar a cabo pueden dividirse en continuos o discretos (intermitentes). En los registros continuos, las anotaciones se llevan a cabo cuando aparece

alguna de las conductas contempladas por el sistema de categorías mientras que, en los registros intermitentes, las anotaciones se realizan en función del tiempo transcurrido (por ejemplo, se toman notas a cada minuto). Para este autor, el modo más correcto de nombrar a estas formas de anotación es, respectivamente, *Registro Activado por Transiciones* (RAT) y *Registro Activado por Unidad de Tiempo* (RAUT).

Dentro del RAUT, podemos decidir anotar el número de veces que ocurre la conducta en cada intervalo o si esta acción tiene lugar o no en este tiempo, mediante 1 y 0, respectivamente (*muestreo de Hansen*). Aunque lo más usual es encontrar trabajos basados en registros de 0 y 1, podemos anotar la frecuencia real si nuestros objetivos de investigación lo hacen indispensable. Las cuatro opciones de muestreo temporal se muestran en la tabla 4.

(ver página siguiente)

| Nombre | Registrar | Ejemplo |
|---------------|---|---|
| RAT | Cada vez que aparezca la conducta (sin importar el momento) | Cada vez que un niño comience a jugar, se registra |
| RAUT-A | Se anota si al pasar el umbral temporal la conducta está teniendo lugar | Si al pasar el minuto 1 el niño juega, se anota; vuelve a comprobarse a las 120 segundos, a los 180, etc. |
| RAUT-B | Se anota si la conducta tiene lugar en cualquier momento dentro del intervalo | Si entre el minuto 1 y 2 el niño juega en algún momento, se anota en el intervalo 1-2min. Aunque juegue dos o más veces, sólo se registra una vez. |
| RAUT-C | Se anota si la conducta tiene lugar desde el principio y hasta el final del intervalo | Si el niño comienza a jugar en el minuto 1:50 y sigue jugando hasta el minuto 3:20, se registra la ocurrencia del intervalo 2-3min. Si se interrumpe el juego dentro de este minuto, aunque sólo sea un momento, no se anota. |

Tabla 4. Modalidades de muestreo, según temporalidad.

El RAT tiene la ventaja de ser exhaustivo, aunque el tiempo que dedicamos a anotar la conducta puede conllevar pérdidas de información. Los RAUT tienen la ventaja de ser temporalmente económicos a la vez que sistemáticos, si bien cada uno de sus tipos presentan distintas características: el RAUT-C tiende a ofrecer resultados más discretos, ya que exige una actividad ininterrumpida durante todo el intervalo (es útil cuando observamos conductas prolongadas en el tiempo); el RAUT-B ofrece los resultados más numerosos, ya que sólo necesita una ocurrencia dentro del intervalo para registrarlo como positivo; por último, el RAUT-A es el más económico, ya que sólo exige atención durante unos instantes (cuando se acerca el momento de cambiar de intervalo) y ofrece un número de positivos intermedio entre los anteriores.

Por supuesto, hay otras decisiones que tomar en torno a las unidades de medida (tabla 5).

(ver página siguiente)

| Variable | Decisiones |
|------------------|---|
| Tiempo | Nº de sesiones a realizar Tiempo de observación por sesión Duración de intervalos temporales (si es un RAUT) |
| Situación | Si se supone que el ambiente influye sobre la conducta a estudiar, hay que decidir en qué contextos se realizará el registro |
| Sujetos | Selección de sujetos concretos a observar Si observamos grupos, puede programarse una rotación entre distintos sujetos en función del tiempo (5 minutos para el sujeto A, 5 minutos para el sujeto B, etc.) Si se hace rotación, el tiempo de observación debe ser igual para todos los sujetos |

Tabla 5. Otras características típicas del diseño de medidas para la observación.

4.3. Diseñar el formato de registro conductual

Hay ciertos datos que todo registro conductual debe poseer. De un modo genérico, debemos incluir campos para consignar:

- Fecha de la sesión y horas de inicio y final.
- Observador que realiza el registro.

- Lugar o contexto de observación.
- Sujeto o sujetos a los que se observan (si estudiamos sujetos concretos).
- Área para notas, por si el observador estima oportuno hacer constar alguna circunstancia específica.

El diseño del registro depende del tipo de dato que queramos obtener. Existe una gran variedad de combinaciones derivadas de mezclar distintos valores de: número de sujetos (idiográfico o nomotético), número de sesiones (observación puntual o de seguimiento), número de variables (unidimensional o multidimensional) (Anguera, y cols. 2001), relación entre eventos registrados (emisiones aisladas o secuencias), etc. A continuación se proponen algunos ejemplos de registro.

Registro de conductas no secuenciales (/aisladas): Es la rejilla de uso más frecuente. En ella, disponemos una serie de categorías de conducta en las filas y los intervalos de observación en las columnas. Es la herramienta de registro, por lo que debe ser funcional; sólo aparecen los nombres o abreviaturas de las categorías, mientras que los indicadores y ejemplos quedan fuera (pueden añadirse en otra página por si

surgen dudas, pero el observador debe conocer la herramienta lo bastante bien como para no necesitarlos). Vemos un ejemplo prototípico para la observación de la conducta de un niño en el patio de recreo (cuadro 8).

| Intervalos | | | | | | | | | |
|-------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|----------|
| Conductas | 0-1' | 1-2' | 2-3' | 3-4' | 4-5' | 5-6' | 6-7' | 7-8' | ∴ |
| Juego solitario | | | | | | | | | |
| Juego social | | | | | | | | | |
| Habla solo | | | | | | | | | |
| Habla niños | | | | | | | | | |
| Habla adultos | | | | | | | | | |
| Lectura | | | | | | | | | |
| ... | | | | | | | | | |

Cuadro 8. Ejemplo de registro conductual para la conducta de juego.

Registro de conductas secuenciales (/interactivas): en este tipo de rejilla, el interés se centra en conocer la relación entre conductas (cuál sigue a cuál), y no tanto en el momento en que cada una aparece. Por ello, en vez de ocupar las columnas con intervalos temporales, vamos a colocar otras conductas.

Cuando la interacción se dé entre dos sujetos (no necesariamente humanos), podemos asignar a uno de ellos las filas y al otro las columnas. En Escudero, Rogers y Gutiérrez (2001) encontramos un ejemplo de rejilla para registrar los diálogos dentro de la pareja, según si antecedentes y consecuentes implican dominación, sumisión o nivelación (cuadro 9).

| | | Conducta consecuente | | |
|-------------------------|------------|----------------------|----------|------------|
| | | Dominación | Sumisión | Nivelación |
| Conducta antecedente | Dominación | | | |
| | Sumisión | | | |
| | Nivelación | | | |

Cuadro 9. Ejemplo de registro para conductas secuenciales.
Adaptado de Escudero, Rogers y Gutiérrez (2001).

Mediante esta tabla, logramos registrar datos sobre a qué estilo de conducta sigue a ciertos comportamientos. Sin embargo, con esta propuesta desconocemos qué miembro de la pareja realiza cada una de las acciones, y por razones evidentes, conocer un posible desnivel de poder es potencialmente relevante. Esto puede paliarse introduciendo una sencilla variación en la rejilla (cuadro 10).

| | | Sujeto 2 | | |
|----------|------------|------------|----------|------------|
| | | Dominación | Sumisión | Nivelación |
| Sujeto 1 | Dominación | | | |
| | Sumisión | | | |
| | Nivelación | | | |

Cuadro 10. Modificación para diferenciar el inicio del intercambio de la respuesta.

La zona de la celda en que se registra el evento nos informa de quién emitió en primer lugar y quién en el segundo: cuando el sujeto 1 emita en primer lugar, anotaremos en la zona inferior, mientras que cuando sea el sujeto 2 quien emita y el 1 el que responda, lo anotaremos en la zona superior de la celda.

Las rejillas mostradas son cuadradas y simétricas, puesto que ambos interlocutores pueden emitir todos los tipos de conducta. Pero si existen asimetrías, como las provocadas por discapacidades (en las que uno de los sujetos está imposibilitado para parte de las conductas, como un invidente para la categoría *mantener la mirada*) o la interacción entre humanos y otros animales, las tablas de

registro pueden adaptarse y contemplar distinto número y/o tipos de categorías. Podemos ejemplificarlo mediante la interacción entre estado emocional de un paciente en cuidados paliativos y el comportamiento de un animal de compañía (se añaden las diagonales para poder controlar antecedentes y consecuentes de cada contingencia-ver cuadro 11).

| | | Perro | | | |
|----------|------------------|-----------------|-----------------|-------------|----|
| | | Mantiene mirada | Contacto físico | Tumba suelo | .. |
| Paciente | Muestra alegría | | | | |
| | Expresión neutra | | | | |
| | Muestra enfado | | | | |
| | (...) | | | | |

Cuadro 11. Ejemplo de rejilla asimétrica para registrar interacciones, separando el inicio del intercambio de las respuestas.

También es posible registrar interacciones en grupos con más de dos sujetos. Uno de los ejemplos clásicos es el *Interaction Process Analysis* (IPA) de Bales (1950), en que tanto filas como columnas representan sujetos (y no conductas ni intervalos). La tabla es cuadrada, de modo que cada uno de los sujetos tiene una celda en la que coincide con cada uno de los demás.

El formato del IPA ha sido usado en gran cantidad de estudios, y distintas variantes pueden encontrarse en la literatura actual. La situación base consiste en reunir a un grupo de desconocidos para que busquen solución a un problema propuesto; las respuestas se clasifican en 12 categorías, cuyas referencias (de 1 a 12) son las que se anotan en la rejilla. En este tipo de rejilla, también es posible anotar quién emite la conducta y quién responde a ella, de un modo similar al anteriormente expuesto.

La estructura básica de una rejilla para registrar interacciones en grupos se muestra en el cuadro 12 (usando como código números: 1=Pedir ayuda, 2=Aportar soluciones, etc.)

| | | Receptor | | | | | | |
|--------|------|----------|------|------|------|------|------|-----|
| | | Suj1 | Suj2 | Suj3 | Suj4 | Suj5 | Suj6 | ... |
| Emisor | Suj1 | | | | | | | |
| | Suj2 | | | | | | | |
| | Suj3 | | | | | | | |
| | Suj4 | | | | | | | |
| | Suj5 | | | | | | | |
| | Suj6 | | | | | | | |
| | ... | | | | | | | |

Cuadro 12. Ejemplo de rejilla de interacción para grupos.

Mapa de conducta: Es una herramienta que sirve para registrar la conducta espacial del sujeto o grupo observado. En ella se incluye una representación gráfica del contexto de observación y algún tipo de registro conductual, ya sea en forma de tabla (en el ejemplo) o de resumen de las observaciones (con la frecuencia o índice observado). En el mapa adjunto (gráfico 2) nos permite registrar las conductas llevadas a cabo por un animal en cautividad en las distintas zonas de su recinto.

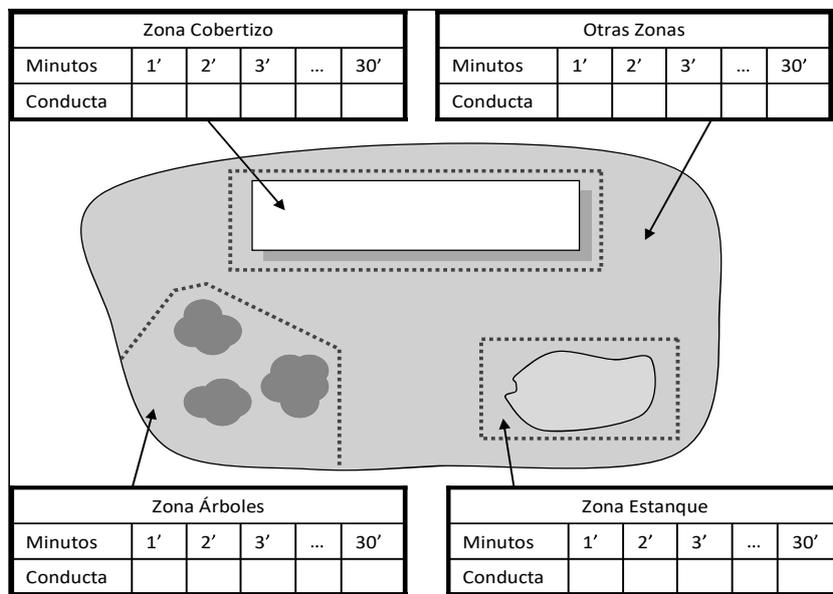


Gráfico 2. Ejemplo de mapa de registro conductual.

4.4. Análisis de la calidad del dato

En el esquema de investigación presentado al principio del capítulo, el momento de la recogida de datos aparece como previo al análisis de los mismos. Sin embargo, ya se comentó que esta idea necesita ser matizada. El diseño de las herramientas para el registro es circular, en el sentido de que un sistema de categorías no puede darse por definitivo hasta que se comprueba su validez y fiabilidad. Una vez contrastadas estas características, podemos dar por bueno el

sistema de registro y analizar los resultados, pero en caso contrario sería necesario modificar la herramienta de evaluación.

El concepto de *validez* presenta dos acepciones, según Anguera (1985):

1. *Validez interna o de contenido*. Hace referencia a si la herramienta de registro evalúa lo que pretende evaluar, o si es representativa del fenómeno estudiado (León y Montero, 2003). En este sentido, cuanto mayor sea el grado de inferencia que asumamos, mayores dudas surgirán acerca de si estamos midiendo lo que esperamos medir. Cuando no tenemos elementos de juicio para determinar si una evaluación es válida en este sentido, podemos comparar nuestros resultados con otros instrumentos contrastados (por ejemplo, con los datos obtenidos usando otros registros conductuales, tests, etc.; ver cuadro 13). Cuando las herramientas de evaluación poseen suficiente validez interna, la fiabilidad tiende a ser elevada.

Se intenta evaluar el grado de *depresión* en un paciente. Para ello, determinamos varias categorías que, consideramos, sirven para detectarla

- Salir de su cuarto
- Interactuar con otras personas
- Participar en actividades de ocio
- Verbalizaciones con tono emocional positivo
- Verbalizaciones con tono emocional negativo

¿Estas categorías son suficientes para detectar la depresión?
= **validez interna**

Comparar resultados de nuestra evaluación con otra herramienta de evaluación más contrastada, como el BDI (Beck, y cols., 1961; ver capítulo 4)
= **validez concurrente**

Cuadro 13. Ejemplo de evaluación de la validez interna.

2. *Validez predictora*. Hace referencia a la utilidad de los datos que nos ofrece nuestra herramienta de evaluación de cara a afrontar retos futuros. Por ejemplo, es conocido que los seres vivos con un nivel de activación fisiológica media rinden mejor durante una situación de aprendizaje (como un perro durante el adiestramiento) que los que presentan niveles muy bajos o muy altos, y podemos proponer un sistema de categorías que nos ayude a clasificar a los perros en

función de su nivel de activación fisiológica. Posteriormente, podemos comprobar el desempeño de estos perros durante el adiestramiento, de modo que se puede poner en relación con la evaluación inicial; si los resultados son los esperados, podemos considerar que la evaluación ha sido correcta.

Por otra parte, tenemos la *fiabilidad* (también denominada *confiabilidad*). Ésta viene definida como el grado de coincidencia entre varias medidas, bien sea entre los registros de distintos jueces (*fiabilidad entre jueces* o *interjueces*), o bien por un mismo juez en distintos momentos, como cuando se visiona un fragmento de sesión en vídeo (*fiabilidad intrajuez*).

La fiabilidad puede aumentarse mejorando la herramienta de evaluación (por ejemplo, eliminando ambigüedades en el momento de codificar) y con un entrenamiento exhaustivo de los observadores.

El modo más sencillo de obtener una medida de fiabilidad consiste en el *índice de acuerdo* (o *tasa*; cuadro 14). Normalmente, obtener un valor por encima de 0,8 (esto es, el 80%) es suficiente para dar por fiable la medida.

| |
|---|
| $\text{Índice de acuerdos} = \frac{\text{N}^\circ \text{ acuerdos}}{\text{N}^\circ \text{ acuerdos} + \text{N}^\circ \text{ desacuerdos}}$ |
| $\text{Tasa de acuerdos} = \frac{\text{N}^\circ \text{ acuerdos}}{\text{N}^\circ \text{ acuerdos} + \text{N}^\circ \text{ desacuerdos}} \times 100$ |

Cuadro 14. Cálculo de índice y tasa de acuerdo entre jueces.

Estos índices pueden calcularse para el total del registro, o bien por separado para cada categoría de conducta (útil para comprobar si alguna categoría ofrece mayor ambigüedad que el resto).

Sin embargo, estos índices no tienen en cuenta el efecto que el azar puede tener sobre las coincidencias y divergencias entre jueces. Por ello, cuando trabajemos con sistemas de categorías en los que las conductas codificadas sean mutuamente excluyentes, y en los que contemos con al menos tres categorías, podemos usar el coeficiente Kappa de Cohen (Anguera, 1990; Cohen, 1960; León y Montero, 2003; ver cuadro 15). El índice Kappa se da por válido cuando supera el valor de 0,70 (Bakeman y Gottman, 1989).

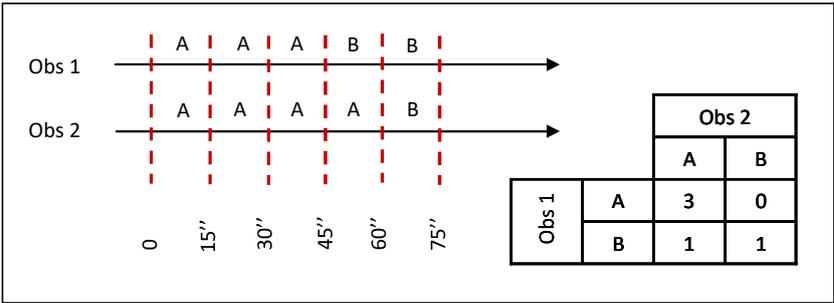
$$K = \frac{\text{Índice de acuerdo} - \text{proporción de acuerdos al azar}}{1 - \text{proporción de acuerdos al azar}}$$

$$\text{Índice de acuerdo} = \frac{\text{Nº acuerdos}}{\text{Nº acuerdos} + \text{Nº desacuerdos}}$$

$$\text{Proporción de acuerdos al azar} = \sum [(\text{proporción de ocurrencia de la cat. } i \text{ para el observador 1}) * (\text{proporción de ocurrencia de la cat. } i \text{ para el obs. 2})]$$

Cuadro 15. Cálculo del índice Kappa de Cohen.

El índice Kappa puede utilizarse también para calcular el grado de concordancia entre registros cuando estos contemplan la duración de los mismos. Basta con dividir el tiempo en unidades fijas (como minutos o segundos), y comprobar las coincidencias frente a las diferencias de registro (ver cuadro 16).



Cuadro 16. División de la línea temporal en intervalos para calcular el índice Kappa.

Este proceso, siempre que se realice de manera sistemática, es de sencilla resolución. A continuación (cuadro 17) vemos un ejemplo del cálculo de estos índices.

En el patio de un colegio, dos observadores registran la conducta de un niño en particular a lo largo de 10 intervalos de 30 segundos. Los resultados son:

| Intervalo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|---|---|---|---|---|---|---|---|---|----|
| Obs. 1 | A | C | C | D | D | C | B | A | C | A |
| Obs. 2 | A | C | B | D | D | C | B | A | B | A |

[A=Juego en solitario; B=Juego en grupo; C= Habla con otros niños; D= Habla con un adulto]

Acuerdo interjueces = $(8 / 8 + 2) = 0,80$

Preparamos una matriz para conocer las coincidencias y calcular Kappa.

| | | Observador 2 | | | | |
|--------------|-------|--------------|------|------|------|-------|
| | | A | B | C | D | Total |
| Observador 1 | A | 3 | 0 | 0 | 0 | 3/10 |
| | B | 0 | 1 | 0 | 0 | 1/10 |
| | C | 0 | 2 | 2 | 0 | 4/10 |
| | D | 0 | 0 | 0 | 2 | 2/10 |
| | Total | 3/10 | 3/10 | 2/10 | 2/10 | |

Prop. acuerdos por azar = $[(3/10)*(3/10) + (1/10)*(3/10) + (4/10)*(2/10) + (2/10)*(2/10)] =$
 $= [(9/100) + (3/100) + (8/100) + (4/100)] = (24/100) = 0,24$

Kappa = $(0,80 - 0,24) / (1 - 0,24) = (0,56) / (0,76) = (0'737)$

Cuadro 17. Ejemplo de cálculo del índice Kappa de Cohen.

La fiabilidad puede también calcularse para más de dos jueces. Una forma posible para ello consiste en realizar el cálculo de la fiabilidad interjueces en parejas y hallar la fiabilidad promedio; por ejemplo, si trabajamos con tres observadores (A, B, y C), necesitaremos calcular el coeficiente Kappa en tres ocasiones (parejas AB, AC y BC), sumar sus resultados y dividir entre 3 (el número de observadores).

Otra opción consiste en utilizar el coeficiente de concordancia W de Kendall, que es un coeficiente de correlación (y por lo tanto, sus valores absolutos oscilan entre 0, acuerdo nulo, a 1, acuerdo total). Sin embargo, esta fórmula puede necesitar de la estimación de una constante de ponderación, lo que hace que no siempre sea sencilla de aplicar. La fórmula para su cálculo es la siguiente:

$$W = \frac{12 \times S}{O^2 \times K \times (K - 1) \times C}$$

Donde:

$S = \sum [(suma \ de \ registros \ por \ categoría) - (media \ de \ registros \ para \ todas \ las \ categorías)]^2$

O = Número de observadores

K = Número de categorías de conducta utilizadas

C = Constante de ponderación

Cuadro 18. Cálculo del coeficiente de concordancia de Kendall.

Existen otras fórmulas posibles para calcular el grado de fiabilidad a través de varias medidas, de uso menos frecuente, y que suelen depender del software informático utilizado para codificar los registros. Por ello hemos considerado no incluirlos en este capítulo.

4.5. Revisar el sistema de categorías

El último paso en el proceso de diseño de nuestro sistema de categorías es la comprobación de que el instrumento se ajusta a los criterios establecidos. La concordancia entre el objeto de estudio elegido y los datos obtenidos (validez interna o de contenido), así como el grado de fiabilidad entre jueces (influido por la exhaustividad, exclusividad y grado de explicitación de nuestras categorías mediante indicadores y ejemplos) son los indicadores de calidad que nos permiten decidir acerca de la adecuación de nuestra herramienta.

En el caso de que la validez y fiabilidad no sean adecuadas, debemos volver a recorrer los pasos descritos. A continuación se describen algunos de los problemas más frecuentes:

- Si hemos usado un criterio externo para comprobar la validez de nuestra observación y los resultados son dispares, hay que asegurarse de que el instrumento usado para comparar los resultados define las variables estudiadas del mismo modo que nosotros. En ocasiones, una misma etiqueta puede usarse con distintas connotaciones (*depresión, ansiedad, estrés...* pueden tener distintos significados para distintos autores).
- Si hemos comprobado la validez predictiva y los resultados no son los esperados, debemos tener en cuenta el número de sujetos observados, ya que los estudios nomotéticos ofrecen una mayor solidez que los idiográficos.
- En la validez predictiva, un control deficiente de variables influyentes puede aminorar la utilidad de nuestros datos. Es recomendable revisar el proceso desde el principio, sobre todo si se creó el sistema de categorías mediante una estrategia racional (partiendo de supuestos teóricos, en vez desde la observación no sistemática).

- Si la fiabilidad es baja por el desacuerdo entre los jueces, hay varias cuestiones a revisar: comprobar la correcta explicitación de las categorías a través de indicadores y ejemplos; comprobar si la falta de acuerdo es general o se refiere a unas categorías concretas; aumentar el entrenamiento de los observadores para minimizar los sesgos personales (usando material grabado, por ejemplo)...

Por supuesto, toda investigación es mejorable, y resulta deseable que los planteamientos utilizados se refinan conforme sea posible. No obstante, si las evidencias sobre la calidad de los datos son satisfactorias, podemos pasar al análisis de datos como tal para comprobar las hipótesis con que iniciamos la investigación.

5. REFERENCIAS

- Álvarez, A. (1975). *Manual de Psicología Experimental: Metodología*. Madrid: Biblioteca Nueva.
- Anguera, M. T. (1985). *Metodología de la Observación en las Ciencias Humanas*. Madrid: Cátedra.

- Anguera, M.T. (1990). Metodología observacional. En J. Arnau, M.T. Anguera y J.L. Gómez. *Metodología de la Investigación en Ciencias del Comportamiento* (pp. 125-238). Murcia: Secretariado de Publicaciones de la Universidad de Murcia.
- Anguera, M.T., Blanco, A. y Losada, J.L. (2001). Diseños observacionales, cuestión clave en el proceso de la metodología observacional. *Metodología de las Ciencias del Comportamiento*, 3, 135-160.
- Anguera, M. T. (2003). La observación. En Moreno, C. *Evaluación psicológica. Concepto, proceso y aplicación en las áreas del desarrollo y de la inteligencia* (pp. 271-307). Ed. Sanz y Torres.
- Arnau, J. (1995). Metodología de la investigación psicológica. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo. *Métodos de investigación en psicología. Síntesis psicología*.
- Bakeman, R. y Quera, V. (1996). *Análisis de la Interacción: Análisis Secuencial con SDIS y SGEQ*. Madrid: RA-MA.
- Bakeman, R. y Gottman, J.M. (1989). *Observación de la Interacción: Introducción al Análisis Secuencial*. Madrid: Morata.
- Bales, R.F. (1950). *Interaction Process Analysis*. Chicago: Chicago U. Press.
- Blanco, N. (2010). Metodología de investigación. En A. Blanco y L. Rodríguez-Franco (coords.). *Manual de investigación en ciencias de la salud. Guía práctica de iniciación* (pp. 53-68). Sevilla.
- Cohen, S. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological measurement*, 20, 37-46.

- Escudero, V., Rogers, E. y Gutiérrez, E. (2001). Observación de patrones de interacción de la pareja. En M.T. Anguera (coord.). *Observación en Psicología Clínica: Aplicaciones*. Barcelona: Edicions de la Universitat de Barcelona.
- Fernández-Ballesteros, R. (2004). La observación (pp. 161-201). En R. Fernández-Ballesteros (Dir.). *Evaluación Psicológica: Conceptos, métodos y estudios de caso*. Madrid: Pirámide.
- Jackson, D.D. Watzlawick, P. y Beavin, J.H. (1967). *Pragmatics of Human Communication*. New York: Norton & Company Inc.
- Kerlinger, F.N. y Lee, H.B. (2001). Observación del comportamiento y sociometría (pp. 661-685). En F.N. Kerlinger y H.B. Lee. *Métodos de Investigación en Ciencias Sociales*. (4ª ed.) México: McGraw-Hill.
- León, O. G. y Montero, I. (2003). *La observación*. En O.G. León e I. Montero. *Métodos de Investigación en Psicología y Educación*. (3ª ed.) Madrid: McGraw-Hill.
- Quera, V. (1989). Estimación de frecuencia y duración en el muestreo temporal de la conducta. *Anuario de Psicología*, 43, 33-62.