

LA CONVERGENCIA DE LO COGNITIVO Y LO SOCIAL EN LOS ERRORES HUMANOS*

Anna Estany
Universitat Autònoma de Barcelona
Anna.Estany@uab.cat

Abstract

The aim of this paper is to show that not only there is no incompatibility between the cognitive and social factors in the analysis of scientific and technological phenomena, but that they are complemented. This idea is exemplified with the study of the human error from a double perspective: the one of a cognitive scientist like James Reason and the other one of an engineer like Henry Petroski. The analysis of catastrophes because of failures in architectonic structures or industrial processes show that only considering the cognitive and social elements we can diminish the failures and increase the successes of the scientific and technological activity.

Key Words: Human error, engineering failures, design.

Resumen

El objetivo de este artículo es mostrar no sólo que no hay incompatibilidad entre los factores cognitivos y sociales en el análisis de fenómenos científicos y tecnológicos, sino que se complementan. Esta idea se ejemplifica con el estudio del error humano desde una doble perspectiva: la de un científico cognitivo como James Reason y la de un ingeniero como Henry Petroski. El análisis de catástrofes a causa de fallos en estructuras arquitectónicas o en procesos industriales muestra que sólo teniendo en cuenta sus elementos cognitivos y sociales podemos disminuir los fracasos y aumentar los éxitos de la actividad científico-tecnológica.

Palabras clave: Error humano, fracaso ingenieril, diseño.

*Recibido: 06-06-07. Aceptado: 25-06-08.

*Este artículo se ha realizado en el marco del proyecto de investigación “Cognición y entornos tecnológicos. El papel de las ciencias cognitivas en el diseño de artefactos tecnológicos” (HUM2005-01552/FISO) financiado por el Ministerio de Ciencia y Tecnología de España.

La introducción del elemento cognitivo en la filosofía de la tecnología no hay que considerarla independiente y desligada de los elementos contextuales (sociales, políticos, culturales, éticos) en los que se han centrado (por no decir, limitado) los estudios de Ciencia-Técnica-Sociedad (CTS). Una de las situaciones donde se muestra claramente la relación entre lo cognitivo y lo social es cuando se cometen errores en el curso de cualquier proceso tecnológico, provocando desgracias personales y sociales que, en muchos casos, se atribuyen a “errores humanos”. Por ello tomaré los errores humanos y los fracasos en la satisfacción de necesidades humanas como uno de los fenómenos donde se entrelazan aspectos cognitivos y sociales y, no solamente esto, sino que el abordarlos independientemente es, en muchas ocasiones, la causa del error y el fracaso.

El objetivo de este artículo es mostrar la interrelación de los elementos cognitivos y sociales a partir del fenómeno de la comisión de errores. El estudio de los errores es un tema muy amplio y que puede abordarse desde perspectivas muy diversas. Aquí me centraré en el factor humano desde una doble perspectiva: cognitiva e ingenieril.

Si actualmente se habla tanto de los problemas de la tecnología es porque, en el fondo, subyace la idea de que la relación intrínseca, que durante muchos siglos pero quizás de forma más persistente durante el siglo XIX, se pensó había entre progreso y desarrollo tecnológico, se rompió a finales del XX, y así sigue. Es necesario preguntarse por qué el futuro ha devenido un asunto controvertido interpretado en términos de riesgo y no de progreso (Lübbe 1993, 18). Para los estudios CTS éste ha sido su punto de mira y su objetivo principal, pero no son los únicos que han abordado la problemática. Petroski (1995, 2) señala que en los últimos años ha habido tantos accidentes y con un costo tan alto de vidas humanas, miseria, ansiedad, etcétera, que nos preguntamos dónde está el progreso tecnológico. Desde la filosofía estas cuestiones se han analizado en el marco de la filosofía de la tecnología¹. Estas cuestiones no pueden abordarse de forma unilateral, sino que necesitan un trata-

¹La lista sería larga pero, a modo de ejemplo, podemos citar en el ámbito hispánico, entre otros muchos, a M. Bunge (1966 y 1976), M. A. Quintanilla (1989) y J. Echeverría (2003).

miento multidisciplinar y multifactorial. Esto no significa que, luego, los estudios se centren en algún aspecto específico. Además, no sólo están focalizados en alguno de los factores sino que, en buena parte, son elegidos en función del desarrollo tecnológico de dichos campos y de las consecuencias sociales en poblaciones que tienen capacidad de decisión sobre los campos de investigación. Por ejemplo, las centrales nucleares (Chernobyl, 1986) y las industrias químicas (Bhopal, 1984) han sido objeto de estudio de errores humanos, buscando las causas y la forma de detectarlos y corregirlos. Otro tipo de catástrofes tiene que ver con la ingeniería de estructuras que tiene a su cargo la construcción de puentes, edificios, aviones, etc. (el hotel Hyatt de Kansas City, 1981; Mianus Bridge, 1967; DC-10 Chicago, 1979). Estos son sólo algunos de los ejemplos de accidentes que dieron la vuelta al mundo, en los que se presentó un cúmulo de errores con responsabilidades también a todos los niveles. Las centrales nucleares han sido claves para el mundo desarrollado, aunque ahora se cuestionen. Los países que las tienen y que las han promovido son los que tienen la capacidad política y económica para investigar sus fallos. En cuanto a los puentes y, en general, las obras de ingeniería han sido la base de la revolución industrial del primer mundo.

Nadie pone en duda las especiales circunstancias de nuestra época para justificar la percepción de riesgo y peligro de las innovaciones tecnológicas. Entre las nuevas situaciones de riesgo podemos señalar las siguientes por sus especiales características: los impactos dañinos no están limitados ni en el tiempo ni en el espacio; las causas y responsabilidades por los daños ocasionados no pueden atribuirse a decisiones individuales o a organizaciones sino a colectividades; los nuevos riesgos son abstractos. Esto significa que hay riesgos que no pueden experimentarse directamente, que no pueden calcularse económicamente ya que poseen su propia dinámica, y que el carácter abierto de la tecnología moderna hace posible reconocer sus riesgos y posibles peligros sólo cuando han sido introducidos (Lau 1989, 20-21).

Vemos que hay dos conceptos en torno a los cuales gira esta conexión entre lo cognitivo y lo social: error y diseño, frente a los que vamos a tomar dos puntos de vista: el del científico cognitivo, a través de Ja-

mes Reason (1990); y el del ingeniero, de la mano de Henry Petroski (1995), como marcos de referencia². Tanto el error como el diseño nos llevan a comprender mejor las distintas dimensiones del riesgo, mostrando la interrelación entre lo cognitivo y lo social. El estudio del riesgo es otro de los campos que han sido motivo de interés y preocupación por parte de la filosofía de la tecnología y se ha estudiado, fundamentalmente, desde la perspectiva sociológica y política. Aquí nos centraremos en las relaciones que el riesgo pueda tener con la comisión de errores y con el diseño en tanto que un mal diseño pueda ser la causa principal de errores humanos y, en consecuencia, aumentar el riesgo en determinadas tecnologías.

Luego veremos si otros marcos teóricos procedentes de otras disciplinas pueden contribuir a resolver algunos de los problemas planteados por Reason y Petroski.

1. La naturaleza del error

Si partimos de que los errores humanos constituyen un factor causal de las catástrofes provocadas por los sistemas tecnológicos es fundamental el estudio de su naturaleza. Esto es posible porque, desde el punto de vista psicológico, los errores no son tantos ni tan distintos y, por tanto, es posible su sistematización y generalización y, en consecuencia, su predicción y prevención, aunque no de forma determinista. La predicción de los errores depende, en gran medida, de nuestra comprensión de los factores que los ocasionan. Para ello necesitamos una teoría que articule los tres aspectos de la producción de errores: las características de la tarea, los mecanismos que rigen la actividad y las particularidades del sujeto.

El “error”, en sentido genérico, cubre todos los casos donde una secuencia planificada de actividades mentales o físicas no consigue los fines deseados, y cuando los fracasos no pueden ser atribuidos a la in-

²Quiero señalar que el tomar como referencia a estos autores no significa que sean los únicos sino que son lo suficientemente relevantes para el análisis del tema propuesto en este artículo.

tervención del azar. Siguiendo a Reason (1990, 9) veamos los tipos de errores:

- “Deslices” (*slips*) y “lapsos” (*lapses*) son errores que resultan de un defecto en la ejecución y/o en el almacenaje de una secuencia de acciones, independientemente de la adecuación del plan que les guía a su objetivo. La diferencia entre deslices y lapsos es que mientras los primeros son potencialmente observables (en el habla, la redacción o la acción), los segundos responden a formas de errores menos observables causados, principalmente, por defectos de memoria.
- “Equivocaciones” (*mistakes*) son deficiencias o defectos en el proceso de decisión y/o inferencia, que están implicadas en la selección de un objetivo o en la especificación de los medios para alcanzarlo, independientemente del hecho de que las acciones basadas en este esquema de decisión se desarrollen (o no) conforme al plan previsto.

Esta tipología de errores hay que entenderla como una clasificación desde el punto de vista psicológico y en función de los procesos cognitivos que tienen lugar en una acción cualquiera. Esto es importante señalarlo porque si partimos de un contexto concreto, seguramente podríamos caracterizar muchos y muy diversos tipos de errores. Por ejemplo, como veremos más adelante, en los tipos de fallos en la producción hay distintos tipos de errores que se cometen en las diversas fases de la producción, pero, desde el punto de vista psicológico, podemos encajarlos en alguno de los tipos señalados por Reason.

Estos tipos de error están ligados a las diversas etapas que van desde la concepción de una acción, con la intención de alcanzar un fin determinado, a la realización de la misma. Durante la planificación se suelen cometer equivocaciones, en el almacenaje de información lapsos y en la ejecución los errores más comunes son los deslices.

2. Errores y memoria

Los mecanismos de recuperación de la memoria están íntimamente relacionados con los tipos de error. Así, los deslices y los lapsos corresponden al nivel de actividad de los automatismos, es decir, se producen por fallos en la recuperación de información basada en acciones automáticas (conducir cuando se lleva años haciéndolo). Entre las equivocaciones podemos distinguir aquellas que corresponden al nivel de actividad basado en reglas (poner en funcionamiento el aparato de video) y las que corresponden al nivel de actividad basado en los conocimientos (recurrir a los conocimientos generales que tenemos sobre cocina cuando no logramos que nos salga bien un plato).

Toda tentativa de modelizar los fundamentos de la cognición humana debe abordar dos cuestiones: las propiedades de lo que constituye la base de los conocimientos y sus modos de representación; y el conjunto de reglas heurísticas para la selección de la estructura de los conocimientos almacenados que serán activados en cada una de las situaciones posibles.

Este mecanismo de selección de respuesta proporciona al modelo no solamente las características del tratamiento de la información sino que le permite también determinar las formas de error humano. En este modelo ideal, cada problema específico debería permitir encontrar la solución más apropiada de todas las que tenemos almacenadas. Pero la realidad está generalmente alejada de este ideal y comporta dos fuentes fundamentales de confusión. Una consiste en que las condiciones de llamada que están asociadas a un problema específico correspondan a muchas estructuras almacenadas, o a ninguna. También puede ocurrir que estas estructuras estén degradadas (borrosas) o ausentes. La otra se basa en que las estructuras del conocimiento pueden ser incompletas (algunos hechos son desconocidos) o falsas. Por lo que respecta al sistema cognitivo, estos dos tipos de confusión son funcionalmente equivalentes. Representan dos tipos de subespecificación cognitiva y crean las condiciones en las que los sesgos de similitud y frecuencia tienen muchas posibilidades de manifestarse y, en consecuencia, de ser fuente de errores.

3. Errores y atención

La falta de atención también juega un papel importante en la comisión de errores, muy especialmente en los deslices y lapsos. Esto es así por las características de la cognición humana. Rouse (1981, 104) dice que los seres humanos, si se les deja escoger, prefieren comportarse por reconocimiento de configuraciones específicas del contexto más que intentar contabilizar y optimizar.

La hipótesis central del *sistema genérico de modelización de los errores* (GEMS) es que los seres humanos, cuando se enfrentan a un problema, presentan un sesgo de resistencia: buscan y encuentran una solución pre programada al nivel de las reglas, antes de volver al nivel de los conocimientos, que es necesariamente más costoso, incluso cuando este último es requerido desde el principio.

El nivel de los automatismos se pone en marcha cuando las actividades se desarrollan en circunstancias familiares. Se recurre al nivel de las reglas cuando se pone en marcha la atención, en el curso de una actividad, una vez se ha comprobado que ha habido cambios respecto a la planificación prevista. El paso del nivel de las reglas al del conocimiento se produce cuando el sujeto se da cuenta que ninguna solución de su repertorio de reglas es adecuado para resolver el problema. El fenómeno de recurrir, sólo en última instancia, al conocimiento, se conoce como “conservadurismo cognitivo” (Luchins y Luchins 1950, 120) “el hábito manda sobre el sujeto en lugar del sujeto sobre el hábito”.

Si el estudio del error humano tiene que contribuir de forma útil a la seguridad y a la eficacia de las tecnologías complejas e incluso peligrosas, debe proporcionar a los que diseñan la tecnología, principios generales que sean aplicables a las características del sistema de tratamiento de la información que poseen los humanos. En este sentido es relevante saber que el sistema cognitivo tiende a generar las respuestas más familiares, más frecuentes y más contextualizadas. Todo ello lleva a la comisión de errores en cuanto hay datos incompletos o ambiguos, conocimientos incompletos o incorrectos, olvidos de la memoria prospectiva, sobrecarga de la capacidad del espacio de trabajo consciente, limitaciones de la intención y fallos normales o de origen patológico en la atención. Habría

que señalar que la asignación de acciones a la categoría de automatismos, reglas y conocimiento es relativa a la situación de la persona que comete el error. Por ejemplo, leer y conducir al principio son actividades basadas en reglas pero luego se convierten en automatismos.

4. Detección y corrección de errores

El estudio de los mecanismos cognitivos del conocimiento es relevante porque nos permite conocer la probabilidad de cometer errores y, en consecuencia, prevenirlos y evitarlos en la medida de lo posible. Sin embargo, dado que no estamos en un sistema determinista, los errores no se eliminarán completamente. Aunque sea una perogrullada, no está de más señalar que la tecnología es un producto humano, diseñada, dirigida y manipulada por humanos, es decir, por máquinas potentes pero falibles (así son nuestros cerebros). El estudio del funcionamiento de esta máquina falible es una cuestión cognoscitiva pero la decisión de adaptar la tecnología y el proceso de producción a dicha máquina es una cuestión de política empresarial. En la toma de decisiones habría que plantearse el riesgo que se está dispuesto a correr, que estará en función de qué es lo que va a priorizarse (seguridad, costos, impacto ambiental, etc.). Esta cuestión es estrictamente social pero sería una irresponsabilidad moral no tener en cuenta lo que la ciencia cognitiva nos dice sobre la naturaleza psicológica de los errores. Por tanto, es importante estudiar los mecanismos de detección y corrección de errores a fin de mitigar en lo posible sus consecuencias.

Entre los modos de detección de errores podemos señalar el auto-control consciente, la información del entorno y la detección de errores por terceros. El primero está relacionado con la atención. El segundo y el tercero son muy importantes en sistemas complejos y, de hecho, es lo que suele ocurrir en las centrales nucleares y en las industrias químicas. Por ejemplo, muchos de los errores de los operarios suelen ser detectados por otros que no estaban realizando la actividad donde se ha cometido el error. El poner señales que faciliten la información y que sean una llamada para que no se cometa un lapso o un desliz depende,

en buena parte, de los diseñadores y de las autoridades políticas, económicas y empresariales.

Tanto la detección como la corrección de errores son cruciales para evitar catástrofes. En este sentido es muy importante la distinción entre errores activos y errores latentes. Los primeros están asociados a la actividad de los operarios que están “en primera línea” de un sistema complejo (pilotos, controladores de navegación aérea, los equipos de la sala de control, etc.). En el caso de Chernobyl serían los operarios que estaban al frente del reactor que explotó. Los segundos son los que se producen en las actividades de los que están alejados del control directo, tanto en el tiempo como en el espacio (los diseñadores, los dirigentes de alto nivel, los obreros que construyen los sistemas, los directores y personal de mantenimiento, etc.). En Chernobyl sería el ingeniero que diseñó el reactor y en Kansas el arquitecto responsable de la construcción del hotel Hyatt. La tesis de Reason es que detrás de todas las catástrofes hay algún error latente, aunque haya un error activo que ha funcionado como detonante. Además los errores cometidos por individuos alejados del control directo (y eventualmente de los riesgos directos) suelen tener consecuencias potencialmente más peligrosas para el sistema que los cometidos por los que están en “primera línea”. Entre los casos de catástrofes podemos señalar algunos en los que convergen una serie de errores latentes: Three mile island, 1979; Bhopal, 1984; Challenger, 1986; Chernobyl, 1986; Zeebrugge, 1987.

Sin embargo, la realidad es que la evolución de las tecnologías en las últimas décadas nos ha llevado a sistemas cada vez más automatizados en los que los operarios están cada vez más alejados de los procesos que deben controlar, a veces necesariamente, ya que los procesos son demasiado peligrosos para ser controlados directamente. Frente a ello se han desarrollado Dispositivos de Seguridad Automática (DSA). Por ejemplo, en las centrales nucleares modernas se ha instalado lo que se llama “protecciones en serie”, en el sentido de que frente a cualquier problema hay dispositivos que activan medidas de seguridad. A pesar de ello, continúan produciéndose accidentes. Una de las razones es que los sistemas de seguridad son presa de los errores humanos, en especial de los

errores latentes. Nos encontramos con la paradoja de que estos sistemas concebidos especialmente para garantizar, ellos solos, la seguridad de la instalación son también su “talón de Aquiles”. Son, lo que podríamos llamar, “ironías de la automatización”. En principio los diseñadores de sistemas dejan para los operarios (actores humanos) lo que no ha podido ser automatizado. Lo que se les pide es que se aseguren que el sistema automático funcione correctamente. Pero se sabe que incluso los operarios más responsables no pueden mantener un nivel de vigilancia eficaz más allá de periodos de tiempo bastante cortos. Para resolver este problema los diseñadores dotan al sistema de signos automáticos de alarmas (información del entorno).

Otra de las funciones de los operarios es volver al sistema manual cuando el automático falla, pero resulta que los operarios tienen muy poca experiencia en resolver problemas manualmente, con lo cual sólo operarios muy cualificados podrán hacerlo. Brainbridge (1987, 278) dice:

[S]in duda el colmo de la ironía es que los sistemas en los que más se ha logrado la automatización, y en los que raramente se necesita la intervención manual, son también los que necesitan una mayor inversión en la formación de los operarios.

Sólo sacrificando la seguridad, la automatización significaría disminución de gastos y, por tanto, más beneficios económicos.

5. Fallos en la producción

A partir del estudio de casos Reason propone una concepción global de la causalidad de los accidentes en los sistemas complejos en la que entran en juego los diversos elementos que intervienen en la producción. A cada uno de estos elementos les corresponden unos errores o fallos posibles debido a la intervención humana (Reason 1990, 275).

1. *Los que toman decisiones*: arquitectos y gestores de alto nivel. Una gran parte de sus funciones consiste en asignar los recursos, siempre limitados, a los distintos departamentos: renovación de maquinaria,

personal, seguridad, publicidad, etc. En esta etapa la mayoría de los accidentes son fallos en la toma de decisiones. Incluso en las empresas mejor dirigidas encontramos fallos, y éstos proceden, en mayor medida, de la dificultad de hacer compatible la producción y la seguridad (Reason 1990, 277). En este punto el factor económico es primordial.

Los fallos en esta etapa corresponderían a equivocaciones y, probablemente, estarían basadas en conocimiento más que en reglas. Sin embargo, precisamente la reflexión de Reason sobre la dificultad de compatibilizar producción y seguridad, me hace pensar que una parte de los fallos no son debidos tanto a errores humanos sino al haber aceptado un grado de riesgo superior al que la prudencia hacía aconsejable. Aquí es donde las cuestiones sociales, políticas y éticas intervienen y donde la mayoría de los estudios CTS han hecho aportaciones cuando han estudiado los riesgos³.

2. *La gestión de la producción*: consiste en poner en marcha las estrategias de los que han tomado las decisiones (formación, ventas, mantenimiento, seguridad, etc.). En este nivel habría que decir que la gestión de la producción puede tanto exacerbar como atenuar los fallos de la toma de decisiones. A veces, los operarios con experiencia pueden evitar accidentes, del mismo modo que los novatos pueden provocarlos o, al menos, no evitarlos. Habría que estudiar hasta qué punto las prejubilaciones de las empresas para poner, de golpe, personas sin experiencia y así pagar sueldos más bajos, constituyen un peligro añadido. Éste es el caso Bhopal (India) que en un plan de viabilidad, debido a un descenso de la producción, despidieron a muchos trabajadores con mucha experiencia y contrataron personas jóvenes, sin ninguna experiencia pero con sueldos bajos.

Los fallos en este nivel de actividad están más directamente relacionados con estados psicológicos y capacidades cognitivas, aunque las decisiones de carácter estrictamente empresarial pueden aumentar o disminuir el riesgo.

³López Cerezo y Luján (2000), Vallverdú (2002) y Shrader-Frechette (2005). Estos autores son sólo una muestra del análisis de riesgos.

Para este punto tienen especial relevancia algunos factores que concurrieron en el accidente de Chernobyl. El modelo elegido por los ingenieros para el reactor 4 era del tipo RBMK, cuya característica es que debe funcionar siempre a plena potencia ya que por debajo de 700 megavatios el reactor se vuelve inestable. Lo más seguro es no dejarlo bajar de 1000 megavatios. Cuando un reactor se vuelve inestable la solución es insertar en el núcleo unas barras (barras de control) de material inerte que absorben los neutrones sobrantes. El problema es que, por un lado, las reacciones se descontrolan en menos de 2 segundos pero, por otro, el diseño del RBMK exige más de 20 segundos en insertar las barras de control en el uranio. Estas características del diseño del reactor RBMK requieren que no se opere nunca por debajo de 1000 megavatios y tener siempre en el núcleo 30 barras de control para asegurar que en el caso de inestabilidad del reactor el uranio pueda ser absorbido y no producirse ningún accidente. En este caso, los operarios violaron todas las reglas: el reactor funcionaba a baja potencia y no había ninguna barra de control en el núcleo del reactor. La pregunta que se plantea inmediatamente es por qué violaron las reglas, ¿fue por falta de formación?, ¿por órdenes incorrectas? Es difícil pensar que fue por desidia suya ya que les iba la vida. Quizás la explicación haya que buscarla en otro plano. La cuestión es que la explosión se produjo al poner a prueba el reactor para comprobar hasta dónde podían bajar la potencia. Es relevante señalar que a mayor potencia mayor consumo de uranio y que el uranio es escaso y caro. Sin embargo, el plutonio generado como subproducto por el reactor es constante, independientemente de la potencia a la que funcione el reactor. No es difícil concluir el interés de las autoridades soviéticas del momento en intentar reducir la potencia del reactor. En una de estas pruebas se produjo el accidente.

3. *Las condiciones iniciales*: son los requisitos para que la producción sea un éxito (fiabilidad de las máquinas apropiadas, fuerza de trabajo competente, códigos de buena conducta). Es importante en esta etapa la distinción entre “tipos de fallos” y “signos de fallos” (Hudson 1988, 278). Una formación deficiente de los operarios es un tipo de agente patógeno que puede mostrarse a través de distintos signos. Hay que ob-

servar bien los signos para evitar posibles fallos. Algunos pueden evitarse porque son más evidentes (llevar el casco de seguridad), pero otros que afectan al estado psíquico del personal (divorcios, depresiones, etc.), son más difíciles de detectar, por lo cual, hay que introducir medidas para que los efectos no sean graves.

En el caso anterior los elementos humanos (psicológicos) y sociales están interrelacionados. Por ejemplo, el que haya fuerza de trabajo apropiada es una decisión de política empresarial (pueden elegir personas menos competentes para bajar los salarios) pero incluso con personal competente pueden producirse fallos aunque en mucho menor grado y, por tanto, corriendo menos riesgos.

4. *Las actividades productivas*: se trata de la sincronización precisa de las actividades técnicas y humanas para fabricar el buen producto en el buen momento. En esta etapa una actividad de riesgo, desde el punto de vista psicológico, es un error o una violación cometidos en presencia de un peligro potencial, es decir, de algún objeto, energía o agente tóxico que, si no son debidamente controlados, pueden causar heridas o males mayores.

En este punto, al aumentar el riesgo, la responsabilidad de poner todos los medios humanos y sociales para que no suceda es mucho mayor. Se sabe desde el principio que la actividad comporta riesgos y que hay peligros reales, por tanto, es dudoso que aquí pueda haber equivocaciones de planificación basados en conocimiento general. Si no hay irresponsabilidad ética los errores que podrían esperarse son lapsos, deslices y, en todo caso, equivocaciones basadas en reglas, que por falta de competencia o falta de atención se produjera algún fallo.

5. *Los sistemas de producción*: cuando la producción está expuesta a peligros naturales o implícitos, hay que estar preparado para posibles desgracias humanas y técnicas. Los fallos en la protección del sistema suelen venir por una serie de errores latentes que coinciden con incidentes en un momento determinado y que provocan el accidente. Las protecciones en serie pueden a veces no funcionar porque a los operarios les falta una visión global del sistema y al actuar sobre la seguridad de un elemento provocan fallos en otro. Los errores latentes son pro-

ducidos por equivocaciones pero el no haberlos atendido a tiempo es responsabilidad de la dirección, sea la empresa, la institución o una persona.

Todos los elementos implicados en una actividad tecnológica (en realidad en toda actividad humana pero aquí nos hemos centrado en la tecnología) incluyen elementos cognitivos y sociales. Cualesquiera de las etapas y sus causas de errores tienen fallos cognitivos y sociales (políticos, económicos) que, en último término, apuntan a una responsabilidad moral.

6. Evaluación y meta-evaluación del riesgo

Uno de los aspectos en el estudio de los errores es la evaluación probabilística del riesgo a partir de diversas técnicas de análisis de fiabilidad humana para tratar diferentes tipos de errores y diferentes situaciones. Hay que señalar que nos referimos sólo a los riesgos causados por la comisión de errores, en el entendido de que no todas las situaciones de riesgo están provocadas por errores. Diría más, muchos riesgos a los que estamos expuestos actualmente no están provocados por errores en el sentido instrumental que los tratamos aquí. Por ejemplo, el riesgo que comportan determinados pesticidas, emisiones de CO_2 , y un largo etcétera no está relacionado con errores humanos, por tanto, su tratamiento tiene que proceder de la política y la ética, además de los informes de los expertos.

Volviendo a los riesgos provocados por errores humanos, hay que decir que muchos de los estudios y consideraciones sobre el tema se han hecho pensando en las centrales nucleares, aunque esto no significa que no puedan extrapolarse a otros campos. Existen diferentes técnicas de análisis en torno a las que ha surgido un debate sobre las cualidades y defectos de cada una de ellas. Fruto de ello es la (meta) evaluación de las diversas técnicas desde la perspectiva de los ingenieros y desde la de los psicólogos especialistas en factores humanos. Entre las técnicas de detección y tratamiento de errores hemos visto los seguros en

serie, la información del entorno, etc. La meta-evaluación consiste en evaluar las distintas técnicas de detección y análisis de errores.

Los resultados de la evaluación pueden marcar el camino de cómo eliminar las ocasiones de error. Aquí es donde, como el mismo Reason reconoce, las ideas de D. Norman (1986 y 2004) sobre la adaptación de los diseños a las capacidades humanas pueden ser un factor importante para evitar ocasiones de error. Además están los sistemas inteligentes de ayuda a la toma de decisiones, de ayuda a la memoria, los programas de formación y el aumento de la tolerancia del sistema al error. En último término, lo que tenemos es una adaptación de las máquinas a los seres humanos pero también de éstos a la tecnología.

Por otro lado, la evaluación de los riesgos nos lleva a la gestión de los mismos, campo en el que los estudios CTS tienen un papel relevante. A veces, el enfoque CTS de gestión de riesgos se ve contrapuesto a un estudio técnico y científico de los mismos. Se cree que las grandes catástrofes de las últimas décadas (los accidentes de Harrisburg y Chernobyl, Bhopal, etc.) han restado credibilidad al estudio de los riesgos; sin embargo, es la mejor forma de prevenirlos. Esto no significa que los dictámenes científicos sean suficientes para evitar los accidentes en su totalidad ya que, como hemos visto, hay que afrontar este problema desde la multidisciplinariedad, contemplando, tanto los factores cognitivos como los sociales. Por tanto, la perspectiva objetiva que tiene como objetivo el análisis científico y técnico de los riesgos es una condición necesaria aunque no suficiente para prevenir las catástrofes y los efectos perversos de las tecnologías.

7. Éxitos y fracasos de la actividad ingenieril

La mirada de Reason desde las ciencias cognitivas de las catástrofes es buscar las causas en fallos humanos, la de Petroski, desde la ingeniería, es buscar las causas de los fracasos desde el punto de vista técnico.

La ingeniería se rige por los éxitos y fracasos de sus construcciones. Cuando se trata de valorar las obras de ingeniería se habla de fracasos,

en buena parte causados por errores de todo tipo. También pueden ser provocados por violaciones o sabotajes. Los últimos, afortunadamente, no son habituales, pero las violaciones (voluntarias o involuntarias) lo son demasiado (pensemos en las infracciones de los conductores).

La ingeniería de estructuras es la ciencia y el arte de diseñar y de hacer, con economía y elegancia, edificios, puentes, construcciones y otras estructuras similares de tal forma que puedan resistir de forma segura, las fuerzas y pesos a las que están sujetas. Comparte ciertas características con los postulados de las teorías científicas, pero en lugar de hacer hipótesis sobre un universo dado, sean átomos, colmenas de abejas o planetas, los ingenieros hacen hipótesis sobre combinaciones de cemento y acero con los que construyen su propio mundo. Por tanto, cada nuevo edificio o puente puede considerarse una hipótesis en sentido propio. En particular, una hipótesis de un ingeniero de estructuras debe ser que determinado puente sobre un río, bajo determinadas condiciones de tráfico y mantenimiento, durará tantos años sin caerse. Si dicho puente se cae sin que ocurran circunstancias extraordinarias, significa que la hipótesis no era válida.

No cabe duda que un ingeniero debe tener en cuenta las funciones y usos de aquello que diseña, pero como no puede considerarlas todas, el estudio de las circunstancias contextuales que acompañan el producto (puente, edificio, etc.) son claves para no correr riesgos innecesarios. Más adelante abordaremos la cuestión de la “fatiga” que tiene mucho que ver con este problema.

Ahora bien, el éxito de una estructura ingenieril confirma la hipótesis de su función en el mismo sentido que la hipótesis de que cada día sale el sol es confirmada. Es decir, las hipótesis en ingeniería, como las hipótesis científicas, carecen de certeza absoluta (problema de la inducción aplicado a la ingeniería). Esta falta de certeza absoluta es lo que lleva a Cranor (2004, 113) a decir que debido a que las conclusiones científicas derivadas de estudios en animales y otros sistemas no humanos, o incluso de estudios epidemiológicos en humanos, no pueden ser establecidas con certeza por el carácter abierto de la ciencia, y debido a la complejidad de la naturaleza, los científicos tienden a “eludir” las conclusiones en los

artículos científicos. Este comentario va dirigido a mostrar que la gestión de los riesgos no puede estar en manos de los científicos. Cranor sigue la línea de los CTS más relativistas al culpar a la ciencia de los desastres tecnológicos y de su incapacidad para paliarlos. Hay que decir que la ciencia es abierta porque no es un dogma, pero esto no significa que no puedan sacarse conclusiones.

En buena parte, estos comentarios de Cranor son fruto de no distinguir entre riesgos procedentes de errores humanos (en cualquiera de las fases de la producción) y aquellos fruto de políticas científicas, económicas y sociales determinadas que, en el fondo, requieren un planteamiento ético.

8. El método de ensayo y error

En buena parte los métodos de la ingeniería son de ensayo y error. Las pirámides de Egipto y las catedrales medievales son ejemplos de construcciones por este método. Por ejemplo, en 1284 la catedral de Beauvais se derrumbó y este incidente supuso un punto de inflexión en el desarrollo de las estructuras góticas. Tanto las pirámides como las catedrales pertenecen a la edad pre-racional de la ingeniería estructural, que estaba basada más en el experimento físico y en la corrección a medida que se iba construyendo que en un plan previo sobre la versión final de la estructura.

Durante los siglos XIX y XX se desarrollaron las grandes estructuras de los puentes, pasando de la piedra al acero y al cemento. Los puentes adquirieron una gran importancia al ser el lugar por donde circulaban los ferrocarriles, con lo cual necesitaban de unos soportes más reforzados. La historia de los puentes modernos es también la historia de un enfoque cada vez más científico del diseño de grandes estructuras. Sin embargo, la construcción de puentes de acero y cemento no estuvo exenta de fallos, con la consiguiente catástrofe que ello ocasionaba. El número de accidentes era tan elevado que se llegaron a hacer chistes y comentarios críticos sobre los puentes de acero, pero los puentes de ma-

dera no eran la solución ya que la madera podía podrirse y ser pasto de las llamas.

Petroski (1995) ejemplifica el análisis de la actividad ingenieril a partir del estudio de casos como el del hotel Hyatt de Kansas City, en julio de 1981, con 114 muertos, más de 200 heridos y la mitad de la población de Kansas directa o indirectamente relacionada con la tragedia; el Mianus Bridge en 1967; y el DC-10 Chicago, en 1979, entre otros. En todos ellos había indicios de que algo iba mal antes de los accidentes. En realidad no eran más que los errores latentes de los que nos habla Reason. Otros indicios que pueden considerarse errores latentes son las grietas, aviso de posibles accidentes si no se les pone remedio. El proceso de construcción de la campana de Filadelfia que se agrietó la primera vez que sonó en 1752, además de otros casos como el Big Ben del parlamento de Londres, nos llevan a la conclusión de que del 50 % al 90 % de los fallos de estructura, incluyendo campanas, puentes, aviones y otros productos tecnológicos, son debido a un aumento de las grietas (Petroski 1995, 109).

Del estudio de casos, tanto antiguos como recientes, Petroski concluye que se puede aprender mucho de los fallos. Cuando se produce un accidente se sabe qué es lo que no hay que hacer, mientras que de los éxitos simplemente se repite lo hecho sin introducir cambios. Hay una especie de proceso cíclico en el cual cuando hay un fallo, el diseño es más conservador (al correr menos riesgos), pero luego viene una etapa sin fallos y la construcción se relaja, hay un fallo y vuelve a empezar. Para evitar en lo posible los fallos, hay que tener en cuenta el factor de seguridad que tiene la función de proporcionar un margen de error que permite un considerable número de corolarios de la Ley de Murphy sin amenazar el éxito de la actividad ingenieril. El factor de seguridad se calcula dividiendo el peso requerido para causar un fallo por el máximo peso que se espera que vaya a tener que soportar la estructura. Por ejemplo, si una cuerda con una capacidad de 6000 kilos se utiliza para un montacargas que con mucho tiene que levantar 1000 kilos, el factor de seguridad será $6000:1000 = 6$. Sin embargo, hay que tener en cuenta las posibles condiciones adversas como viento, fatiga, más peso en un

momento determinado, con lo cual, a lo mejor el factor 6 no sería suficiente. La idea fundamental que hay detrás del factor de seguridad es un medio de hacer explícito el fallo y para poder calcular el peso que puede causarlo.

Relacionado con los fallos está el concepto de “fatiga”, una de las causas de accidentes y catástrofes. En algún sentido podemos relacionarlo con el uso de las cosas. Por ejemplo, de las teclas de un ordenador se estropearán más fácilmente las que corresponden a las letras que más se utilizan en una lengua que otras que raramente salen en las palabras. El “enter” puede ser la que más se estropee. Se estropean las bombillas que más se encienden y se apagan. Todos estos son problemas de fatiga. Tenerlas en cuenta puede evitar accidentes. El error, en este caso, reside en no haber diseñado teclas más resistentes para las letras más habituales.

La importancia de los fracasos hizo prosperar lo que se ha llamado “ingeniería forense” o de análisis de fallos que consiste en analizar el fallo una vez ha sucedido, como un forense analiza un cadáver, y las causas de su muerte como en una investigación policial, buscando los indicios a la “Sherlock Holmes”.

En cuanto a la instrumentalización utilizada por los ingenieros es especialmente relevante la regla de cálculo durante la primera mitad del siglo XX. A principios de los sesenta se pasó a la tecnología digital y más adelante al ordenador. Esta evolución trajo ventajas e inconvenientes y representa el paso de las tecnologías transparentes a las opacas, con todo lo que supone de posibles errores humanos. Petroski insiste en la fuente de fallos que ha supuesto la introducción del ordenador pero hay que señalar que la última edición de este libro es de 1995. Habría que revisar los comentarios sobre la utilización de los ordenadores y, en general, de la ciencia de la computación para el diseño de estructuras ingenieriles. En todo caso lo que es relevante es que en cuanto se introduce tecnología digital hay que adaptar dicha tecnología al procesamiento de la información y a los procesos cognitivos de los humanos tal como preconiza Norman cuando aborda el uso de los ordenadores.

En el caso de Petroski, aunque aborde los problemas técnicos y sociales no son desde el punto de vista de la responsabilidad moral sino

que parte de la base de que el diseñador es responsable y actúa con las mejores intenciones. Los fallos son por falta de conocimiento de las especiales circunstancias en las que se aplicará la tecnología. Todos sabemos que no siempre es así, pero entonces entramos en el campo de la responsabilidad ciudadana, del código deontológico de los científicos y de la ética en general.

9. Causalidad y límites del diseño

El análisis de las causas de los accidentes ha sido abordado desde distintos puntos de vista. A modo de ejemplo podemos señalar dos aportaciones, una la de Thomas McKaig, 1962, *Building Failures*, quien sostiene que normalmente los edificios fallan por ignorancia, falta de cuidado o codicia. Presenta una lista anónima de causas de fallos de la que dice que es difícil pensar en un problema que no esté contenido en esta lista, según la cual los fallos pueden agruparse en cuatro bloques: ignorancia; economía de gastos; lapsos, deslices y falta de cuidado; y circunstancias inusuales (terremotos, tormentas, etc.).

Otra aportación interesante es la de D. I. Blockley, *The nature of structural design and safety* (1980), quien también da una lista de las causas de fallos estructurales agrupadas en tres bloques: estados límite (sobrepeso, deterioro, etc.), situaciones azarosas (fuego, terremotos, inundaciones) y errores humanos (en el diseño, en la construcción, etc.).

Las listas de McKaig y de Blockley revelan diferencias importantes respecto al papel que dan al individuo en las causas de los fallos. La de McKaig es mucho más antropocéntrica. Pero incluso si hay acuerdos sobre listas de causas de fallos, todavía puede haber diferentes interpretaciones sobre el papel que se le da a cada categoría en un accidente concreto. Aquí los colectivos con poder económico, político o social constituyen grupos de presión para inclinar la balanza en un sentido u otro.

Como contrapunto a la búsqueda de causas de los fallos en la actividad ingenieril es relevante el análisis de Ch. Perrow (1984). Según este autor no hay forma de predecir todas las combinaciones de posibles fa-

llos que puedan ocurrir en un sistema con un alto grado de complejidad. Así, el término “normal” tiene el sentido de propiedad inherente al sistema y señala: “tal vez el aspecto más original del análisis es que se enfoca en las propiedades del sistema mismo, antes que en los errores de los propietarios, diseñadores y operarios” (Perrow 1984, 63). Aunque sería simplista pensar que para Perrow no puede hacerse nada para prevenir accidentes en determinados sistemas tecnológicos (por ejemplo, en las centrales nucleares), sí veo un cierto fatalismo en su postura. Que no hay sistemas sin un cierto grado de riesgo, es cierto, que cuanto más complejo el sistema más riesgo, también, pero que los errores de propietarios, diseñadores y operarios queden en segundo plano creo que no está justificado ni es lo que nos enseñan la experiencia ingenieril y los estudios cognitivos.

El estudio de los fallos nos lleva a los límites de la ciencia del diseño sobre los que caben algunas reflexiones. Un ingeniero puede preguntarse si llegará un día en que los diseñadores serán capaces de afirmar con seguridad y determinación “éste es un diseño sin fallos”. La cuestión es compleja por lo que es difícil dar una respuesta categórica, en un sentido u otro. Por un lado, podríamos contestar afirmativamente ya que el proceso puede converger en un diseño más fiable y más razonable; pero también negativamente ya que nunca puede garantizarse que se produzca un producto sin fallos.

Teniendo en cuenta la importancia del método de ensayo y error, la experiencia en la producción y construcción de objetos y materiales juega un papel importante en los éxitos y fracasos. Por ejemplo, durante siglos se fabricó espadas, por tanto, hubo mucho tiempo para hacerlas cada vez mejor y que las hojas de metal no tuvieran ningún fallo ni defecto. Por el contrario hace no más de 250 años que se construyó el primer puente en Coalbrookdale (1777-1779) al comienzo de la revolución industrial. De todas formas, en último término, conocer la historia de la tecnología no es suficiente para que los errores no se repitan.

Por tanto, hay que aceptar que la innovación implica un cierto riesgo. La valoración y gestión del mismo es una cuestión distinta. En cuanto a la evaluación hay que tener en cuenta las causas de los fallos, los me-

canismos de detección de errores y tener presente que la forma de todas las cosas diseñadas es el producto de una elección arbitraria. Si uno varía los términos de su compromiso —digamos, velocidad, más calor, menos seguridad, menos comodidad, coste más bajo— entonces uno está variando la forma del artefacto diseñado.

Finalmente, hemos dicho que la ingeniería está entre la ciencia y el arte. Petroski propone varias analogías, el ingeniero como poeta, como planificador de viajes o como creador de una obra literaria, pero hay diferencias sustanciales. Cuando hay una errata en un libro, en el siguiente se corrige y el problema queda resuelto. En las obras ingenieriles un error puede suponer vidas humanas, catástrofes naturales, etc. Las consecuencias de las erratas literarias y los errores ingenieriles son totalmente distintas.

10. Conclusiones

- Los errores humanos están a la base de muchos accidentes y catástrofes ocasionados por el desarrollo tecnológico. Esto no significa ni que todos los efectos negativos de la tecnología sean achacables a errores humanos ni que éstos sólo se produzcan en entornos tecnológicos. Sin embargo, hay un punto de conexión entre errores humanos y accidentes por fallos tecnológicos que hace relevante el estudio de esta interrelación.
- La separación de los elementos cognitivos y sociales ha sido un hecho en las últimas décadas del siglo XX, especialmente por todos los que más se han interesado por las consecuencias negativas del desarrollo tecnológico. El análisis que aquí se presenta muestra que esta desconexión no tiene sentido, antes al contrario que es absolutamente necesaria para mitigar estos efectos negativos. No es suficiente pero sí necesario.
- La aproximación desde la psicología cognitiva nos proporciona información sobre los procesos cognitivos relacionados, funda-

mentalmente, con la memoria y la atención, dos facultades claves para evitar cometer errores.

- La aproximación ingenieril nos lleva a tener en cuenta, por un lado, las cuestiones científico/técnicas y, por otro, los factores sociales, políticos y económicos que entran en juego en el diseño, organización y proceso tecnológico.
- La aproximación a los errores puede hacerse desde muchas perspectivas, pero no cabe duda que una de ellas es la que aquí hemos abordado como muestra de la conexión entre lo cognitivo y lo social.

Bibliografía

- Asimov, M. 1974. A philosophy of engineering design. En Rapp 1974, 150-157.
- Bechman, G. 2004. Riesgo y sociedad post-moderna. En Luján y Echeverría 2004, 17-34.
- Blockley, D. I. 1980. *The nature of structural design and safety*. West Sussex: Chichester; Ellis Horwood Limited.
- Bunge, M. 1966. Technology as applied science. *Technology and Culture* 1: 329-347.
- ——— 1976. The philosophical richness of technology. *Philosophy of Science* 2: 153-172.
- Clark, Andy. 2003. *Natural-born cyborgs. Minds, technologies, and the future of human intelligence*. Oxford: University Press.
- Clark, A. y D. Chalmers. 1998. The extended mind. *Analysis* 58/1: 7-19.
- Cranor, C. F. 2004. Conocimiento experto y políticas públicas en las sociedades tecnológicas. En búsqueda del apoyo científico

- apropiado para la protección de la salud. En Luján y Echeverría 2004.
- D'Andrade, R. 1989. Culturally based reasoning. En Gellatly, Rogers y Sloboda 1989.
 - Dekker, S. 2006. *The field guide to understanding human error*. Hampshire (England): Ashgate Publishing.
 - Echeverría, J. 2003. *La revolución tecnocientífica*. México; Madrid: Fondo de Cultura Económica.
 - Estany, A. 1999. *Vida, muerte y resurrección de la conciencia. Análisis filosófico de las revoluciones científicas en la psicología contemporánea*. Barcelona: Paidós.
 - ——— 2001. The thesis of theory-laden observation in the light of cognitive psychology. *Philosophy of Science* 68: 203-217.
 - ——— 2005. Progress and social impact in design sciences. En González 2005, 135-159.
 - Fauconnier, G. y M. Turner. 2002. *The way we think*. Oxford: Basic Books.
 - Gellatly, A. y D. Rogers y J. Sloboda, eds. 1989. *Cognition in social worlds*, New York: McGraw-Hill.
 - Godoy, L. A., C. Escaudar, R. Jaca y F. 2001. Revisión crítica de algunas teorías de accidentes asociadas a la infraestructura. *Revista Internacional de Desastres Naturales, Accidentes e Infraestructura Civil* 1/2: 127-139.
 - González, W. J., ed. 2005. *Science, Technology and Society: A Philosophical Perspective*, A Coruña: Netbiblo.
 - Hall, A. D. 1974. Three-dimensional morphology of systems engineering. En Rapp 1974.

- Hutchins, E. 1995. *Cognition in the wild*, Cambridge (MA): The MIT Press.
- Jacques, R. y J. A. Powell, eds. 1981. *Design: Science: Method*. England: Westbury House.
- Knorr-Cetina, K. D. 1981. *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*, Oxford: Pergamon.
- Kotarbinski, T. 1962. Praxiological sentences and how they are proved. En Nagel, Suppes y Tarski 1962, 211-223.
- ——— 1965. *Praxiology. An introduction to the science of efficient action*. New York: Pergamon.
- Latour, B. y S. Woolgar. 1979. *Laboratory life: the social construction of scientific facts*. Londres: Sage.
- López Cerezo, J. A. y J. L. Luján. 2000. *Ciencia y política de riesgo*. Madrid: Alianza.
- Luján, J. L. y J. Echeverría, eds. 2004. *Gobernar los riesgos. Ciencia y valores en la sociedad del riesgo*, Madrid: Biblioteca Nueva.
- McCrory, R. J. 1974. The design method-A scientific approach to valid design. En Rapp 1974, 158-173.
- McKaig, T. K. 1962. *Building failures: case studies in construction and design*. New York: McGraw-Hill.
- Miller, G., Galanter, E. y Pribram, K. 1960. *Plans and structure of behaviour*. New York: Holt.
- Nadler, G. 1967. An investigation of design methodology. *Management science* 13/10: B642-B655.
- Nagel, E., P. Suppes y A. Tarski, eds. 1962. *Logic, Methodology and Philosophy. Proceedings of the 1960 International Congress*.

- Niiniluoto, I. 1993. The aim and structure of applied research. *Erkenntnis* 38: 1-21.
- Norman, D. A. 1985. *El aprendizaje y la memoria*, Madrid: Alianza.
- ——— 1986. Cognitive engineering. En Norman y Draper 1986, 31-61.
- ——— 1990. *La psicología de los objetos cotidianos*. Madrid: Editorial Nerea.
- ——— 1992. *Turn signals are the facial expressions of automobiles*. Reading (MA): Addison-Wesley.
- ——— 1993. *Things that make us smart. Deffending human attributes in the age of the machine*. Cambridge (MA): Perseus Books.
- ——— 2004. *Emotional design. Why we love (or hate) everyday things*. New York: Basic Books.
- Norman, D. A. y S. W. Draper, eds. 1986. *User centered system design. New perspectives on human-computer interaction*. Hillsdale (NJ): Erlbaum.
- Perrow, C. 1984. *Normal accidents: living with high-risk tecnologies*. New York: Basic Books.
- Petroski, H. 1982. *To engineer is human. The role of falilure in successful design*. New York: St. Martin's Press.
- ——— 1995. *Design paradigms. Case histories of error and judgment in engineering*. Cambridge (MA): University Press.
- ——— 2003. *Small things considered. Why there is no perfect design*. New York: Alfred A. Knopf.

- ——— 2006. *Success Through Failure: The Paradox of Design*. Princeton: University Press.
- Quintanilla, M. A. 2005. *Tecnología: un enfoque filosófico y otros ensayos sobre filosofía de la tecnología*. México; Madrid: Fondo de Cultura Económica.
- Rapp, F., ed. 1974. *Contributions to a philosophy of technology*. Dordrecht (Holanda): D. Reidel.
- Reason, J. 1990. *Human error*. Cambridge (MA): University Press.
- Reason, J. y A. Hobbs. 2003. *Managing maintenance error: A practical guide*. Hampshire: Ashgate Publishing.
- Shrader-Frechette, K. 2005. Objectivity and professional duties regarding science and technology. En González 2005, 51-79.
- Simon, H. 1996. *The science of the artificial*, Cambridge (MA): The MIT Press.
- Vallverdú, J. 2002. *Marc teòric de les controvèrsies científiques: el cas de la sacarina*. Tesis doctoral, Barcelona: Universidad Autónoma de Barcelona.