

Modelos para la evaluación del error humano en estudios de fiabilidad de sistemas

Juan M. Ruiz-Moreno^{1*} y Humberto M. Trujillo²

¹Torraspapel, S.A. Grupo LECTA

²Facultad de Psicología, Universidad de Granada

Resumen: El análisis probabilístico de riesgos aplicado a sistemas industriales o técnicos complejos demuestra que el error humano es un elemento muy importante del riesgo total, aunque su cuantificación exacta es dudosa. Por otra parte, las técnicas actuales para el modelado y cálculo de los errores humanos presuponen, de manera más o menos explícita, la elección de algún modelo de comportamiento humano; de tal manera que buena parte de las limitaciones que en esas técnicas refleja la literatura científica parecen debidas al empleo de los modelos de comportamiento denominados parciales, mientras que los modelos denominados generales, más ambiciosos, adolecen aún de desarrollo metodológico. El presente trabajo describe los métodos aplicados actualmente en los estudios de fiabilidad humana y señala la relación que subyace entre éstos y los modelos de representación del comportamiento humano, poniendo de manifiesto las líneas de investigación en curso para cubrir las carencias metodológicas. Para el avance de los estudios sobre fiabilidad humana, se manifiesta el protagonismo de la Psicología, tanto en la fundamentación de los modelos de comportamiento humano que subyacen en las técnicas de análisis de la fiabilidad, como en la crítica de estas mismas técnicas, en especial hacia las simplificaciones que, por hacerlas más manejables, tienen como efecto perverso un encubrimiento de los mecanismos psicológicos del error.

Palabras clave: análisis de fiabilidad; error humano; comportamiento humano.

Title: Models for assessment of human error in system reliability studies.

Abstract: The probabilistic analysis of risks applied to complex industrial systems demonstrates that human error is a very important part of the total risk, although its exact quantification is doubtful. On the other hand, the current techniques for modeling and calculating human errors presuppose, in a way not always explicit, the election of a certain pattern of human behavior; in such a way that many limitations revealed by the scientific literature seems due to the use of so named 'partial' behavior models, while the 'general' models, that are more ambitious, still need better methodological development. The present work describes the methods applied currently to human reliability studies and it points out the relationship that underlies between these and the models for human behavior representation, showing the investigation lines ongoing for covering methodological lacks. We declare the protagonism of Psychology about defining the human behavior models that underlie the reliability analysis techniques, and also about criticizing these techniques, especially on the simplifications that, to make them more governable, they have the perverse effect of hiding psychological mechanisms of errors.

Key words: reliability analysis; human error; human behavior.

Introducción

Los estudios sobre análisis probabilístico de riesgos (Apostolakis, Bier, y Mosleh, 1988; Williams, 1989; Kirwan, 1994; CSNI, 1998; Rea y Sandoval, 2000; Pyy, 2000; Creus, 2001; Johnson, 2003; Maluf, Gawdiak, y Bell, 2008) reflejan que la contribución de los errores humanos es muy significativa y, de forma casi unánime, exponen que las cantidades estimadas para estas contribuciones incluyen grandes incertidumbres. Por otra parte, las técnicas de modelado y cálculo de los errores humanos, incluidas en los estudios de fiabilidad de sistemas, guardan relación, más o menos explícita, con los modelos de comportamiento humano que subyacen en ellas (Quesada, 1999; EATMP, 2002).

En la mejora de la fiabilidad de los sistemas industriales, el punto clave se encuentra en la interacción hombre-máquina (Creus, 1992; EATMP, 2002; Johnson, 2003; Kim y Seong, 2004; Kennedy, Siemieniuch, Sinclair, Kirwan, y Gibson, 2007; Forester y cols., 2009); pero es obvio que resulta mucho más complicado analizar el comportamiento del hombre que el de la máquina, y ésta es una dificultad que ha dado pie a muchas líneas de investigación multidisciplinar, especialmente en aquellos sectores industriales donde el impacto de los posibles errores humanos es más fuerte: energía nuclear, aviación e industria química.

Los estudios de fiabilidad humana, como parte de los de

fiabilidad de sistemas, tuvieron un antes y un después a partir del famoso accidente nuclear de Three-Mile Island, en 1979. Antes de ese momento, la inclusión del análisis específico del error humano quedaba a criterio de los analistas, en función de la apreciación o de los prejuicios que tuvieran respecto a las posibles causas de fallo del sistema. Pero, a partir de ese accidente, el análisis de la fiabilidad humana se hizo obligatorio en la industria nuclear y, por extensión, a cualquier actividad susceptible de fallo catastrófico.

Durante los años ochenta se produjo el desarrollo de las primeras técnicas específicas; pero ha sido a partir de los noventa cuando, además de continuar el desarrollo de nuevos métodos, se ha pretendido profundizar en los mecanismos causantes del error humano.

De las distintas técnicas para cuantificar y predecir el error humano -véase una relación amplia en (CSNI, 1998)-, la metodología THERP es la más antigua y, con las adaptaciones realizadas en los años ochenta más la incorporación de la variable temporal a los errores de diagnóstico, sigue siendo la más utilizada en la práctica y puede considerarse como el método paradigmático respecto de los puntos fuertes y deficiencias que caracterizan a las técnicas de fiabilidad humana en vigor: buen análisis cualitativo de los modos de error y buena computabilidad y verificabilidad de los resultados, frente a los elementos de subjetividad en la apreciación de cada analista y al hecho de ignorar los mecanismos psicológicos del comportamiento.

En los últimos años el avance de los estudios sobre fiabilidad humana va muy de la mano de las aportaciones en el campo de la Psicología: por un lado, en la fundamentación

Dirección para correspondencia [Correspondence address]: Juan M. Ruiz-Moreno. Torraspapel, S.A. Grupo LECTA. Camino de la Vía, s/n. 18600 Motril (Granada) (España).
E-mail: ojmruiz@torraspapel.es

de los modelos de comportamiento humano que subyacen en las técnicas de análisis de la fiabilidad y, por otro, en la crítica de estas mismas técnicas, en especial hacia las simplificaciones de tipo operativo, que tienen como efecto perverso un encubrimiento de los mecanismos psicológicos del error, y ello, a su vez, impone una limitación fuerte al desarrollo de mejoras a los sistemas. La aportación más prometedora corresponde a los llamados modelos generales (Rasmussen, Norman, Shallice...) de representación del comportamiento humano, como solución a los defectos de los modelos mecanicistas.

Fiabilidad de sistemas. Revisión metodológica

Para situar el campo del análisis del error humano, dentro del espectro metodológico de los estudios de fiabilidad de sistemas, empezaremos con un recorrido por los conceptos y herramientas matemáticas de uso más frecuente.

Conceptos generales

El objetivo de partida es la necesidad de garantizar la fiabilidad y seguridad de los sistemas complejos, mediante procedimientos con suficiente respaldo científico. Y denominamos *sistema* a toda entidad formada por elementos o componentes discretos que interactúan mutuamente (Creus, 1992). Por tanto, su *complejidad* estará directamente relacionada con la cantidad de elementos que lo componen y con la cantidad o variedad de interacciones que puedan darse entre ellos.

Es esencial en la definición del sistema, que se delimiten:

- El *contorno externo*, también denominado alcance u objeto del análisis. Define el criterio por el cual se determina claramente la pertenencia o no de los elementos al sistema.
- El *contorno interno*, que establece el nivel de detalle o resolución con que se definen los elementos constitutivos del sistema; considerando que cada elemento podría ser, a su vez, un subsistema formado por subelementos; estos a su vez compuestos de otros subelementos, y así sucesivamente.

Una vez definido el sistema, interesa analizar la fiabilidad individual de cada elemento componente. Y entendemos por *fiabilidad* la *probabilidad de que un elemento (aparato, dispositivo, persona...) desarrolle una determinada función, bajo ciertas condiciones y durante un período de tiempo determinado (entre 0 y t)*. Es por ello que su análisis se denomina *análisis probabilístico de riesgos*.

Se suele representar por $R(t)$ la fiabilidad (*Reliability*) hasta el instante t , y partimos de la seguridad de que el elemento funciona en el instante 0 : $R(0)=1$. Con la misma seguridad, predecimos que en un tiempo largo el elemento habrá fallado: $R(\infty)=0$. Si analizamos una cierta cantidad N de elementos iguales, la fiabilidad en el instante t vendrá dada por:

$$R(t) = \frac{N_s(t)}{N(0)} = \frac{N(0) - N_f(t)}{N(0)} = 1 - \frac{N_f(t)}{N(0)} \quad (1)$$

en la que:

$N_s(t)$ = Número de elementos que continúan funcionando.

$N_f(t)$ = Número de elementos que han fallado.

$N(0)$ = Número de elementos ensayados.

En términos de probabilidad, $R(t)$ se denomina también función de supervivencia. Si llamamos T al momento en que falla la unidad:

$$R(t) = P(T > t) \text{ para } t > 0$$

Y derivando la ecuación (1) respecto a t , obtenemos la función de densidad de probabilidad de no fallo:

$$\frac{dR(t)}{dt} = -\frac{1}{N(0)} \frac{dN_f(t)}{dt} \quad (2)$$

En consecuencia, la probabilidad de fallo (infiabilidad) será:

$$Q(t) = 1 - R(t) = P(T \leq t)$$

Y la función de densidad de fallos:

$$f(t) = \frac{dQ(t)}{dt} = -\frac{dR(t)}{dt}$$

Despejando $dN_f(t)/dt$ en la ecuación (2) y dividiendo por $N_s(t)$, obtenemos la tasa instantánea de fallos:

$$\frac{1}{N_s(t)} \frac{dN_f(t)}{dt} = -\frac{N(0)}{N_s(t)} \frac{dR(t)}{dt} = -\frac{1}{R(t)} \frac{dR(t)}{dt} = \frac{f(t)}{R(t)} = \zeta(t) \quad (3)$$

e integrando y despejando $R(t)$:

$$R(t) = e^{-\int_0^t \zeta(t) dt} \quad (4)$$

que es la expresión general de la fiabilidad en función de la tasa instantánea de fallos.

Sustituyendo en (3) y despejando $f(t)$, queda la expresión general de la función de densidad de fallos:

$$f(t) = \zeta(t) e^{-\int_0^t \zeta(t) dt}$$

Cuando la tasa instantánea de fallos es constante $\zeta(t)=\lambda$, tenemos la conocida función exponencial

$$R(t) = e^{-\lambda t} \quad \text{para la función de fiabilidad,}$$

y

$$f(t) = \lambda e^{-\lambda t} \quad \text{para la densidad de fallos.}$$

Funciones de distribución de fallos

Como análisis probabilístico de riesgos, en los estudios de fiabilidad se emplean varios tipos de distribuciones estadísticas (Bunday, 1991; Creus, 1992; Hoyland y Rausand, 1994; Biorolini, 2007), las más importantes de las cuales se describen a continuación. La adopción de una o de otra dependerá del grado de ajuste de los datos a la función considerada.

Distribución normal. Se dice que una variable T tiene una distribución normal de media μ y varianza σ^2 , $N(\mu, \sigma^2)$,

cuando la función de densidad de fallos es

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad \text{para } -\infty < t < \infty$$

A la distribución $N(0,1)$ se le denomina *distribución normal estándar* y habitualmente se le denota por $\Phi(\cdot)$.

La función de distribución de $N(\mu, \sigma^2)$ puede escribirse como

$$F(t) = P(T \leq t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

Para aplicar la distribución normal en los estudios de fiabilidad, habremos de considerar que la variable está truncada por la izquierda en el punto $T=0$; por tanto, la función de supervivencia habría de adoptar la forma

$$R(t) = P(T > t | T > 0) = \frac{\Phi((\mu-t)/\sigma)}{\Phi(\mu/\sigma)} \quad \text{para } t \geq 0$$

Distribución binomial. Llamando q a la probabilidad de que un componente sea defectuoso y r a la de que no lo sea, y siendo n el número de componentes, la probabilidad de que x de ellos sean defectuosos se representa por:

$$Q(x) = \binom{n}{x} q^x r^{n-x}$$

que es la llamada distribución binomial, o de Bernouilli, y puede aproximarse a una distribución normal cuando n es grande.

La media y desviación estándar de la distribución son, respectivamente:

$$\mu = nq \quad \sigma = \sqrt{nqr}$$

Distribución de Poisson. Si consideramos que el tiempo de reparación o sustitución de un dispositivo es despreciable, el evento 'fallo del dispositivo' puede tratarse como evento \mathcal{A} de un *proceso de Poisson* en un intervalo de tiempo dado.

Si denotamos $N(t)$ el número de veces que sucede el evento \mathcal{A} en el intervalo $(0, t]$, y $P(N(t)=n)$ la probabilidad de que \mathcal{A} suceda exactamente n veces en el intervalo $(0, t]$, puede demostrarse (Hoyland y Rausand, 1994) que:

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad \text{para } n = 0, 1, 2, \dots \quad (5)$$

Esta es la llamada *distribución de Poisson* de parámetro λt , cuya media es

$$E(N(t)) = \sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t$$

El número de ocurrencias de \mathcal{A} por unidad de tiempo ($t=1$) es λ , la *intensidad del proceso*.

La distribución de Poisson puede representarse también como un caso particular de la distribución binomial cuando el número de componentes defectuosos es pequeño, caso frecuente en los estudios de fiabilidad. También puede aproximarse a la normal de media $\lambda = nq$ y desvia-

ción típica $\sigma = \sqrt{nqr}$, donde $r = 1 - q$.

Distribución exponencial. Si asumimos que el tiempo T hasta el fallo de una unidad está exponencialmente distribuido con parámetro λ , la densidad de probabilidad será entonces

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{para } t > 0 \text{ y } \lambda > 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

La función de supervivencia o fiabilidad es

$$R(t) = P(T > t) = \int_0^{\infty} f(u) du = e^{-\lambda t} \quad \text{para } t > 0$$

Y la tasa de fallos:

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

En general, esta distribución representa la fase normal de operación del componente y es la más utilizada en los estudios de fiabilidad.

Distribución de Weibull. El tiempo T hasta el fallo responde a una distribución de Weibull de parámetros $a(>0)$ y $\lambda(>0)$ si la función de distribución viene dada por

$$F(t) = P(T \leq t) = \begin{cases} 1 - e^{-(\lambda t)^a} & \text{para } t > 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

La correspondiente densidad de probabilidad es

$$f(t) = \frac{d}{dt} F(t) = \begin{cases} \alpha \lambda (\lambda t)^{\alpha-1} e^{-(\lambda t)^a} & \text{para } t > 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

donde λ se denomina *factor de escala* y a es el *factor de forma*.

Mayor flexibilidad obtenemos si, en lugar de dos, consideramos la distribución Weibull de tres parámetros (a, λ, μ) , que se expresa como

$$F(t) = P(T \leq t) = \begin{cases} 1 - e^{-[\lambda(t-\mu)]^a} & \text{para } t > \mu \\ 0 & \text{en cualquier otro caso} \end{cases}$$

La función de densidad correspondiente es

$$f(t) = \frac{d}{dt} F(t) = \alpha \lambda [\lambda(t-\mu)]^{\alpha-1} e^{-[\lambda(t-\mu)]^a} \quad \text{para } t > \mu$$

Al tercer parámetro μ se le denomina *parámetro de posición inicial o de garantía*, puesto que la probabilidad de que ocurra un fallo en un tiempo anterior a μ es 0.

Distribución log-normal. Se dice que el tiempo T hasta el fallo tiene una distribución log-normal si la variable $Y = \ln T$ tiene una distribución normal. Si μ y σ son, respectivamente, la media y la varianza de Y , entonces la función densidad de T es

$$f(t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma t} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} & \text{para } t > \mu \\ 0 & \text{en cualquier otro caso} \end{cases}$$

La función de fiabilidad o supervivencia resulta

$$R(t) = P(T > t) = P(\ln T > \ln t) \\ = P\left(\frac{\ln T - \mu}{\sigma} > \frac{\ln t - \mu}{\sigma}\right) = \Phi\left(\frac{\mu - \ln t}{\sigma}\right)$$

La distribución log-normal se emplea habitualmente como distribución del tiempo de reparación y en el análisis de fallos por fatiga.

Fiabilidad de sistemas

Después de considerar los conceptos de fiabilidad y probabilidad de fallos para los elementos individuales, es necesario analizar su interacción en sistemas complejos (Hoyland y Rausand, 1994; Birolini, 2007), teniendo en cuenta las relaciones entre esos elementos.

Función estructura de un sistema. Supongamos que el sistema está formado por n componentes y que el estado de cada componente i está descrito por la variable $X_i(t)$, que toma el valor 1 si funciona y el valor 0 si no funciona en el instante t . El vector de estado y la función estructura las denotaremos mediante

$$\mathbf{X}(t) = (X_1(t), \dots, X_n(t)) \text{ y } \varphi(\mathbf{X}(t))$$

Son de interés, las siguientes probabilidades:

$$P(X_i(t) = 1) = E[X_i(t)] = p_i(t) \text{ para } i = 1, 2, \dots, n \quad (6)$$

$$P(\varphi(\mathbf{X}(t)) = 1) = E(\varphi(\mathbf{X}(t))) = p_S(t) \quad (7)$$

Las expresiones (6) y (7) son las *funciones de supervivencia*, respectivamente, del componente i y del sistema.

Sistemas en serie. Los sistemas en serie se caracterizan porque sólo funcionan si lo hacen todos y cada uno de sus componentes. En este caso, la función estructura del sistema es

$$\phi(\mathbf{X}(t)) = \prod_{i=1}^n X_i(t)$$

Considerando que los eventos de fallo en cada componente sean independientes, la fiabilidad del sistema es

$$E(\phi(\mathbf{X}(t))) = E\left(\prod_{i=1}^n X_i(t)\right) = \prod_{i=1}^n E(X_i(t)) = \prod_{i=1}^n p_i(t) \quad (8)$$

Sistemas en paralelo. Los sistemas en paralelo se caracterizan porque fallan sólo si todos sus componentes fallan a la vez. La función estructura del sistema es, por tanto,

$$\phi(\mathbf{X}(t)) = \prod_{i=1}^n X_i(t) = 1 - \prod_{i=1}^n (1 - X_i(t))$$

Y la fiabilidad,

$$E(\phi(\mathbf{X}(t))) = 1 - \prod_{i=1}^n (1 - E(X_i(t))) = 1 - \prod_{i=1}^n (1 - p_i(t)) = \prod_{i=1}^n p_i(t) \quad (9)$$

Estructuras k-de-n. Se denominan estructuras “k-de-n” a aquellas compuestas por n elementos y que funcionarán mientras estén en funcionamiento al menos k de ellos. Su función estructura será, por tanto,

$$\phi(\mathbf{X}(t)) = \begin{cases} 1 & \text{si } \sum_{i=1}^n X_i(t) \geq k \\ 0 & \text{si } \sum_{i=1}^n X_i(t) < k \end{cases}$$

Consideremos, para simplificar, que todos los elementos tienen la misma fiabilidad, $p_i(t) = p(t)$ para $i = 1, \dots, n$,

y que los fallos de componentes individuales son sucesos independientes. Entonces, en un instante dato t , la variable

$$Y(t) = \sum_{i=1}^n X_i(t) \text{ tendrá una distribución binomial } (n, p(t)):$$

$$P(Y(t) = y) = \binom{n}{y} p(t)^y (1 - p(t))^{n-y}$$

En estas condiciones, la fiabilidad es

$$p_S(t) = P(Y(t) \geq k) = \sum_{y=k}^n \binom{n}{y} p(t)^y (1 - p(t))^{n-y} \quad (10)$$

Análisis cuantitativo del árbol de fallos. Veremos ahora cómo puede analizarse una estructura que haya sido modelada mediante un árbol de fallos, muy frecuente en el análisis de la fiabilidad humana.

Sea n el número de eventos básicos diferentes en el árbol de fallos. Diremos entonces que el árbol es de orden n y los diferentes eventos (fallos) pueden representarse por variables de estado $Y_i(t)$, y el vector de estado para la estructura en el instante t es

$$\mathbf{Y}(t) = (Y_1(t), \dots, Y_n(t))$$

El propósito del “análisis cuantitativo” de un árbol de fallos es determinar la probabilidad del *evento superior* (fallo del sistema), evento que puede describirse mediante una variable binaria $\psi(\mathbf{Y}(t))$ denominada función estructura del árbol de fallos.

Sea $q_i(t)$ la probabilidad de que el evento i suceda en el instante t , para $i = 1, \dots, n$. Y sea $Q_0(t)$ la probabilidad de que el evento superior ocurra en el instante t . Entonces

$$Q_0(t) = P(\psi(\mathbf{Y}(t)) = 1) = E(\psi(\mathbf{Y}(t)))$$

y

$$P(Y_i(t) = 1) = q_i(t) = 1 - p_i(t) \text{ para } i = 1, \dots, n$$

donde $p_i(t)$ es la probabilidad de que el componente i esté funcionando en el instante t . A $q_i(t)$ se le denomina *infiabilidad del componente i* en el instante t , mientras que $Q_0(t)$ representa la *infiabilidad del sistema* en el mismo instante.

Árbol con una puerta tipo AND. Si el evento superior sucede si y sólo si suceden los eventos básicos simultáneamente, la función estructura será

$$\psi(\mathbf{Y}(t)) = Y_1(t) \cdot Y_2(t) \cdots Y_n(t) = \prod_{i=1}^n Y_i(t)$$

Y si los eventos básicos son independientes,

$$Q_0(t) = E(\psi(\mathbf{Y}(t))) = E(Y_1(t) \cdot Y_2(t) \cdots Y_n(t))$$

$$= q_1(t) \cdot q_2(t) \cdots q_n(t) = \prod_{i=1}^n q_i(t)$$

Árbol con una puerta tipo OR. Si el evento superior sucede si sucede al menos uno de los eventos básicos, la función estructura será

$$\psi(\mathbf{Y}(t)) = 1 - \prod_{i=1}^n (1 - Y_i(t))$$

Y con los eventos básicos independientes,

$$\begin{aligned}
 Q_0(t) &= E(\psi(\mathbf{Y}(t))) = 1 - \prod_{i=1}^n E(1 - Y_i(t)) \\
 &= 1 - \prod_{i=1}^n E(1 - Y_i(t)) = 1 - \prod_{i=1}^n (1 - q_i(t))
 \end{aligned}$$

Modelos de Markov

Los modelos de Markov describen las probabilidades de los distintos estados posibles de un sistema en función del tiempo. Las variables significativas son, por tanto, *estado* y *tiempo*; pudiendo éstas, a su vez, ser discretas o continuas, lo que implica la existencia de cuatro tipos diferentes de modelos de Markov.

En un sistema, cuya fiabilidad se esté estudiando, los distintos estados posibles representan la situación de fallo u operación en que se encuentre cada uno de los componentes del sistema; y el cambio de un estado a otro viene determinado por los valores de régimen de transición que representan las tasas de fallo y de reparación de esos componentes, configurando las denominadas *ecuaciones de estado*, cuya solución expresa la probabilidad de que el sistema esté en un determinado estado y el tiempo de permanencia en cada uno de ellos.

Ecuaciones de estado. Denotaremos por $X(t)$ la variable estado del sistema en el instante t . Supongamos que el sistema tiene $r+1$ estados posibles, numerados $0, 1, 2, \dots, r$. El evento $X(t)=j$ significa que el sistema se encuentra en el estado j en el instante t , para $j=0, 1, 2, \dots, r$. La probabilidad de este evento es

$$P_j(t) = P(X(t) = j) \text{ para } j = 0, 1, 2, \dots, r$$

La transición entre estados puede ser descrita mediante un proceso estocástico y, en muchos casos, con la propiedad de Markov:

Dado un sistema en un estado i en el instante t (es decir, $X(t)=i$), los estados futuros $X(t+v)$ no dependen de los estados previos $X(u)$, $u < t$. En otras palabras: si conocemos el estado actual, la probabilidad de cualquier comportamiento futuro del sistema no estará condicionada por el comportamiento pasado.

Si el proceso satisface la propiedad de Markov se le denomina *proceso de Markov*, y a las probabilidades condicionales $P(X(t+v)=j | X(t)=i)$ para $i, j = 0, 1, \dots, r$ se les denomina *probabilidades de transición*, que satisfacen:

$$P_{ij}(t) > 0 \text{ para } t > 0 \quad (11)$$

$$\sum_{j=0}^r P_{ij}(v) = 1 \text{ para } t > 0 \quad (12)$$

$$P_{ij}(t+v) = \sum_{k=0}^r P_{ik}(t) \cdot P_{kj}(v) \text{ para } t, v > 0 \quad (13)$$

Definimos la *tasas o coeficiente de transición* desde el estado i al j como

$$a_{ij} = \lim_{\Delta t \rightarrow 0} \frac{P(X(t+\Delta t) = j | X(t) = i) - P_{ij}(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t} = \dot{P}_{ij}(0) \quad (14)$$

Modelado de sistemas en red

Redes de Petri. Se trata de una herramienta gráfica para modelar las interacciones entre componentes de sistemas complejos y se usa en combinación con otras herramientas gráficas como los diagramas de bloques o los árboles de fallos.

Una red de Petri consta de los siguientes elementos:

- $P = \{p_1, p_2, \dots, p_{np}\}$ es el conjunto de np nodos, denominados *asentamientos* y dibujados como círculos.
- $T = \{t_1, t_2, \dots, t_{nt}\}$ es el conjunto de nt *transiciones*, dibujadas como barras.
- I = entradas de transición, representadas por flechas que salen de los asentamientos y se dirigen a las transiciones.
- O = salidas de transición, representadas por flechas que salen de las transiciones y se dirigen a los asentamientos.
- $M = (m_1, m_2, \dots, m_{np})$ es la representación de una *mana*.

El estado dinámico se representa mediante una o varias "fichas" que pueden ir saltando entre los nodos, siguiendo las reglas:

1. Un estado de transición en una marca M es permitido si todas las entradas disponen de al menos una ficha.
2. Una transición permitida es disparada extrayendo una ficha por arco de cada asentamiento de entrada y añadiendo una ficha por arco a cada posición de salida.

Aunque inicialmente aplicadas en la descripción de sistemas electromecánicos, la capacidad de adaptación de las redes de Petri ha quedado ampliamente demostrada, en las últimas décadas, en prácticamente todas las áreas de conocimiento, incluida la fiabilidad humana.

Redes neuronales artificiales. Las redes neuronales artificiales (RNA) son, en esencia, modelos matemáticos simples, sobre la base de funciones $f: X \rightarrow Y$. Cada tipo de RNA se caracteriza por la clase de funciones que utiliza.

El concepto de red, en este caso, se corresponde con el hecho de que las funciones $f(x)$ están compuestas por otras funciones $g_i(x)$ que, a su vez, pueden estar compuestas de otras funciones. Estas composiciones pueden representarse mediante una estructura en red.

Un tipo de composición usado con frecuencia en *redes acíclicas* es la suma ponderada no lineal, donde

$$f(x) = K \left(\sum_i w_i g_i(x) \right)$$

siendo K alguna función no lineal predefinida.

En el caso de *redes recurrentes* la función f se hace depender de sí misma, aplicando alguna dependencia temporal.

Una aplicación interesante del formalismo de RNA es la posibilidad de modelar procesos de aprendizaje; lo cual significa que, dada una tarea a realizar y una clase de fun-

ciones F , el sistema ha de encontrar la $f^* \in F$ que soluciona la tarea de forma óptima (en algún sentido). En general, la utilidad de las RNA se basa en la posibilidad de inferir una función a partir de un conjunto de observaciones; lo que resulta particularmente interesante en aplicaciones a partir de una cierta complejidad.

Evaluación del error humano

El factor humano siempre se ha considerado de la mayor importancia a la hora de determinar las posibilidades de éxito o fracaso de cualquier empresa, en el amplio sentido del término. Los estudios sobre la contribución del hombre al fallo del sistema calculan una responsabilidad entre el 50 y el 80 % de los casos de accidentes importantes. En la aviación, por ejemplo, un 70 % de los accidentes se estima que son causados por actuaciones indebidas (Creus, 1992).

Por similitud con la definición general de fiabilidad, consideraremos la *fiabilidad humana* como:

La probabilidad de que el comportamiento humano o su efecto sobre el sistema excedan el límite de aceptabilidad.

Y, por tanto, el error humano se presenta cuando el comportamiento humano o su influencia sobre el sistema exceden el límite de aceptabilidad, pero este límite debe definirse claramente, junto con los factores capaces de influir en el comportamiento humano.

Por contraposición al papel de la persona como fuente de error, debemos considerar también el componente de sobrefiabilidad que representa, ya que es capaz de anticipar y recuperar las desviaciones no previstas del sistema. En este sentido, el ser humano aventaja a la mayoría de los dispositivos técnicos por:

- . Su capacidad de adaptación ante situaciones no previsibles.
- . Por su capacidad de aprendizaje, es capaz de construir de varias maneras una estrategia de resolución y de adaptarla para el futuro.
- . Por la posibilidad de anticipar los acontecimientos y, por tanto, de modificar su estrategia inicial.
- . Y por la facultad de corregir sus propios errores.

Factores de forma. Para el estudio de las variaciones en el rendimiento humano, se definen los llamados factores de forma del comportamiento (PSF, Performance Shaping Factor), cuya relación simplificada (Creus, 1992) es la siguiente:

- . Situación del entorno de trabajo
- . Forma de las instrucciones de trabajo
- . Circunstancias de la tarea
- . Factores de estrés psicológico
- . Factores de estrés fisiológico
- . Factores orgánicos

Tipos de fallos. En las operaciones que realiza el hombre pueden presentarse los siguientes tipos de fallos:

1. *Error de omisión.* Por olvido de alguna acción a efectuar.
2. *Error de comisión.* Consistente en realizar mal alguna acción y que puede producirse como: *error de tiempo*, cuando la acción no se realiza en el momento adecuado; *error de selección* del control erróneo o elección de procedimiento incorrecto; *error de secuencia*, acciones llevadas a cabo en orden incorrecto; *error cualitativo*, acción llevada a cabo en mucho o en poco.
3. *Errores de tipo extraño.* Cualquier otro caso no asimilable a los anteriores.

Métodos específicos de análisis del error humano

Para el análisis de la seguridad de los sistemas hombre-máquina el primer intento sería aplicar los métodos de estudio de la fiabilidad técnica al estudio de la fiabilidad humana; sin embargo eso sería olvidar las características del ser humano descritas anteriormente. Por otra parte, es necesario considerar la *variabilidad de conducta* de la persona (en función de su resistencia a la fatiga, sus características psicológicas, afectivas y cognitivas, sus actitudes, su formación y su experiencia) y la *variabilidad interpersonal*, que pueden manifestarse como elementos de infiabilidad difícilmente controlables.

Todo ello hace que las posibilidades de transferir los métodos de la fiabilidad técnica a la fiabilidad humana sean limitadas y resulte necesario emplear métodos específicos.

Cuantificación de los errores humanos

Como hemos indicado, el error humano se presenta cuando el comportamiento humano o su influencia sobre el sistema exceden el límite de aceptabilidad prefijado, su cuantificación puede hacerse mediante la llamada *tasa de error*:

$$\lambda = \frac{N_{ED}}{N_D}; \text{ o bien } \lambda_D = \frac{N_{EP}}{T}$$

donde

N_{ED} = número de errores por demanda

N_D = número de demandas

N_{EP} = número de errores en el período de observación

T = período de observación

En los próximos apartados describimos los métodos de análisis y valoración del error humano más destacados:

THERP

Esta técnica (iniciales de *Technique for Human-Error Rate Prediction*) fue ideada por Swain, Rook y su equipo en el Laboratorio Sandia en 1962. La técnica ha evolucionado y ha sido perfeccionada (Swain y Guttman, 1983; Bladh, Holmberg, y Pyy, 2009), de tal modo que constituye un estándar mayoritariamente aceptado.

La hipótesis de base de THERP es que se puede considerar a la persona como fuente de fallos, de manera similar a un dispositivo técnico. Las tareas de la persona se descomponen en actividades elementales y su secuencia se visualiza en un *árbol de eventos*, juntamente con sus posibles desviaciones en forma de error de omisión o de comisión del operador. El evento básico que representa el error humano puede representarse como un subconjunto de nudos del árbol; de modo que para calcular su probabilidad basta multiplicar las probabilidades que se encuentran a lo largo del camino que conduce hacia él.

La técnica THERP es ampliamente aceptada en diversos campos de actividad por su capacidad de integrarse en un análisis probabilístico de riesgos y ha influido en el desarrollo de otras técnicas de análisis; pero presenta dificultades, ya que es de aplicación muy laboriosa y los análisis tienen un componente de subjetividad importante, a causa de las múltiples elecciones de valores que se deben realizar en tablas de datos, de modo que frecuentemente no hay uniformidad entre distintos analistas en la identificación de errores y en los PSF a considerar.

Las críticas más fuertes (Arquer, 2007; Quesada, 1999) se centran en dos aspectos:

- En el desglose de la tarea en acciones elementales, porque ello no siempre es coherente con la organización psicológica de la tarea.
- En el origen y la aplicabilidad a situaciones diversas de los datos que se dan en las tablas.

Los críticos de orientación cognitiva exigen un modo de análisis del error más afianzado teóricamente y son escépticos sobre la asignación de probabilidades razonables a acciones relacionadas con situaciones muy específicas. Además, para las tareas de diagnóstico o que requieran un cierto nivel de decisión, esta técnica no permite identificar recomendaciones para disminuir la probabilidad de fallos humanos de este tipo. Sólo una persona experta en fiabilidad humana puede determinar con criterio las modificaciones ergonómicas necesarias para incidir sobre las causas psicológicas de algunos errores.

SHARP

(*Systematic Human Action Reliability Procedure*) Es un método (Hannaman y Spurgin, 1984) que define una estructura basada en la experiencia y el análisis sistemático de las interacciones. El método aplica siete pasos, que pueden verse en la figura 1 con sus interacciones.

Paso 1: Definición. Se trata de identificar y asegurar que todas las acciones humanas importantes se han incluido en el análisis. Para ello, en primer lugar se clasifican como:

- Tipo 1: Relacionadas con pruebas y mantenimiento, anteriores al incidente.
- Tipo 2: Acciones que, a partir de un error, generan anomalía del sistema.
- Tipo 3: Acciones con las que, de acuerdo a un procedimiento, el operador puede recuperar control del sistema.

- Tipo 4: Acciones incorrectas tras un incidente, que empeoran la situación.
- Tipo 5: Durante la secuencia de un incidente, acciones del operador, no procedimentadas, con las que puede recuperar el control.

Paso 2: Selección. Consiste en elegir, de entre las acciones identificadas en el paso anterior, las más importantes para la seguridad del sistema, en dos fases: cualitativamente al principio del análisis y cuantitativamente después.

Para la determinación cuantitativa se aplicarán unos valores preliminares o genéricos, en función del tipo de tarea y del nivel de actuación requerido (rutina, reglas o conocimiento). Se seleccionarán para el análisis en los pasos siguientes aquellas acciones humanas que aparezcan con una contribución al fallo del 1 % o más.

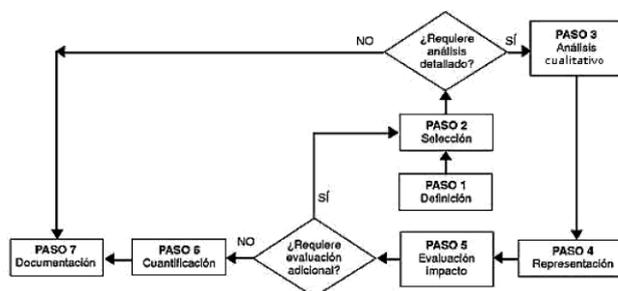


Figura 1. Procedimiento SHARP de análisis de fiabilidad humana.

Paso 3: Análisis cualitativo. Los dos objetivos de este paso, son: a) detallar los conocimientos del operador y qué tipo de acciones puede realizar en una determinada situación; b) identificar cómo el comportamiento de un operador puede ser modificado por una disfunción del sistema.

Paso 4: Representación. Para la inclusión de los resultados del análisis de fiabilidad humana en el análisis global del sistema, los modos de fallo humano requieren una representación de las acciones, que refleje cómo pueden afectar a la operación del sistema (Faig, 2007).

Paso 5: Evaluación del impacto. Las acciones del paso anterior deben revisarse con el objeto de verificar que la representación no afecta a ningún suceso más en el modelo.

Paso 6: Cuantificación. El propósito ahora es establecer, en todas aquellas acciones a las que se le haya requerido un análisis detallado, el valor de la probabilidad de fallo para cada acción humana. Para ello (Faig, 2007):

- Para los errores tipo 1, se aplicará la metodología THERP.
- Para los tipo 3, 4 y 5, la cuantificación se realizará mediante el modelo TRC (ver más adelante) para la parte cognoscitiva y THERP para la manual.
- Para acciones fuera del ámbito de aplicación de los modelos mencionados, se utilizarán bases de datos, estimaciones de expertos, o la comparación con análisis anteriores.

Paso 7: Documentación. La metodología SHARP postula documentar adecuadamente los pasos realizados, de forma que se asegure la trazabilidad de las hipótesis, datos, y modelo seleccionado.

SHERPA

(Systematic Human Error Reduction and Prediction Approach)

Esta otra técnica (Embrey, 1986) pretende evaluar cualitativa y cuantitativamente la fiabilidad humana y elaborar recomendaciones concretas para reducir la probabilidad de errores humanos, especialmente en lo que se refiere a procedimientos, formación de personal y diseño de equipos.

Está basada en el análisis funcional de la conducta humana según el modelo de Rasmussen (ver más adelante) y se desarrolla en varios módulos de análisis:

Análisis jerárquico de las tareas. Se trata de un método sistemático de identificación de las diferentes subtareas, de sus objetivos respectivos y de la forma en que estos objetivos se combinan para alcanzar un objetivo global. Se obtiene entonces una lista detallada de actividades y objetivos elementales que pueden clasificarse en cuatro categorías de procesos mentales:

1. Basado en habilidades o automatismos.
2. De diagnóstico basado en reglas.
3. Basado en reglas del tipo: si x , entonces y .
4. Basado en conocimientos.

La técnica predictiva de los errores sólo es aplicable para las tres primeras categorías; para la cuarta, están en fase de desarrollo otras técnicas.

Análisis de los errores humanos. Consiste en una búsqueda sistemática de los posibles errores, en varias etapas:

1. Definición de los pasos de una tarea que no pueden ser omitidos o modificados, sin perjudicar el objetivo.
2. Análisis de la fiabilidad de la tarea.
3. Análisis del impacto del error sobre los elementos relacionados con él.
4. Recomendaciones para reducir los errores.

Cuantificación. Este módulo es discrecional. Cuando es suficiente con una evaluación cualitativa, la técnica SHERPA puede ir directamente al módulo de reducción de errores. El análisis cuantitativo se realiza mediante *Success Likelihood Index Methodology (SLIM)*, con la que se obtienen índices de probabilidad de éxito para las tareas (considerando los PSF correspondientes):

$$\log(p) = a \cdot SLI + b$$

donde:

$$a = \frac{\log(p_1 / p_2)}{SLI_1 - SLI_2}$$

$$b = \log(p_1) - a \cdot SLI_1$$

SLI_1 , SLI_2 , p_1 y p_2 son parámetros tabulados correspondientes a situaciones límite tomadas como referen-

cia. SLI es el índice de fiabilidad, calculado como se expone más adelante, en la técnica SLIM.

Reducción de los errores. Los resultados del análisis jerárquico de tareas y del análisis de los errores humanos deberían posibilitar la propuesta de recomendaciones para mejorar los procedimientos de trabajo, la formación y adiestramiento del personal, el diseño del equipo, etc.

Valoración. Al ensayo de las recomendaciones le debe seguir el control de su eficacia mediante una fase de evaluación del resultado.

La técnica SHERPA se ha aplicado en sectores de actividad de alto riesgo (nuclear) y en sectores de servicios (correos). Sus principales ventajas son:

- La aplicación de *modelos cognitivos* facilita la comprensión y reducción de los errores.
- Su orientación hacia la propuesta de recomendaciones.
- Trata el error de modo positivo, como elemento de información.

Esta técnica resulta adecuada para tareas en las que la actividad humana está en el nivel de funcionamiento basado en *habilidades* (automatismos) o en la aplicación *reglas*. Sin embargo, los errores que se producen en el nivel de *conocimientos* son poco predecibles. Otras limitaciones son:

- Escasa madurez de algunos módulos.
- Su inserción en el análisis probabilístico de riesgos está aún en desarrollo.
- No consideración de los efectos temporales.
- Los resultados están muy influidos por el criterio del analista.

Juicio de expertos

Es una técnica aplicable junto con cualquier otra, como método alternativo al uso de tablas para la estimación de probabilidades de error humano. Se recurre a ella cuando, para conocer la probabilidad de error, hay que considerar la influencia de muchos factores interdependientes que afectan a la respuesta de la persona; también se emplea cuando en las bases de datos no se encuentran puntos de referencia aplicables. Combina las aproximaciones analíticas y los métodos de estimación subjetiva de probabilidades desarrollados conforme a las reglas de la teoría de la decisión.

Los resultados del juicio de expertos pueden estar afectados por algunos sesgos de origen diverso. Principalmente por la mayor facilidad para considerar los sucesos o errores más conocidos y por la posibilidad de que uno de los expertos arrastre a los demás hacia su propia opinión. También pueden surgir conflictos por motivos personales.

TRC

(Time Reliability Curve) Se trata de una técnica (Dougherty y Fragola, 1988) especialmente aplicable a las acciones post-accidente. Se basa en suponer que la disponibilidad de tiempo, o la posibilidad del operador de poder pensar, es el factor determinante en la probabilidad de fallo de la acción.

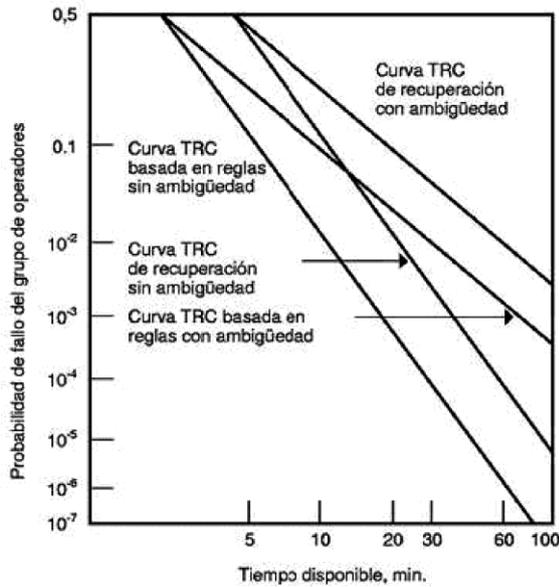


Figura 2. Curvas nominales en el modelo TRC

La figura 2 muestra las curvas de correlación tiempo-probabilidad de la TRC. Existen dos pares de curvas, un par se aplica a los errores basados en reglas, por ejemplo en diagnóstico incorrectas, mala previsión o incorrecta toma de decisiones en el seguimiento de las reglas (procedimientos), una de las curvas se utiliza cuando no hay ambigüedad, la otra cuando la toma de acción se realice con dominio de las dudas. El segundo par de curvas se utiliza cuando la diagnosis se realiza en ausencia de reglas, este es el caso de recuperaciones después de un incidente no procedimentadas.

HCR

(Human Cognitive Reliability model). Trata también (Han-naman, Spurgin, y Lukics, 1885) de una correlación tiempo-fiabilidad en las acciones de diagnóstico o identificación. Para ello, maneja tres tipos de eventos (A1, diagnóstico correcta; A2, diagnóstico errónea; A3: falta de diagnóstico o diagnóstico tardía), de forma que la suma de probabilidades:

$$P(A1) + P(A2) + P(A3) = 1$$

La función de distribución del tiempo requerido para realizar una tarea es representada mediante una distribución de Weibull (ver modelo psicológico de Rasmussen). Y la probabilidad de error en función del tiempo es:

$$P(t) = e^{-\left(\frac{t-a}{T_{0,5}-a}\right)^b}$$

donde:

(a, b, c) = coeficientes representativos del tipo de comportamiento.

Por su parte, el tiempo promedio $T_{0,5}$ resulta de un nominal estimado $T_{0,5/nominal}$ y modificado por tres parámetros de forma (PSF):

$$T_{0,5} = T_{0,5/nominal} (1+K_1) (1+K_2) (1+K_3)$$

donde K_1 , K_2 y K_3 son, respectivamente, los coeficientes para la experiencia del operador, para nivel de estrés y para la calidad de la interfaz operador/planta.

HCR/ORE

(Human Cognitive Reliability / Operator Reliability Experiments). Es una revisión (Spurgin, 1990) de la técnica HCR, basada en experimentos de fiabilidad de operadores e incluye la posibilidad de cuantificar acciones de recuperación de errores.

SLIM

(Success Likelihood Index Methodology). Este método (Embrey, Humphreys, Rosa, y Rea, 1984) sólo tiene en cuenta factores de forma (PSF) para ajustar las probabilidades de error; en particular, el tiempo que dispone el operador para realizar una tarea es también tratado como PSF.

La suma ponderada de esos PSF da un índice de fiabilidad, denominado SLI:

$$SLI = \sum_{i=1}^n (w_i x_i)$$

con

$$\sum_{i=1}^n w_i = 1$$

donde:

x_i = grado o efecto de incremento de fiabilidad que produce el PSF i en la acción investigada;

w_i = peso relativo del PSF i ;

n = número de PSFs relevantes.

Los cálculos de las probabilidades se realizan en base a dos modelos:

1. Calculando la probabilidad q de no fallo ($q=1-p$):

$$\log(q) = a \cdot SLI + b$$

$$a = \frac{\log((1-p_1)/(1-p_2))}{SLI_1 - SLI_2} ; b = \log(1-p_1) - a \cdot SLI_1$$

2. O calculando la probabilidad p de fallo:

$$\log(p) = a \cdot SLI + b$$

$$a = \frac{\log(p_1/p_2)}{SLI_1 - SLI_2} ; b = \log(p_1) - a \cdot SLI_1$$

SLI_1 , SLI_2 , p_1 y p_2 son parámetros tabulados correspondientes a situaciones límite tomadas como referencia válida para el caso en estudio.

INTENT

Trata (Gertman, Blackmann, Haney, Seidler, y Hahn, 1992) de la estimación de probabilidades de error humano en casos de errores basados en procesos de decisión (errores de intención), como complemento de otros métodos que no consideran este tipo de procesos. INTENT tabula 20 tipos

de errores nominales, para cada uno de los cuales aporta un límite inferior (LB) y un límite superior (UB) de probabilidad de ocurrencia. Adicionalmente, se definen 11 factores de forma (PSFs).

De forma similar al método SLIM, la toma en consideración ponderada de los PSFs lleva al cálculo de un índice de fiabilidad, SLI. No obstante, difiere en el cálculo de la probabilidad de error: INTENT asume que p es igual al cuantil $(1-SLI)100\%$ de una distribución log-normal en la que LB es el cuantil 5% y UB es el cuantil 95%.

$$\ln(p) = \sigma \cdot \Phi^{-1}(1 - SLI) + \mu$$

$$\sigma = \frac{\ln(UB/LB)}{3,29} \quad ; \quad \mu = \frac{\ln(UB \cdot LB)}{2}$$

OAT

(*Operator Action Type / Operator Action Tree*). Es un método (Hall, Fragola, y Wreathall, 1982) que caracteriza cada una de las partes de la acción humana y sirve para cuantificar los errores, en particular cuando el tiempo desempeña un papel importante. El método se basa en modelos de acciones humanas ligadas a la detección del evento, su razonamiento, su diagnóstico y a la respuesta requerida.

EdF's PHRA

(*Method of Electricité de France for Probabilistic Human Reliability Analysis*). Esta técnica (Mosneron, 1989) da prioridad al análisis cualitativo más que a la cuantificación de las probabilidades de error. No obstante, esta cuantificación es también tenida en cuenta como función del tiempo disponible, mediante curvas basadas en datos de simulador, para acciones de diagnóstico.

Para acciones post-diagnóstico, aplica el llamado "modelo de ejecución":

$$P_E = P_B \cdot K_F \cdot P_{NR}$$

donde:

P_E = probabilidad de que falle la ejecución de la acción;

P_B = probabilidad básica de fallo;

K_F = factor para considerar el contexto de la acción;

P_{NR} = probabilidad de que falle la recuperación.

HEART

(*Human Error Assessment and Reduction Technique*). Es otro método (Williams, 1988) basado exclusivamente en el tratamiento de los PSF (aquí llamados EPC, error producing conditions) y se utiliza fundamentalmente para el análisis a nivel de tareas complejas, más que en el nivel de acciones elementales. El modelo es multiplicativo y tiene la siguiente forma:

$$p = p_0 \prod_{i=1}^n [1 + (K_i - 1)x_i]$$

donde:

p = probabilidad de error humano en el evento;

p_0 = probabilidad nominal;

K_i = máximo efecto multiplicativo del EPC i ;

x_i = tasa estimada de influencia del EPC i ;

n = número de EPCs considerados.

Resumen comparativo

A modo de resumen comparativo, hemos compuesto el Cuadro 1, indicando las ventajas y limitaciones que, sobre las distintas técnicas de análisis del error humano, se ponen de manifiesto en la literatura técnica.

Modelos de representación del comportamiento

En cada una de las técnicas descritas en los apartados anteriores subyace una forma de entender el comportamiento humano, más o menos mecanicista, que está siendo criticada desde la Psicología, por su escasa profundización en las causas que desencadenan el error humano. Por ello, en este punto vamos a repasar los distintos modelos de representación del comportamiento, para exponer sus ventajas e inconvenientes cara al análisis de la fiabilidad humana.

Los modelos para representar la actividad o el comportamiento humano pueden clasificarse, en primer lugar, en *parciales* y *generales*. Los primeros se limitan a analizar una actividad concreta, por ejemplo la detección de una señal, mientras que los segundos intentan analizar en conjunto todas las causas posibles del error y hacer propuestas de mejora. Estos últimos los separaremos a su vez entre los basados en teorías psicológicas y los basados en el análisis de la tipología de los procesos humanos.

Modelos parciales

Se limitan, como hemos dicho, a analizar una actividad elemental concreta entre las que podemos considerar, como paradigma, el *modelo de la detección de señales* que distingue dos tipos de error:

1. Omisión: una señal presente no ha sido percibida y
2. Falsa alarma: una señal ausente es percibida como presente.

Presupone un mecanismo de respuesta de la persona que conjuga la discriminabilidad de la señal y el criterio de decisión adoptado.

Para la estimación de las probabilidades de error de las actividades elementales se usan tablas de probabilidades, obtenidas a partir de juicios de expertos y/o de datos históricos. El instrumento de base para el análisis, cuando la actividad puede representarse en tareas o actividades más simples (p.e. en la metodología THERP), es un tipo de *árbol de sucesos* llamado *diagrama en árbol de probabilidades*.

Cuadro 1. Resumen de técnicas de análisis del error humano.

Técnica	Ventajas	Limitaciones	Observaciones
THERP	Estándar aceptado. Integrable en un análisis de probabilidades de riesgos. Amplia base de datos disponible.	Importante componente de subjetividad. Excesivo énfasis en los errores externamente observables. Difícil aplicación de los datos en tablas a situaciones distintas. No permite hacer recomendaciones de mejora para errores de diagnóstico.	Aplicable a todos los sectores industriales.
SHARP	Adaptable a la experiencia de los analistas. Buen tratamiento de las interacciones. Buena integración en el análisis probabilístico de riesgos.	Sólo la pueden aplicar expertos en la actividad analizada.	Marco útil para la integración de diversos métodos.
SHERPA	Utilización sencilla para no especialistas. Uso de modelos cognitivos. Orientado a la propuesta de recomendaciones.	Escaso tratamiento de los errores basados en conocimientos. Falta de madurez. Escasa integración en el análisis de probabilidades de riesgos. No considera los aspectos temporales. Excesiva subjetividad.	Se ha aplicado, tanto en sectores industriales, como en servicios.
Juicio de expertos	Sencillez de uso. Posibilidad de recoger información complementaria para proponer mejoras.	Posibilidad de sesgo por influencia de unos expertos en otros.	Aplicable a todo tipo de actividades simples. Imprescindible en ausencia de datos asimilables.
TRC	Considera el factor tiempo. Integrable como parte de cualquier otra técnica.		Aplicable a acciones post-accidente.
HRC	Aplicación del modelo cognitivo de Rasmussen.	No hay amplia aceptación sobre los parámetros aplicables. Pendiente de mayor desarrollo.	
SLIM	Orientado al tratamiento de los PSF. Flexible.	Poco validada. Desprecia las interrelaciones entre PSFs. Precisa de sofisticadas estimaciones de expertos.	
OAT	Técnica pionera en el tratamiento de errores de diagnosis y en el efecto temporal.	Integración poco clara en el análisis probabilístico de riesgos.	
HEART	Aplicación simple.	Aislamiento excesivo de las distintas tareas.	
INTENT	Especializada en errores de decisión.	Falta de madurez.	

La consideración de la influencia de algunos factores (PSF) sobre el desempeño de la persona es la única manera en que las técnicas basadas en el análisis de tareas tienen en cuenta la naturaleza humana del individuo, en tanto que fuente de error. Estos factores (experiencia, nivel de estrés, etc.) se utilizan para modificar las probabilidades nominales de error. Específicamente la metodología THERP ofrece veintisiete tablas de probabilidad de errores humanos cuyos valores son probabilidades nominales.

En resumen, la probabilidad $P(E)$ de un error elemental se obtiene por la fórmula:

$$P(E) = P_1 \cdot K \cdot P_2$$

en la que:

P_1 = Probabilidad nominal de error.

K = Coeficiente de corrección.

P_2 = Probabilidad de no recuperación del error.

Para tareas de diagnóstico, se establece la probabilidad de fracaso en función del tiempo disponible para realizarlo. Se trata de un cálculo sencillo, pero sin información sobre la causa del error y

partiendo del supuesto, no siempre cierto, de que el carácter anómalo de la situación es detectado de forma evidente. Por otra parte, la obtención de información y la elaboración de bases de datos útiles presentan problemas ocasionados, entre otras razones, porque:

- sólo se suele informar de los errores humanos que se traducen en consecuencias no deseables; sería necesario asegurar el registro de todos los errores;
- existen limitaciones a causa de las diferencias entre unas situaciones y otras;
- puede ser difícil definir el límite exacto entre la conducta humana adecuada y la errónea.

Una alternativa para soslayar en parte las limitaciones del uso de bases de datos consiste en la estimación de probabilidades de error mediante juicios de expertos.

Modelos generales basados en teorías psicológicas

Rasmussen. El modelo de Rasmussen es, entre los mode-

los generales, el más utilizado (casi el único) y distingue *ocho etapas en el proceso de tratamiento de la información y toma de decisiones*: activación, observación, identificación, interpretación, evaluación, selección de objetivo, elección del procedimiento y ejecución. También diferencia *tres niveles de comportamiento* de la persona, según base su actividad en: automatismos, reglas o procedimientos y conocimientos. Dichos niveles corresponden a grados decrecientes de familiaridad con el entorno y la tarea.

En la figura 3 se representa el esquema de 8 etapas del tratamiento de la información y se indica los saltos que se dan entre ellas, en función del nivel de comportamiento: cuando una persona tiene un comportamiento basado en automatismos pasará de las etapas de “activación y observación” directamente a la de “ejecución”; cuando tiene un comportamiento basado en las reglas y procedimientos, el salto hacia la “ejecución” se hará desde la etapa de “identificación” y, por último, cuando el comportamiento se basa en los conocimientos, la persona recorre las ocho etapas mencionadas.

La ventaja de este modelo cognitivo es que conjuga la identificación del origen funcional del error (etapa de tratamiento) con el nivel de funcionamiento implicado y así permite reparar los errores por etapas de tratamiento y proponer medidas de prevención adaptadas al nivel de comportamiento de la persona.

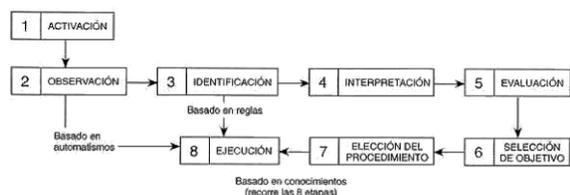


Figura 3. Etapas de tratamiento de información y toma de decisiones según Ramussen.

Norman y Shallice. D. Norman y Shallice (1980) quisieron establecer una teoría de la acción humana que fuese adecuada tanto para explicar la ejecución correcta como las variedades más predecibles del error humano. Las formas sistemáticas de error y la ejecución correcta se enfocan como caras de la misma moneda.

El modelo se denomina ‘Atención para la acción’ y se pensó básicamente para explicar las diferencias entre el procesamiento automático y el controlado. Propone dos componentes para describir los distintos niveles de control necesarios para llevar a cabo varias secuencias de acción o pensamiento. El primero, la *agenda de contenidos*, controla la ejecución de actos rutinarios y habilidades sobreaprendidas: los “esquemas de acción”. Éstos son conjuntos de respuestas asociadas con un estímulo ambiental específico para el que el esquema es activado. El segundo componente es el *sistema supervisor*, la unidad de planificación general necesaria para resolver problemas no rutinarios, proveyendo un control atencional consciente para modular la ejecución. Se activa cuando el proceso de selección por la agenda de contenidos falla o cuando un es-

quema apropiado no es viable.

Norman y Shallice proponen que el sistema supervisor debería ser activado para:

1. Situaciones que involucren planificación o toma de decisiones.
2. Situaciones que involucren corrección de errores.
3. Situaciones donde la respuesta es nueva y no está bien aprendida.
4. Situaciones que se juzgan como difíciles o peligrosas.

En el modelo, el *procesamiento automático* puede referirse a procesos en los cuales la selección de esquemas implica tan poca activación que la conducta puede ser realizada en paralelo con otras actividades. De manera alternativa, se puede referir a procesos controlados por su propio esquema individual, disparados directamente por algún estímulo. El *procesamiento controlado* podría corresponder a procesos que requieren seleccionar más de un esquema de la agenda de contenidos. No obstante, dado que un esquema sobreaprendido puede, según la teoría, ser seleccionado mientras esté funcionando un segundo esquema más potentemente activado, la selección puede ocurrir en un continuo de valores de activación. Por lo tanto, la distinción entre procesos automáticos y controlados se convierte en una cuestión de grado más que en una dicotomía.

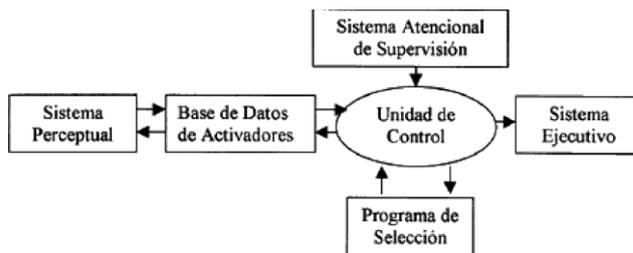


Figura 4. Esquema de atención para la acción de Norman y Shallice.

Las explicaciones para los errores derivadas del marco de Norman y Shallice se basarían en que los procesos automatizados pueden prescindir, y de hecho prescinden en muchas ocasiones, del control consciente, y esto puede hacer que las situaciones que son novedosas por alguna característica sutil sean tratadas, erróneamente, como situaciones conocidas.

Teorías basadas en el concepto de esquema. Bartlett (1931) invocó la noción de esquema para explicar los errores sistemáticos en el recuerdo de material pictórico y verbal. Las reproducciones hechas de memoria de esos materiales eran más regulares, más significativas y más convencionales que los dibujos o las historias originales. Las características poco comunes del material a recordar se sometían a las expectativas y hábitos de pensamiento de la persona.

Enfatizó tres aspectos fundamentales de los esquemas:

1. que eran estructuras mentales inconscientes,
2. que estaban compuestas de conocimiento (de experiencias pasadas) y
3. que la memoria a largo plazo consiste en estructuras ac-

tivas de conocimiento y no en imágenes pasivas.

Por tanto, los esquemas, más que reproducir, reconstruyen las experiencias pasadas. Y este proceso lleva a unas predecibles predisposiciones en el recuerdo.

En la década de los 70, las ideas de Bartlett se retomaron desde distintos campos, como la 'visión computerizada', la 'interpretación de historias' y el 'aprendizaje de habilidades motoras'. Común a estas teorías es la idea de que los esquemas son estructuras de conocimiento de alto nivel que contienen 'slots informacionales' o variables. Cada 'slot' sólo aceptará un tipo particular de información. Si los 'inputs' que provienen del mundo en ese momento no proporcionan la información específica para esos 'slots', éstos toman valores por defecto: valores estereotípicos derivados de la experiencia previa con el ambiente en situaciones parecidas. Esta idea de volver a los valores por defecto es central en la explicación de los errores humanos.

Para estas teorías, la rapidísima manipulación de la información característica de la cognición humana es posible porque las regularidades del mundo y nuestras rutinas para afrontarlas han sido representadas internamente como esquemas. El precio que pagamos por este procesamiento automático de la información es que las percepciones, la memoria, los pensamientos y las acciones tienen tendencia a errar en la dirección de lo familiar y lo esperado.

Modelo de Norman

Una teoría específica para explicar errores humanos, basada en el concepto de esquema, es la de D. A. Norman (1981), llamada *esquema desencadenado por activación* (EDA). En ésta una secuencia de acción es representada por un esquema padre y varios esquemas hijos, entre los cuales puede haber activos varios esquemas de acción al mismo tiempo. Cada esquema tiene un conjunto de características de activación y un valor de activación. El camino de la intención a la acción consiste en la activación del esquema padre que corresponde a la intención, la activación de los esquemas hijos para las partes que componen la acción, y después el desencadenamiento de los esquemas cuando se dan las condiciones que requieren para operar. Este sistema permite organizar los errores en tres categorías mayores:

1. Errores en la formación de la intención.
2. Fallos en la activación de los esquemas.
3. Errores por fallo en el desencadenamiento del esquema.

Modelos basados en la tipología de los procesos humanos

Este otro modelo humano se funda en el análisis básico de los procesos que realiza el hombre:

- . Procesos de tipo cognoscitivo, con una toma de decisiones de "alto nivel".
- . Procesos de ejecución, que precisan la toma de decisiones de "bajo nivel".
- . Errores cometidos en cualquier nivel y que dependen de

las condiciones ambientales, estrés, etc.

Para representar estos procesos de decisión y de ejecución se utilizan con frecuencia lenguajes de inteligencia artificial (p.e, LISP), siendo la simulación de los mecanismos de error humano objeto de estudio especial, debido a su dificultad.

Resumen y conclusiones

Los modelos del comportamiento humano revisados anteriormente, y especialmente los denominados modelos parciales, son el soporte teórico de diversas técnicas (THERP, SHARP,...) que se usan para el análisis y cuantificación del error humano. Sin embargo, los llamados modelos generales, con la excepción del de Rasmussen, tienen todavía una aplicación incipiente en esas técnicas de análisis.

Modelos parciales

Hemos visto que los *modelos parciales*, influidos por la tradición 'mecanicista', proceden diseccionando las tareas y limitando el análisis a cada actividad elemental concreta, a la que se asocian los tipos de error que le son significativos y que, por la propia naturaleza elemental de la actividad, serán pocos y aceptablemente bien caracterizados. Considerados como procesos estocásticos, para cada tipo de error se busca su probabilidad de ocurrencia, aplicando, según sea el caso, tablas o registros históricos, simulación o juicio de expertos.

Este tratamiento mecanicista hemos visto también que admite ajustes orientados, por un lado, a la consideración de los denominados *factores de forma* (PSF) y, por otro, a la consideración del factor tiempo; sujetos, en cualquier caso, a la apreciación del analista.

También se han descrito los problemas para obtener datos que sean útiles:

- . por la dificultad de asimilar situaciones distintas,
- . por la necesidad de disponer de grandes períodos de tiempo y
- . por la dificultad de encontrar el límite exacto entre conducta adecuada y errónea.

Pero la crítica fundamental es que, como resultado de la visión parcial del comportamiento humano, proceden sobre la ignorancia o el tratamiento superficial de los elementos cognitivos que, tal como proclaman las teorías psicológicas, están en la base de los denominados errores de comisión; en consecuencia, no aportan información sobre la causas del error y ayudan poco al objetivo final, que no es otro que reducir su probabilidad de ocurrencia.

Modelos generales

El recurso a *modelos generales* busca un marco teórico que englobe los procesos cognitivos presentes en el comportamiento humano; permitiendo así, no solo tipificar los errores, sino comprender el mecanismo que los desencadena.

El modelo de Rasmussen, basado en ocho etapas de tra-

tamiento de la información y toma de decisiones y tres niveles de comportamiento, es el más utilizado entre los modelos generales y conjuga la identificación del origen funcional del error (etapa de tratamiento) con el nivel de funcionamiento implicado (basado en conocimientos, en reglas, o automático).

Norman y Shallice proponen una teoría de la acción humana adecuada tanto para explicar la ejecución correcta como las variedades habituales del error humano. Se denomina 'Atención para la acción' y postula la existencia de dos componentes para describir los distintos niveles de control: la *agenda de contenidos*, que controla la ejecución de actos rutinarios y habilidades sobrepasadas, denominadas 'esquemas de acción', y el *sistema supervisor*, que es la unidad de planificación general necesaria para resolver problemas no rutinarios.

Por otro lado, hay un grupo de teorías que profundizan en el concepto de esquema, enfatizando, entre otros aspectos, que la memoria a largo plazo se compone de estructuras activas de conocimiento, en vez de imágenes pasivas. Como ejemplo de estas teorías, hemos señalado la de Norman, denominada *esquema desencadenado por activación* (EDA). Describe el comportamiento humano como un proceso de activación y desencadenamiento de esquemas.

Por último, hemos mostrado la simplificación que representan los modelos basados en la tipología de los procesos humanos. Estos modelos distinguen entre procesos de tipo

cognoscitivo y procesos de ejecución. La ventaja de estos modelos es que admiten representación mediante lenguajes de inteligencia artificial.

Como crítica global del conjunto de los modelos generales, hemos puesto de manifiesto la escasez de desarrollo para constituir métodos aplicables al entorno industrial; por lo que existe un amplio campo de trabajo por hacer en el desarrollo y perfeccionamiento de técnicas basadas en estos modelos.

Cuadro resumen

Por tanto, podemos resumir que, aunque los llamados modelos globales de comportamiento humano atienden a la cuestión de explicar el origen del error, que es el principal defecto de los modelos parciales, el uso de éstos sigue siendo, hoy por hoy, imprescindible, puesto que son la base de técnicas que, como THERP, están bien aceptadas en la industria, y especialmente en las industrias de alto riesgo.

En el cuadro 2 puede verse una comparación esquemática entre los distintos modelos de comportamiento, su relación con las técnicas en vigor de análisis del riesgo humano y una referencia sucinta a sus debilidades y fortalezas en orden a dar un soporte consistente a esas técnicas.

Cuadro 2. Resumen de modelos.

Modelo de representación del comportamiento	Técnicas de aplicación	Fortalezas y debilidades
Modelos parciales; p.e.: detección de señales.	Arboles de sucesos, aplicados en THERP, OAT... Arboles de fallos, aplicados en SHARP. Estimaciones de expertos.	Madurez y aceptación. No buscan la causa funcional del error.
Rasmussen	Distribución de Weibull (aplicable en todas las técnicas). SHERPA. HCR.	Explicación del origen funcional del error. Buenas posibilidades de integración en técnicas actuales. Dificultad de ajuste de los coeficientes.
Atención para la acción (Norman y Shallice)	SHERPA (análisis cualitativo).	Explicación del error con el mismo mecanismo que la acción correcta. Poca integración en técnicas cuantitativas actuales.
EDA (Norman)	En desarrollo.	Explicación del origen funcional del error y de sus relaciones con los factores ambientales. Escasa integración en técnicas actuales.
Tipología de procesos	Inteligencia artificial.	Buenas posibilidades de modelado mediante esquemas lógicos. Incompleta explicación de causas de error.

Conclusiones y líneas de investigación

Del conjunto de experiencias e investigaciones publicadas hasta el momento destaca el hecho de que es la interacción hombre-máquina el factor clave a considerar en la mejora de la fiabilidad de los sistemas industriales, aunque resulta mucho más fácil analizar el comportamiento de la máquina que el del hombre.

En consecuencia, la principal conclusión que extraemos es que, para mejorar el cálculo de la probabilidad de error, habría que incorporar los aspectos cognitivos del

comportamiento humano y las múltiples variables que lo determinan.

Líneas de investigación. El ámbito de las metodologías de análisis del error humano se trata de un campo multidisciplinar en el que se mantienen abiertas muchas líneas de investigación que podrían resumirse en dos grupos:

1. *Investigación metodológica.* Se ocupa de la validación y mejora de los actuales métodos de análisis de fiabilidad humana, sin poner en cuestión sus fundamentos teóricos, y tratan de buscar solución a los siguientes problemas:

. Las diferencias entre los resultados cuantitativos de dife-

rentes analistas (usando el mismo método) o de diferentes métodos.

- . El tratamiento superficial que se hace de las dependencias entre acciones.
- . El tratamiento superficial de los errores de diagnóstico o errores de decisión.
- . Falta de tratamiento de las dependencias entre factores de forma (PSF).

2. *Investigación sobre modelos.* Persigue ampliar el ámbito de validez de los análisis de fiabilidad, mediante la propuesta de nuevos métodos y/o nuevos enfoques. Un objetivo impor-

tante de estas investigaciones consiste en extender el ámbito de los estudios probabilísticos de riesgos a los llamados 'errores de comisión'. Con ello se pretende suplir el defecto de los actuales métodos de análisis (árboles de eventos y árboles de fallos), bien adaptados a los 'errores de omisión', pero no a los de 'comisión'.

Agradecimientos.- Los autores quieren agradecer la revisión y sugerencias realizadas por el prof. Vicente Bargueño (ETS de Ingenieros Industriales, UNED) en el primer borrador de este trabajo.

Referencias

- Apostolakis, G., Bier, V., y Mosleh, A. (1988). A critique of recent models for human error rate assessment. *Reliability Engineering & System Safety*, 22, 201-217.
- Arquer, M. I. de. (2007). *Fiabilidad humana: métodos*. <http://www.insht.es>, extraído el 23/09/07. INSH. (NTP 377)
- Bartlett, F. (1931). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Birolini, A. (2007). *Reliability engineering. Theory and practice*. New York: Springer.
- Bladh, K., Holmberg, J.-E., y Pyy, P. (2009). *An evaluation of the enhanced bayesian Therp method using simulator data*. Taylor & Francis Group. Disponible en <http://www/inf/julkaisut/muut/2009/Esrel-2008-THERP.pdf>
- Bunday, B. D. (1991). *Statistical methods in reliability theory and practice*. New York: Ellis Horwood.
- Creus, A. (1992). *Fiabilidad y seguridad. Su aplicación en procesos industriales*. Barcelona: Ed. Marcombo.
- Creus, A. (2001). *Jornadas de fiabilidad y error humano*. E.T.S.I.I. de Barcelona.
- CSNI. (1998). *Critical operator actions human reliability modeling and data issues. Final task report*. Nuclear Energy Agency, Committee on the Safety of Nuclear Installations.
- Dougherty, E. M., y Fragola, J. R. (1988). *Human reliability analysis: a systems engineering approach with nuclear power plant applications*. Wiley-Interscience.
- EATMP. (2002). *Technical review of human performance models and taxonomies of human error in ATM (HERA)*. European Organization for the Safety of Air Navigation. EATMP.
- Embrey, D. (1986). SHERPA. *A systematic human error reduction and prediction approach*. Paper presented at the International Meeting on Advances in Nuclear Power Systems, Knoxville, Tennessee.
- Embrey, D., Humphreys, P., rosa, E., y Rea, K. (1984). *An approach to assessing human error probabilities using structured expert judgement*. Department of Nuclear Energy, Brookhaven National Laboratory, Upton, New York, NY, (SLIM-MAUD, NUREG/CR-3515)
- Faig, J. (2007). *Fiabilidad humana: evaluación simplificada del error*. <http://www.insht.es>, extraído el 23/09/07. INSH. (NTP 619, 620 y 621)
- Forester, J. A., Cooper, s. E., Kolaczowski, A. M., Bley, D. C., Wreathall, J., y Lois, E. (2009). *An overview of the evolution of human reliability analysis in the context of probabilistic risk assessment* (Inf. Téc.). Sandia National Laboratories.
- Gertman, D., Blackmann, H., Haney, I., seidler, K., y Hahn, H. (1992). INTENT. A method for estimating human error probabilities for decision based errors. *Reliability Engineering & System Safety*, 35, 127-136.
- Hall, r. E., Fragola, J. R., y Wreathall, J. (1982). *Post-event human decision errors: Operator action trees/time reliability correlation*. Washington, DC.: USNRC. (NUREG/CR-3010)
- Hannaman, G., y Spurgin, A. (1984). SHARP. *Systematic Human Action Reliability Procedure* (n° NP-3583). Palo Alto, CA: Electric Power Research Institute.
- Hannaman, G., Spurgin, A., y Lukics, Y. (1985). *A model for assessing human cognitive reliability in PRA studies*. Proceedings IEEE Conference on Human Factors and Power Plants, Monterey, CA.
- Hoyland, A., y Rausand, M. (1994). *System reliability theory*. New Jersey: John Wiley and sons, Inc.
- Johnson, C. (2003). *Failure in safety-critical systems: A handbook of incident and accident reporting*. Glasgow, Scotland: Glasgow University Press.
- Kennedy, G., Siemieniuch, C., Sinclair, M., Kirwan, b., y Gibson, W. (2007). Proposal for a sustainable framework process for the generation, validation, and application of human reliability assessment within the engineering design lifecycle. *Reliability Engineering & System Safety*, 92, 755-770.
- Kim, M., y Seong, P. (2004, October). A quantitative model of system man interaction based on discrete function theory. *Journal of the Korean Nuclear Society*, 36, 430-449.
- Kirwan, B. (1994). *A guide to practical human reliability assessment*. London: Taylor & Francis.
- Maluf, D. A., Gawdiak, Y. O., y Bell, D.G. (2008). *On space exploration and human error a paper on reliability and safety*. Disponible en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.111.3314>
- Mosneron, F. (1989). *Méthode pratique de prise en compte du facteur humain dans les études de séquences accidentelles*. EPS FH 006 c, Electricité de France.
- Norman, D., y Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior* (Technical Report no 99). San Diego: Center for Human Information Processing.
- Norman, D. A. (1981). *Perspectives on cognitive science*. Ablex Publishing Corporation y Lawrence Erlbaum Associates.
- Pyy, P. (2000). *Human reliability analysis methods for probabilistic safety assessment*. Finland: Technical Research Centre.
- Quesada, J. (1999, Diciembre). *¿Por qué los errores humanos los cometen las personas habilidosas?* <http://www.andrew.cmu.edu/user/jquesada/humanerrorRPC>; extraído el 23/09/07.
- Rea, R., y Sandoval, s. (2000, mayo-junio). Análisis de riesgos en la industria química. *Boletín IIE-Méjico*, 24, 110-118. Disponible en <http://www.iie.org.mx/calidad2000/aplica.pdf>
- Spurgin, A. J. (1990). *Operator reliability experiments using power plant simulators*. Electric Power Research Institute, EPRI NP-6937, Vol. 1.
- Swain, A., y Guttman, H. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG/CR-1278. Albuquerque N.M. (US): Sandia Nuclear Laboratories.
- Williams, J. C. (1988, June). *A data-based method for assessing and reducing human error to improve operational performance*. Proceedings of the IEEE 4th Conference on Human Factors in Power Plants, Monterey, California.
- Williams, J. C. (1989, June). *Human reliability data: The state of the art and the possibilities*. Proceeding Reliability'89, Vol. 1. United Kingdom.

(Artículo recibido: 16-05-2010, revisión: 18-09-2011, aceptado: 08-10-2011)