

# Desarrollo ejecutivo temprano y Teoría de la Generalizabilidad: bebés típicos y prematuros

Ángel Blanco Villaseñor, Sylvia Sastre i Riba\* y Elena Escolano Pérez\*  
Universidad de Barcelona y \* Universidad de La Rioja

En las áreas científicas de las Ciencias Sociales, de la Salud y del Comportamiento utilizamos los diseños observacionales como unos de los que mejor permiten estudiar el comportamiento humano en situaciones naturales y cuasi-naturales. Una característica común a todos ellos es que sólo observamos y registramos una muestra del comportamiento de los individuos, pudiendo generar un amplio error muestral especialmente en edades tempranas. El error muestral surge como resultado de observar sólo una fracción de todas las ocasiones posibles. El muestreo en diferentes etapas (observadores, sesiones, tareas, edad, etc.), requiere la estimación de los componentes de variancia, que se combinan entre sí para producir estimaciones de coeficientes de generalizabilidad. La Teoría de la Generalizabilidad (TG) postula que en cualquier situación de medida existen múltiples (infinitas) fuentes de variación (facetas). El objetivo primordial es identificar los componentes de variancia que están aportando error a una estimación e implementar estrategias que reduzcan la influencia de estas fuentes de error sobre la medida. En concreto, se estudia el funcionamiento ejecutivo de bebés típicos y prematuros de 1;3 a. (15 meses) y 2;0 años (24 meses) mostrando la existencia de diferencias entre los grupos, las fuentes de variación y la generalizabilidad de los resultados.

*Executive function in early childhood and Generalizability Theory: Typical babies and preterm babies.*  
In the areas of Social, Health, and Behavioral Sciences, observational designs are used as one of the best ways to study human behavior in natural and quasi-natural settings. Observing and recording a sample of individual patterns of behavior is a feature common to all of them, but a broad sampling error can be implicit, especially at early ages. Sampling errors arise also as a result of the observation of only a fraction of each possible case, and this is part of the problem involved in the estimation of generalizations of data recorded by systematic observation. The sampling process usually takes place at different stages (observers, sessions, tasks, age, etc...) and requires estimating variance components that can be combined with each other to produce one or more estimates of board coefficients. The Generalizability Theory (GT) proposes that there are many (infinite) sources of variation (facets) in any measurement. Our main objective is to identify the variance components that contribute to estimation errors and implement strategies to reduce the influence of these error sources on the measurement. Executive functioning in typical and premature babies is studied at 15 and 24 months. The results show differences between groups, sources of variation, and generalizability.

Uno de los temas de estudio actuales dentro de la Psicología del Desarrollo es el de las funciones ejecutivas y su papel diferencial en el funcionamiento cognitivo del ser humano desde edades tempranas, con el objetivo de conocer cuál es su repercusión en el desarrollo cognitivo típico y atípico y sus correlatos neurológicos; considerando, además, que son esenciales para el adecuado rendimiento académico y social.

A pesar de que no existe una definición única de las funciones ejecutivas (Colom, Rubio, Shih y Santacreu, 2006; Zelazo, Müller, Frye y Markovitch, 2003), se considera como tales aquellas habi-

lidades centrales reguladoras que orquestan procesos básicos o de dominio específico con el fin de lograr un objetivo flexiblemente. Entre las funciones ejecutivas estudiadas, los autores distinguen entre aquellas denominadas *cool*, relacionadas con el control cognitivo, y las *hot*, relacionadas con el control emocional. Entre las primeras, relacionadas con el control y actividad cognitiva a las que haremos referencia, destacan las de: planificación, inhibición de un esquema no pertinente, y *shifting* o cambio, dada su relación con la toma de decisiones y resolución de problemas.

La planificación consiste en la capacidad para identificar y organizar los pasos necesarios para conseguir un objetivo, analizar las alternativas, atender y tomar decisiones a lo largo de la ejecución (Zelazo, 2004).

La inhibición consiste en la supresión de información no efectiva o interferencia distractora. Permite seleccionar las representaciones y acciones relevantes y la activación eficiente o *shifting* tras una señal de detención que permite modular o generar una nueva

estrategia resolutiva. Por lo tanto, una inhibición y activación ineficaces conducen a la perseveración o activación repetitiva e inapropiada que dificulta la flexibilidad de funcionamiento y resolución adecuadas, siendo necesario un cambio o *shifting* para alcanzar el objetivo deseado.

Estas funciones ejecutivas cambian a lo largo del desarrollo (Zelazo et al., 2003). Su inicio se sitúa en torno a los 5-7 meses de edad cuando se produce una maduración del córtex prefrontal dorsolateral y el área motora suplementaria (Bell y Fox, 1992; Diamond y Goldman-Rakic, 1989). Desde entonces, y siempre acompañadas por los cambios correspondientes de los sustratos neurológicos que las sostienen, se transforman con especial intensidad hasta los 2-5 años, alcanzando su estándar de funcionamiento a los 12 años. Aún así, continúan modificándose a lo largo de la adultez (Luciana, Conklin, Hooper y Yarger, 2005; Zelazo, Craick y Booth, 2004). Esto significa que a lo largo del desarrollo tienen una epigénesis diferencial que condiciona el tipo de eficacia resolutiva del individuo.

Durante los últimos años, la investigación está demostrando su manifestación diferencial en cursos de desarrollo atípicos o con riesgo: autismo, TDA-H, fenilcetonuria, prematuridad, hipotiroidismo, síndrome de Down, etc. (Blanco-Villaseñor, Sastre-Riba, Escolano-Pérez, Anguera y Pascual-Sufrate, 2008; Diamond, 2001; Geurts, Verté, Oosterlaan, Roeyers y Sergeant, 2004; Sastre, 2006; Sastre, Escolano y Merino, 2004; Shallice, Marzocchi, Cosser, Del Shavio, Meuter y Rumiati, 2002), así como en dificultades de aprendizaje escolar (Klein, 2003; Miranda, Meliá de Alba y Marco 2009) y control conductual posteriores (Bull, Espy y Wiebe, 2008; Elliot, 2003). El interés hacia su estudio está determinado por el papel que se ha demostrado que tienen en: 1) el control de la atención (atención selectiva, atención sostenida, inhibición); 2) el establecimiento y mantenimiento de un objetivo (iniciativa, planificación, organización, etc.) y; 3) la flexibilidad (memoria de trabajo, cambio atencional, monitorización).

En suma, el término “funciones ejecutivas” hace referencia a un macroconstructo que implica la coordinación de diferentes funciones necesarias para la supervisión y autorregulación cognitiva y emocional. Se modifican a lo largo de la edad y tipos de desarrollo (Carlson, Moses y Breton, 2002) y están relacionadas con el proceso de configuración neurobiológica del cerebro y su maduración. Para conocerlas, los investigadores utilizan tareas cerradas relacionadas con estímulos que reclaman el cambio organizado (*shift*) del foco de atención y tipo de respuesta, o bien tareas semiestructuradas (Escolano y Sastre, 2006) que permiten observar las modulaciones o cambios introducidos durante la organización de la actividad manteniendo el foco de atención (Davidson, Amso, Anderson y Diamond, 2006).

Por ello, investigar el funcionamiento ejecutivo en edades tempranas reclama un importante ajuste metodológico que responda a las características de los participantes (debidas a su pertenencia a distintos grupos de desarrollo, edad, etc.) y al número de observaciones repetidas que permitan la generalizabilidad de los resultados.

Suele ser difícil en Ciencias del Comportamiento encontrar una unidad física para medir la ejecución de una actividad (exceptuando estudios de percepción o de tiempos de reacción). Los sistemas de medida se basan, fundamentalmente, en la clasificación de los participantes en una investigación respecto a un comportamiento particular. A través del análisis factorial se tiene una garantía de que estos sistemas de medida son representativos del conjunto de

comportamientos que se pretende investigar. Se pone a punto una prueba, se estandariza y se eligen aquellas cuestiones que diferencian mejor la ejecución de las actividades por parte de los participantes. Ahora bien, una solución de tipo psicométrico conllevaría ciertos inconvenientes (Cardinet, 1987).

En la evaluación de contextos naturales no parece correcto recurrir a la clasificación para construir una escala de medida, ya que el cálculo tendría que ser muy exacto. Por ello es preferible utilizar medidas absolutas y no relativas, dado que su interpretación es más directa y más clara. Tampoco tiene mucho interés que la situación de evaluación esté estandarizada, puesto que el conocimiento del número exacto de respuestas que los individuos son capaces de resolver en un tiempo fijo y en unas observaciones particulares no suele aportar mucha información. Más bien interesaría estimar de forma aproximada su probabilidad de reconocer otras respuestas de forma diferente. No se trata, pues, de conocer el número de respuestas más difíciles de identificar que diferencian mejor, ya que en tal caso mediríamos los límites del aprendizaje y no las adquisiciones cognitivas reales.

Esta es una visión particular desde la TG que ha sido avalada especialmente por Cardinet (1987) y que ha supuesto importantes consideraciones metodológicas en el estudio de la observación directa del comportamiento a través de la metodología observacional (Blanco-Villaseñor, 1993). Estamos, por tanto, de acuerdo con Cardinet (1987) en que este tipo de evaluación no debe suponer una competición entre individuos, sino una forma de valorar la progresiva adquisición de elementos cognitivos del medio-ambiente.

Una solución psicométrica que integrara la noción de validez de contenido (necesaria en la evaluación ambiental) no nos permitiría conocer más allá de la variancia entre los individuos (Lakes y Hoyt, 2008). Se hace necesario, por tanto, acudir a la TG, que permite conocer la existencia de fuentes de variancia múltiples para unificar el tratamiento de todos esos casos diferentes. En la comparación con un diseño general de investigación podemos, además, valorar el principio de simetría de las facetas (o variables) y ver la integración de TG en el mismo.

Cronbach, Gleser, Nanda y Rajaratnam (1972) han desarrollado la TG presuponiendo que hay otras fuentes de variación además de las diferencias individuales e integrando cada una de estas fuentes de variación en una estructura global que permite aplicaciones particulares de la teoría estadística del muestreo. La TG reconoce explícitamente las múltiples fuentes de error de medida (individuos, contextos, tratamientos, observaciones, sesiones, ocasiones de medida, etc.). Podemos estimar cada una de estas fuentes de error así como las diferentes interacciones entre ellas. El error de medida no es más que el efecto de las fluctuaciones debidas a la elección aleatoria de individuos, contextos, tratamientos, observaciones, sesiones, ocasiones de medida..., es decir debidas al muestreo de niveles particulares en cada una de las facetas (variables) del universo de observaciones posibles. Optimizar dicha medida es adaptar nuestro diseño para reducir al máximo la variancia del muestreo debida a estas facetas.

La TG es una teoría de los errores multifaceta de una medición conductual (Cronbach et al, 1972). El objetivo de la teoría es desglosar, en cualquier tipo de medición, la variabilidad real de la variabilidad del error. Para que se cumpla, la teoría necesita de los componentes del análisis de la variancia; las variaciones de las facetas (variables) tales como individuos, contextos, tratamientos, escalas de medida, observaciones, sesiones, ocasiones de registro, etc.

El eje central de la TG, por tanto, se encuentra en los componentes de variancia dado que su magnitud nos aporta información sobre las fuentes de error que están afectando una medición conductual.

La filosofía básica que subyace a la TG es que “*un investigador se pregunte acerca de la precisión o fiabilidad de una medición dado que desea generalizar de observaciones reales a cualquier tipo de observaciones a las que éstas pertenezcan*” (Cronbach, Rajaratnam y Gleser, 1963, p. 144).

La TG nos ofrece un marco más satisfactorio para buscar las estimaciones de fiabilidad y de márgenes de error, ya que es suficientemente globalizadora como para adaptarse a las condiciones particulares de cada objeto de medida (Blanco-Villaseñor, Castellano y Hernández-Mendo, 2000; Blanco-Villaseñor, 1992). Este carácter globalizador, sin embargo, no ha sido puesto de manifiesto por los autores de esta teoría, quienes la han formulado casi exclusivamente en términos apropiados a contextos psicométricos, en su obra de base, *The Dependability of Behavioral Measurement: Theory of Generalizability for Scores and Profiles* (Cronbach et al., 1972). Las extensiones sucesivas de este enfoque han planteado nuevas categorías de problemas —sobre todo con la inclusión de los diseños observacionales—, que han sido propuestos y resueltos por Berk (1979), Cardinet y Tourneur (1985), Frick y Semmel (1978), Mitchell (1979), Plewis (1988) y Smith y Teeter (1982). A su vez, y como consecuencia de estas extensiones, se ha multiplicado el espectro de aplicaciones diversas en el ámbito de la Psicología, como lo demuestran los trabajos recientes en diferentes áreas de conocimiento (Psicología del Desarrollo, Psicología de la Educación, Evaluación Psicológica, Psicología del Deporte, etc.). Todos ellos cuentan con literatura metodológica reciente que solventa problemas analíticos de diversa índole (Algina y Keselman, 1997; Cardinet, 1997) y de forma específica en su aplicación a datos categóricos en los trabajos de Li y Lautenschlager (1997) y de Marcoulides (1989a, 1989b).

La demarcación que vamos a llevar a cabo introduce una distinción entre las fases del análisis de la variancia y las que se fundamentan en los conceptos de TG (Máiquez, Blanco-Villaseñor, Rodrigo y Vermaes, 2000). El modelo del análisis de la variancia tiene en cuenta las observaciones en las que se supone la existencia de fuentes de variancia. Permite precisar la importancia de cada una de estas fuentes de variación, atribuyéndoles una porción de la variancia total. En este modelo, nada evoca la distinción entre puntuación verdadera y del error. Todas las fuentes de variancia son necesarias en una descripción correcta y completa de la realidad observada.

Por lo tanto, este trabajo comparativo y diferencial, tiene un doble objetivo de carácter metodológico y del desarrollo cognitivo temprano, centrándose en conocer cuál es la variabilidad del funcionamiento ejecutivo en distintos grupos de desarrollo (típico y con factores de riesgo al nacer: prematuridad), controlando los factores de variabilidad y la generalizabilidad de los resultados en función del tamaño de la muestra.

## Método

### Participantes

Han sido estudiados 18 bebés, 10 de ellos con curso de desarrollo típico y 8 nacidos prematuramente y que cumplen los criterios de: a) haber nacido entre la 30/35 semana de gestación (inclusivos); b) haber nacido con el peso adecuado a la edad

gestacional; c) no provenir de parto múltiple; d) no padecer patología neurológica asociada y e) no presentar malformaciones congénitas asociadas. La edad de los participantes es de 1;3 años en la primera observación y de 2;0 años en la segunda. En el caso de los bebés prematuros se ha corregido la edad cronológica en función de su edad gestacional.

La extracción de la muestra de bebés típicos se ha realizado entre los asistentes a una Escuela Infantil urbana y los bebés nacidos con prematuridad, mediante la colaboración del equipo de neuropediatría del Hospital San Pedro de Logroño. En todos los casos los padres han firmado el consentimiento informado de la participación voluntaria de sus hijos en la investigación.

### Instrumentos

El material de estímulo configura una tarea abierta que permite la realización flexible de diferentes acciones organizadas o moduladas.

Consiste en: a) tres tubos de plástico de 25 cm de longitud y distinto diámetro: 3 cm, 2 cm y 1,5 cm, siendo el segundo de ellos opaco y el resto transparentes; b) dos embudos con canilla de distinto grosor: 1,5 cm y 2,5 cm; c) tres cubiletes opacos iguales salvo en su base: en uno era sólida y en los otros dos estaba agujereada con distinto grosor, 1 cm y 2 cm; d) seis corchos de distinto grosor: cuatro iguales entre sí, de 1,5 cm de diámetro en su base inferior, y dos iguales entre sí y más gruesos que los anteriores, de 2,5 cm de diámetro en su base inferior y, e) un contenedor de plástico transparente, de dimensiones, 50 × 30 × 15 cm, con piedrecitas de acuario de colores.

Para apresar el funcionamiento ejecutivo de los bebés se construyó y validó el sistema de categorías que se representa en la Tabla 1. Este sistema mixto de análisis (categorías y formatos de campo) se compone de nueve unidades de conducta relativas a las funciones ejecutivas de inhibición, flexibilidad y *shifting*, tanto en su sentido positivo (buen funcionamiento ejecutivo) como negativo (funcionamiento ejecutivo deficiente o inapropiado).

Unidad de conducta	Descripción
SH2 <sup>a</sup>	Nuevo esquema para conseguir el mismo objetivo en una actividad continuada.
SH1 <sup>a</sup>	Cambio de objetivo en una actividad continuada.
AUC <sup>a</sup>	Ensayo nueva resolución tras detección de error, manteniendo el mismo objetivo.
COM <sup>a</sup>	Aplica diferentes acciones para obtener los objetivos.
NES	En el curso de la actividad, esquema no realizado antes por el bebé ni por el adulto (ni modularización, ni sustitución ni elaboración).
MIS	Repetición de una acción realizada, no inmediata.
RE <sup>a</sup>	Producto estable por combinación de distintas acciones.
ACT	Acción o esquema repetitivo (sin modularización).
DET	Paro de la actividad.

<sup>a</sup> = Unidades de conducta relacionadas con el buen funcionamiento ejecutivo y flexibilidad

El programa informático GT —*Generalizability Theory*— (Ysewijn, 1996) fue utilizado para el análisis de datos.

### Procedimiento

Se ha registrado la actividad de los bebés a los 1;3 años y a los 2;0 años con el material de estímulo descrito y en un lugar familiar para ellos. La duración media de la actividad es de 15 minutos durante los cuales los bebés están acompañados por un adulto conocido, que no participa en la actividad infantil.

### Análisis de datos

Con el fin de inferir la participación de las funciones ejecutivas en la actividad de los bebés fue preciso: 1) transcribir fiablemente la actividad espontánea de los bebés y codificar las unidades de conducta relacionadas con el funcionamiento ejecutivo de acuerdo con el sistema de categorías creado previamente; 2) aplicar el análisis de variabilidad con un diseño multifaceta: *Tipos* (bebés típicos/bebés prematuros) x *Participante* x *Edad* x *Categoría* (frecuencia unidades de conducta) aplicando dos procedimientos de cálculo: GLM (General Linear Model) Type I y Type III del paquete de programas estadísticos SAS; 3) calcular el coeficiente de generalizabilidad de la muestra estudiada y de tres proyecciones de ésta hasta  $n = 20$  participantes para cada curso de desarrollo, mediante el programa GT.

El Plan de Medida de nuestra estructura de diseño multifaceta es aquél en el que las facetas *Tipos*, *Edad* y *Categoría* constituirán las facetas de Diferenciación, y la faceta *Participante* será la de Generalización, de forma que nos permita obtener si el número de participantes es suficiente para poder generalizar los resultados obtenidos a las poblaciones de referencia: bebés con curso de desarrollo típico y bebés prematuros.

### Resultados

La tabla 2 presenta los resultados obtenidos a través del diseño multifaceta *Tipos x Participante x Edad x Categoría*. La primera columna representa las Fuentes de Variación (FV) y sus interacciones. Las columnas segunda y tercera representan, respectivamente, la Suma de Cuadrados (SC) y sus Componentes de Varian-

cia (CV), es decir, la variabilidad estimada en el caso de un número infinito de niveles en las variables objeto de estudio. Las columnas del procedimiento GLM representan, respectivamente, si hay o no diferencias significativas en el caso de que cada una de las facetas e interacciones hubieran sido seleccionadas al azar (Type I) y en el caso de que sólo consideráramos el número real de niveles de las facetas estudiadas (Type III). En las últimas dos columnas se presentan los porcentajes de variabilidad correspondientes al modelo de efectos aleatorios (Type I) y a un modelo en el que la faceta *Tipos* (bebés típicos y bebés prematuros) ha sido considerada como faceta fija.

Dado el reducido tamaño de la muestra habitual en este tipo de investigaciones del desarrollo temprano, podemos observar que existen resultados dispares en el modelo de efectos aleatorios y en el modelo de efectos fijos. Concretamente, en la faceta *Tipos* y en la interacción de las facetas *Tipos x Categoría* nos encontramos que en ambos casos existen diferencias en el modelo de efectos aleatorios y no en el de efectos fijos. En el resto de facetas individuales o en interacción, los resultados son similares.

Puesto que uno de los objetivos de este estudio es la incidencia de la edad y del número de participantes en la misma, con respecto a presentar un curso de desarrollo típico o prematuro, y dado que existen diferencias significativas en ambos modelos, se confirman nuestras expectativas relacionados con los cambios en el funcionamiento ejecutivo a lo largo de las edades estudiadas.

Puede observarse también que más del 74% de variabilidad del modelo estudiado, en ambos casos, es explicado por las siguientes interacciones: *Categoría*, *Tipos x Categoría* y *Participante x Categoría*, donde el sistema de categorías está presente en todas las situaciones. Ello, a su vez, implica afirmar la validez del instrumento de observación construido a tal fin, dado la existencia de diferencias significativas en los tres casos en el modelo aleatorio.

Como se desprende de la tabla 2, el valor del residual obtenido (interacción de todas las facetas) es 0, indica que las diferencias entre los grupos de desarrollo estudiados son debidas a la incidencia de las variables o fuentes de variación contempladas (*Tipos* o cursos de desarrollo, *Participante*, *Edad* y *Categoría*) y sus interacciones de primer y segundo orden.

Entre estas facetas, las más explicativas según el peso específico (%) son: *Categoría* (24%), *Tipos x Categoría* (21%), y *Participante x Categoría* (29%), confirmando que los componentes eje-

FV	SC	CV Aleatorios	GLM		%	%
			Type I Aleatorios	Type III Fijos		
Tipos (T)	4241,8	13,7	<0,0001	–	15	8
Participante (P)	2333,3	19,9	<0,0001	<0,0001	2	2
T x P	0	0	–	–	0	0
Edad (E)	316,7	19,9	0,0038	0,0078	1	1
T x E	37,4	-1,5	0,3128	–	0	0
P x E	170,5	1,5	0,4606	0,4606	0	0
T x P x E	0	0	–	–	0	0
Categoría (C)	16496	35,9	<0,0001	<0,0001	24	37
T x C	4709,8	82,3	<0,0001	–	21	11
P x C	12643	-74,9	<0,0001	<0,0001	29	31
T x P x C	0	0	–	–	0	0
E x C	333,9	-0,5	0,4296	0,4225	0	1
T x E x C	199	-3,1	0,7899	–	2	1
P x E x C	1250,4	-20,6	0,8516	0,8516	6	7
T x P x E x C	0	0	–	–	0	0

cutivos de la actividad apresados mediante las categorías previamente construidas (véase tabla 1) varían en función del curso de desarrollo del bebé, típico o prematuro, siendo mucho más relevante que la edad o punto de medida (1%), aun cuando se trata de un estudio longitudinal donde la medición se ha tomado en dos momentos de medida y existen diferencias significativas en la interacción *Tipos*  $\times$  *Edad*.

Dado que otro de los objetivos del estudio es conocer el grado de precisión de la generalización a través de los componentes de variancia obtenidos o coeficiente de generalizabilidad, la tabla 3 recoge los resultados obtenidos al respecto.

	$n_{\text{típicos}}=10$ $n_{\text{prematuros}}=8$	$n=12 \times \text{Tipo}$	$n=15 \times \text{Tipo}$	$n=20 \times \text{Tipo}$
Modelo Aleatorio	0,94	0,95	0,96	0,97
Modelo Mixto	0,90	0,94	0,94	-

La primera columna de la tabla 3 expresa el coeficiente de generalizabilidad o de estimación de precisión de los resultados obtenidos en el estudio con  $n=10$  bebés típicos y  $n=8$  bebés prematuros: en el modelo totalmente aleatorio 0,94 y en el modelo mixto en el que la faceta *Tipos* ha sido considerada como fija 0,90.

La segunda columna representa los resultados del Plan de Optimización aumentando el tamaño de la muestra a  $n=12$  participantes por cada *Tipo* (12 bebés típicos y 12 prematuros), la tercera y cuarta columnas representa los obtenidos con  $n=15$  y  $n=20$ , respectivamente.

Los resultados de precisión de generalización son excelentes, tanto en el modelo aleatorio como en el mixto, pero la obtención de participantes de la muestra de bebés prematuros es tan costosa y compleja que el aumento de precisión ofrece baja rentabilidad coste-beneficio. El alto coste que implicaría aumentar a 15 participantes de cada tipo apenas repercute en un incremento del valor del coeficiente de generalizabilidad (de 0,94 a 0,96 en el modelo aleatorio y de 0,90 a 0,94 en el modelo mixto). Incluso doblando su tamaño (20 bebés de cada tipo), no justifica su aumento.

#### Discusión y conclusiones

A partir de los resultados obtenidos, desde una perspectiva metodológica, la ventaja del tipo de análisis realizado es la capacidad

para diseñar estudios de Optimización más eficientemente en base a la información aportada por los Planes de Estimación y Medida. Si tenemos en cuenta la relación nivel de precisión-costes, es obvio que podremos diseñar un estudio de generalizabilidad óptimo.

En definitiva, la teoría de la generalizabilidad organiza la información oponiendo las aplicaciones exploratorias a las confirmatorias, o bien las aplicaciones a priori a las a *posteriori*.

Una primera dirección del análisis de generalizabilidad se refiere a competencias o dominios todavía no lo suficientemente conocidos. A través del estudio de las fuentes de variancia ligadas a las diferentes facetas, es posible estructurar el conjunto de datos que han sido registrados. Es decir, cuando se han determinado las direcciones principales de la variación, las que son más generalizables, ya podemos saber cuáles son las facetas que deben tenerse en cuenta posteriormente para muestrear esa competencia o dominio. Esta sería una utilización heurística de los coeficientes de generalizabilidad que permitiría un análisis exploratorio de los datos.

De esta manera, el coeficiente de generalizabilidad obtenido a partir de un determinado número de participantes, especialmente en muestras costosas de obtener (por ejemplo, niños nacidos con factores de riesgo o distintas patologías establecidas), es útil para el incremento del conocimiento científico, en general, y para el investigador, en particular.

Desde el punto de vista del desarrollo cognitivo, los resultados obtenidos concuerdan con los de otras investigaciones (Crawford, Espy, Isquith y Gioia, 2005; Zelazo, 2004, 2006) en el sentido de mostrar la existencia de una incipiente disfunción ejecutiva en los bebés prematuros estudiados, ya en las tempranas edades comprendidas entre 1;6 años y 2;0 años y entre distintos cursos de desarrollo, así como la importancia de realizar una evaluación multivariada de las mencionadas funciones ejecutivas para apresar mejor los componentes afectados en los niños con posibles disfunciones del desarrollo y, en este caso, los bebés nacidos con prematuridad.

Por otra parte, esta evidencia temprana tiene una función preventiva y aconseja el establecimiento de programas de intervención que incluyan las funciones ejecutivas para aminorar posibles dificultades de aprendizaje y control conductual que se pueden poner de manifiesto en edades escolares.

Un estudio microgenético de la actividad realizada por cada niño y la participación en ella de las funciones ejecutivas con distinto grado de efectividad, podría optimizar la comprensión de las fuentes de variabilidad en su funcionamiento y, en consecuencia, mejorar la dimensión preventiva a la que se hacía referencia dada su repercusión en la adaptación personal, familiar, escolar y social.

#### Referencias

- Algina, J., y Keselman, H.J. (1997). Detecting repeated measures effects with univariate and multivariate statistics. *Psychological Methods*, 2, 208-218.
- Bell, M.A., y Fox, N.A. (1992). The relations between frontal brain electrical activity and cognitive development during infancy. *Child Development*, 63, 1142-1163.
- Berk, R.A. (1979). Generalizability of behavioral observations: A clarification of interobserver agreement and interobserver reliability. *American Journal of Mental Deficiency*, 83(5), 460-472.
- Blanco-Villaseñor, A. (1992). Aplicaciones de la Teoría de la Generalizabilidad en la selección de diseños evaluativos. *Bordón*, 43(4), 431-459.
- Blanco-Villaseñor, A. (1993). Fiabilidad, precisión, validez y generalización de los diseños observacionales. En M.T. Anguera (Ed.): *Metodología observacional en la investigación psicológica* (Vol. 2: Fundamentación, pp. 151-261). Barcelona: PPU.
- Blanco-Villaseñor, A., Castellano, J., y Hernández-Mendo, A. (2000). Generalizabilidad de las observaciones de la acción del juego en el fútbol. *Psicothema*, 12(Sup. 2), 81-86.
- Blanco-Villaseñor, A., Sastre-Riba, S., Escolano-Pérez, E., Anguera, M.T., y Pascual-Sufrate, M.T. (2008, July). Generalizability and Observational Missing Data in Growth and Developmental Research with Prema-

- ture Infants and Low Birth Weight Babies. In A. Blanco-Villaseñor (Chair), *Advances in Applied Research on Observational Methodology*. Symposium presented at the III European Congress of Methodology, 8-12 July. Oviedo, Spain.
- Bull, R., Espy, K.A., y Wiebe, S.A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33(3), 205-228.
- Cardinet, J. (1987). *La construction de tests d'apprentissage selon la théorie de la généralisabilité* (Recherches; 87.107). Neuchâtel: Institut romand de recherches et de documentation pédagogiques.
- Cardinet, J. (1997). From classical test theory to generalizability theory - the contribution of ANOVA [french]. *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée*, 47(3), 197-204.
- Cardinet, J., y Tourneur, Y. (1985). *Assurer la mesure*. Berne: Peter Lang.
- Carlson, S.M., Moses, L.J., y Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development. Special issue: Executive function and its development*, 11, 73-92.
- Colom, R., Rubio, V.J., Shih, P.C., y Santacreu, J. (2006). Fluid intelligence and executive functioning. *Psicothema*, 18(4), 816-821.
- Crawford, J.S., Espy, K.A., Isquith, P.K., y Gioia, G.G. (2005). Assessment of Executive Functions in Preschool Ages Children. *Mental Retardation and Developmental Disabilities Research Reviews*, 11, 100-104.
- Cronbach, L.J., Gleser, G.C., Nanda, H., y Rajaratnam, N. (1972). *The dependability of behavioral measurements: theory of generalizability for scores and profiles*. New York: John Wiley and Sons.
- Cronbach, L.J., Rajaratnam, N., y Gleser, G.C. (1963). Theory of generalizability: a liberalization of reliability theory. *British Journal of Mathematical and Statistical Psychology*, 16, 137-163.
- Davidson, M.C., Amso, D., Anderson, L.C., y Diamond, A. (2006). Development of cognitive control and executive function from 4 to 13 years: Evidence from manipulation of memory, inhibition and task switching. *Neuropsychologia*, 44(11), 2037-2078.
- Diamond, A. (2001). A model system for studying the role of dopamine in prefrontal cortex during early development in humans: early and continuously treated phenylketonuria. En C. Nelson y M. Luciana (Eds.): *Handbook of Developmental Cognitive Neuroscience* (pp. 433-472). Cambridge, MA: MIT Press.
- Diamond, A., y Goldman-Rakic, P.S. (1989). Comparison of human infants and rhesus-monkeys on Piaget's A-non-B task: evidence for dependence on dorsolateral frontal cortex. *Experimental Brain Research*, 74(1), 24-40.
- Elliott, R. (2003). Executive functions and their disorders. *British Medical Bulletin*, 65(1), 49-59.
- Escolano, E., y Sastre, S. (2006). Actividad lógica de los bebés: un estudio diferencial. *Psicothema*, 18(3), 537-543.
- Frick, T., y Semmel, M.I. (1978). Observer agreement and reliabilities of classroom observational measures. *Review of Educational Research*, 48(1), 157-184.
- Geurts, H.M., Verté, S., Oosterlaan, J., Roeyers, H., y Sergeant, J.A. (2004). How specific are executive functioning deficits in Attention Deficit Hyperactivity Disorder and Autism? *Journal of Child Psychology and Psychiatry*, 45, 836-854.
- Klein, C. (2003). Assessing the development of executive functions. *International Society for the Study of Behavioural Development Newsletter*, 43(1), 8-11.
- Lakes, K.D., y Hoyt, W.T. (2008). What Sources Contribute to Variance in Observer Ratings? Using Generalizability Theory to Assess Construct Validity of Psychological Measures. *Infant and Child Development*, 17, 269-284.
- Li, M.N.F., y Lautenschlager, G. (1997). Generalizability theory applied to categorical data. *Educational and Psychological Measurement*, 57(5), 813-822.
- Luciana, M., Conklin, H.M., Hooper C.J., y Yarger, R.S. (2005). The development of non-verbal working memory and executive control processes in adolescents. *Child Development*, 76(3), 697-712.
- Máiquez, M.L., Blanco-Villaseñor, A., Rodrigo, M.J., y Vermaes, I. (2000). La evaluación de la eficiencia en la intervención familiar: generalizabilidad y optimización del programa experiencial para padres. *Psicothema*, 12(4), 533-542.
- Marcoulides, G.A. (1989a). The application of generalizability analysis to observational studies. *Quality and Quantity: The International Journal of Methodology*, 23(2), 115-127.
- Marcoulides, G.A. (1989b). The estimation of variance components in generalizability studies: A resampling approach. *Psychological Reports*, 65(3), 883-889.
- Miranda, A., Meliá de Alba, A., y Marco, R. (2009). Habilidades matemáticas y funcionamiento ejecutivo de niños con trastorno por déficit. *Psicothema*, 21(1), 63-69.
- Mitchell, S.K. (1979). Interobserver Agreement, Reliability, and Generalizability of Data Collected in Observational Studies. *Psychological Bulletin*, 86(2), 376-390.
- Plewis, I. (1988). Estimating generalizability in systematic observation studies. *British Journal of Mathematical and Statistical Psychology*, 41, 53-62.
- Sastre, S. (2006). Condiciones tempranas del desarrollo y el aprendizaje: el papel de las funciones ejecutivas. *Revista de Neurología*, 42(Suppl. 2), s143-s151.
- Sastre, S., Escolano, E., y Merino, N. (2004). Observación sistemática de la cognición temprana: funciones ejecutivas, lógica e interacción. *Metodología de las Ciencias del Comportamiento, Volumen especial*, 557-566.
- Shallice, T., Marzocchi, G.M., Coser, S., Del Shavio, M., Meuter, R.F., y Rumiat, R.I. (2002). Executive function profile of children with attention deficit hyperactivity disorder. *Developmental Neuropsychology*, 21(1), 43-71.
- Smith, P.L., y Teeter, P.A. (March, 1982). *The Use of Generalizability Theory with Behavioral Observation*. Paper presented at the Annual Meeting of the American Educational Research Association. New York.
- Ysewijn, P. (1996). GT. *Software for Generalizability Studies, version 2.0.E*.
- Zelazo, P.D. (2004). The development of conscious control in childhood. *Trends on Cognitive Science*, 8(1), 12-17.
- Zelazo, P.D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1, 297-301.
- Zelazo, P.D., Craick, F.I.M., y Booth, L. (2004). Executive function across the life-span. *Acta Psychologica*, 115(2-3), 167-183.
- Zelazo, P.D., Müller, U., Frye, D., y Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, 68(3), 1-137.