

Análisis de datos observacionales mediante la Teoría de la Generalizabilidad y la utilización del Modelo Lineal General y Mixto: Un estudio empírico del desarrollo y aprendizaje infantil

Angel Blanco-Villaseñor^{1*} y Elena Escolano-Pérez²

¹ Universidad de Barcelona, Barcelona (España).

² Universidad de Zaragoza, Zaragoza (España).

Resumen: Una adecuada evaluación de las competencias infantiles tempranas es esencial para potenciar un desarrollo óptimo, pues los primeros años de vida son la base de todo el desarrollo y aprendizaje posterior. Sin embargo, todavía existen ciertas limitaciones y deficiencias en el ámbito de la medición del desarrollo y aprendizaje infantil. Con el objetivo último de contribuir a la mejora de esta situación, este trabajo presenta las posibilidades y ventajas que ofrecen nuevas técnicas de análisis de datos, tanto para controlar la calidad de los datos infantiles registrados a través de observación sistemática como para analizar su variabilidad. Se ha observado en tres edades diferentes (18, 21 y 24 meses) la actividad lógica y ejecutiva de 48 niños usando un diseño observacional nomotético, de seguimiento y multidimensional.

Dadas las particularidades de los datos del estudio que presentamos, desde el punto de vista metodológico y su análisis, realizamos análisis pormenorizados a través de la Teoría de la Generalizabilidad en tres vertientes posibles en un estudio observacional: Análisis de la fiabilidad intra e inter-observadores, Análisis de la validez del instrumento de observación y Estimación muestral de las facetas estudiadas (en concreto, la de participantes). De esta forma, se pretende optimizar el número de facetas y niveles necesarios para llevar a cabo un estudio de tales características.

Además, se utilizan otras técnicas analíticas para conocer la variabilidad del desarrollo y aprendizaje infantil, como son el Modelo Lineal General y el Modelo MIXED.

Los resultados indican cómo el uso de la Teoría de la Generalizabilidad permite controlar la calidad de los datos observacionales en una estructura única que integra la fiabilidad, validez y generalizabilidad.

Palabras clave: Observación sistemática; Modelo Lineal General; Teoría de la Generalizabilidad; desarrollo; aprendizaje; infancia.

Title: Observational data analysis using generalizability theory and general and mixed linear models: an empirical study of infant learning and development.

Abstract: Accurate evaluation of early childhood competencies is essential for favoring optimal development, as the first years of life form the foundations for later learning and development. Nonetheless, there are still certain limitations and deficiencies related to how infant learning and development are measured. With the aim of helping to overcome some of the difficulties, in this article we describe the potential and advantages of new data analysis techniques for checking the quality of data collected by the systematic observation of infants and assessing variability. Logical and executive activity of 48 children was observed in three ages (18, 21 and 24 months) using a nomothetic, follow-up and multidimensional observational design.

Given the nature of the data analyzed, we provide a detailed methodological and analytical overview of generalizability theory from three perspectives linked to observational methodology: intra- and inter-observer reliability, instrument validity, and sample size estimation, with a particular focus on the participant facet. The aim was to identify the optimal number of facets and levels needed to perform a systematic observational study of very young children.

We also discuss the use of other techniques such as general and mixed linear models to analyze variability of learning and development.

Results show how the use of Generalizability Theory allows controlling the quality of observational data in a global structure integrating reliability, validity and generalizability.

Key words: Systematic observation; General Linear Model; Generalizability Theory; development; learning; childhood.

Introducción

El desarrollo humano es un fenómeno amplio y complejo (Guralnick, 2015), dado que constituye un proceso de construcción y cambio continuo que surge de la interacción heterogénea y dinámica de numerosos elementos: genes; actividad neural; conducta pre y postnatal; contexto físico, social y cultural; etc. (Massand y Karmiloff-Smith, 2015). Además, estos cambios, que tienen lugar ya desde la concepción y durante todo el ciclo vital, afectan a todas las dimensiones del ser humano: física, social, cognitiva, lingüística, emocional y personal; si bien no suceden simultáneamente, con la misma frecuencia e intensidad en todos los momentos o etapas vitales, ni tampoco en todas las áreas del individuo. El periodo que abarca de los 0 a los 3 años es una etapa especialmente crítica en el desarrollo humano dado los numerosos y complejos cambios que, de manera interaccionada, acontecen en las distintas dimensiones del individuo, implicando múltiples ganancias sobre las que, posteriormente, irán construyéndose otras más complejas. Así, lo que sucede en esta

etapa, constituye las bases del desarrollo y aprendizaje posterior (Scharf, Scharf y Stroustrup, 2016).

Dentro de los numerosos y complejos cambios que acontecen en dicha etapa, destacan especialmente los referidos al desarrollo cognitivo. El desarrollo cognitivo es un componente clave del desarrollo infantil temprano (Nelson y Luciana, 2008), pues implica la construcción de un conjunto de habilidades muy diversas, pero interdependientes, que tienen que ver, esencialmente, con los procesos ligados a la adquisición, organización, retención y uso de la información y conocimiento, permitiendo la adaptación del individuo al contexto continuamente cambiante (Goswami, 2010). Este contexto afecta asimismo a la naturaleza del cambio, dado que es elemento participante activo en él, al igual que lo es el propio individuo. La construcción de estas habilidades cognitivas, o lo que es lo mismo, el desarrollo cognitivo, se encuentra estrechamente relacionado con el desarrollo estructural del córtex, de modo que un incremento en la complejidad de la estructura cerebral da lugar a un incremento en las capacidades cognitivas. Pero a su vez, esta complejidad y organización neural se ve modificada por su propia funcionalidad. La progresiva especialización de las estructuras neurales está dirigida por las experiencias ambientales, activamente elegidas por el niño y a las que se enfrenta también de modo activo y participativo. Así, la puesta en marcha de habilidades cognitivas bási-

*** Correspondence address [Dirección para correspondencia]:**

Angel Blanco-Villaseñor. Facultad de Psicología. Universidad de Barcelona. Campus Mundet. Pº Vall d'Hebrón, 171. 08035 Barcelona (España). E-mail: ablanco@ub.edu

cas posibilita el establecimiento de nuevas conexiones neuronales que permiten el logro de capacidades y aprendizajes cada vez más complejos y elaborados (Karmiloff-Smith, Casey, Massand, Tomalski y Thomas, 2014).

Entre las diferentes habilidades cognitivas que emergen en la infancia temprana, este trabajo se centra en la lógica y las funciones ejecutivas.

La lógica implica la capacidad para captar, elaborar, estructurar e interiorizar la información, y tiene sus orígenes en la acción organizada y significativa que el bebé realiza sobre el medio (Langer, 1986, 1990). En algunas de estas acciones, el bebé centra su atención en la exploración y experimentación de la realidad física, obteniendo información de las propiedades directas de los objetos, es decir, elaborando el conocimiento físico. En otras, focaliza su atención en las propias acciones realizadas sobre ellos; en las relaciones que, a través de su acción, se establecen entre los objetos, obteniendo el conocimiento lógico-matemático (Langer, 1986, 1990). Aunque la lógica basada en la acción existe desde los primeros días de vida, el segundo año es especialmente relevante por los logros alcanzados. Los esquemas de acción básicos como tirar o chupar son sustituidos por acciones diferenciadas, es decir, por acciones que se adaptan a las características específicas de los objetos. Además, estas acciones se coordinan, flexibilizan, combinan y redefinen, se prueban resultados semejantes para obtener nueva información. El niño comienza a ser capaz de juntar objetos distintos, luego elementos iguales, y tras una organización y complejidad progresiva de la acción, llega a establecer relaciones entre objetos que pertenecen a distintos conjuntos de elementos. Entre estos tipos de relaciones, destaca la correspondencia uno a uno. Implica la puesta en relación, de modo secuencial, de todos y cada uno de los objetos de un conjunto con un elemento diferente del otro conjunto, obteniendo como resultado una relación de equivalencia entre los conjuntos. Constituye un componente esencial del desarrollo y aprendizaje de las habilidades matemáticas, habilidades imprescindibles para desenvolvemos con éxito en nuestra sociedad (Izard, Streri y Spelke, 2014). La capacidad para llegar a establecer correspondencias uno a uno se ve afectada por las características de los propios elementos, por el número de elementos que componen cada conjunto, además de por el número de conjuntos entre los que se establece la relación. Así, los elementos que guardan entre sí relaciones espaciales de inserción (los elementos de un conjunto pueden ser introducidos en los elemento del otro conjunto) facilitan el establecimiento de las correspondencias uno a uno (Langer, 1986, 1990; Sinclair, Stambak, Lézine, Rayna y Verba, 1984). En cuanto al número de conjuntos y elementos de cada uno de ellos, a los 18 meses casi la totalidad de niños (91,6%) establecen correspondencias entre dos conjuntos compuestos cada uno de ellos por dos elementos. Sin embargo, no establecen correspondencias entre tres o más conjuntos de dos elementos ni tampoco entre dos conjuntos de tres o cuatro elementos. A los 21 meses todos los niños realizan ya mínimas correspondencias entre dos conjuntos compuestos cada uno por dos objetos. En algunos niños (16,6%) comienzan a aparecer las correspondencias entre dos conjuntos compuestos por tres objetos, así como las correspondencias entre tres conjuntos (8,3%) y cuatro conjuntos (8,3%) de dos elementos. A los 2 años se producen avances importantes. La mayoría de niños (58,3%) ya generan correspondencias uno a uno entre dos con-

juntos de tres elementos. Por primera vez, aparecen las correspondencias entre dos conjuntos de cuatro elementos, siendo realmente alto el porcentaje de niños que las generan (66%). Además, continúa incrementándose el porcentaje de niños que construye correspondencias entre tres conjuntos de dos elementos (25%) y entre cuatro conjuntos de dos elementos (33,3%) (Langer, 1986).

A pesar de la relevancia de la lógica y en concreto de la correspondencia uno a uno, su desarrollo durante los tres primeros años de vida apenas ha sido estudiado. La mayor parte de estudios sobre la correspondencia uno a uno se centran en los años preescolares y escolares, es decir, a partir de los 3 años y dentro del ámbito matemático de la educación formal (Muldoon, Lewis y Towse, 2005), olvidando que el conocimiento que ha construido el niño antes de su entrada en la escuela constituye la base sobre la que se fundamenta el desarrollo y aprendizaje posterior.

En cuanto a las funciones ejecutivas (entendidas como aquellos procesos que permiten el control y autorregulación de la conducta para alcanzar un objetivo en situaciones novedosas y complejas -Barkley, 2012; Guare, 2014-), están siendo objeto de gran interés en la última década, especialmente en lo referente a su desarrollo temprano (Carlson, Zelazo y Faja, 2013), dada su relevancia e implicación no solo en el propio desarrollo cognitivo sino también en el desarrollo social, personal y emocional infantil; destacándose como aspecto esencial para el ajuste y éxito escolar e incluso para la salud (Diamond, 2013; García, Rodríguez, González-Castro, Álvarez-García y González-Piendi, 2016; Guare, 2014; Iglesias-Sarmiento, Carriedo y Rodríguez, 2015). Las funciones ejecutivas (FE) constituyen las bases del aprendizaje y adaptación: permiten sostener la atención, mantener la información y los objetivos en mente, controlarse para no responder automáticamente, resistir las distracciones, considerar las consecuencias de las acciones, reflexionar sobre las experiencias pasadas y planificar las futuras. Su relevancia es tal que diversos estudios han puesto de manifiesto que ya en los años preescolares, aspectos como la capacidad para permanecer sentado, prestar atención o recordar y seguir reglas -cuestiones en las que quedan implicadas las FE-, son más importantes para la adaptación, aprendizaje y éxito en el contexto escolar posterior que las habilidades matemáticas o lingüísticas tempranas, o incluso que el propio IQ (Viterbori, Usai, Traverso y De Franchis, 2015; Wass, 2015). Pero las implicaciones que tienen las FE tempranas en el desarrollo y aprendizaje humano abarcan mucho más allá de la etapa escolar, de modo que un buen desarrollo ejecutivo temprano contribuye de manera importante al éxito laboral, personal, familiar y social en la adultez, así como a un mejor estado de salud y un nivel socioeconómico mayor (Miyake y Friedman, 2012). Es por todo ello que algunos autores afirman que conocer el desarrollo temprano de las FE es cuestión imprescindible para poder llegar a comprender el desarrollo y aprendizaje infantil (Moriguchi, Chevalier y Zelazo, 2016), e incluso el desarrollo y aprendizaje humano en toda su extensión.

Entre los procesos o componentes que conforman las FE, los modelos más actuales incluyen los siguientes (Diamond, 2013; Miyake et al., 2000; Zelazo y Carlson, 2012): la memoria de trabajo (guardar la información en mente durante un corto periodo de tiempo, a la vez que la manipulamos mentalmente); inhibición (suprimir una respuesta predominante o un estímulo

irrelevante para la tarea en curso) y flexibilidad cognitiva (cambiar de manera rápida y adaptativa el curso de pensamiento o la acción de acuerdo a las demandas de la situación continuamente cambiante). Las FE comienzan a desarrollarse en la infancia temprana, -concretamente, hacia el final del primer año de vida-, mostrando un rápido desarrollo entre los 2 y los 5 años. Continúan desarrollándose a un ritmo menor hasta la adolescencia, momento en el que muestran otra mejora importante. Posteriormente continúan mejorando de modo más gradual, hasta alcanzar el máximo nivel de desarrollo alrededor de los 20 años (Best y Miller, 2010; Brydges, Anderson, Reid y Fox, 2013; Flores-Lázaro, Castillo Preciado y Jiménez-Miramonte, 2014). Este desarrollo de las FE es paralelo a la maduración de su principal base neuroanatómica: el córtex prefrontal dorsolateral (Funahashi y Andreau, 2013).

De entre todas las FE cognitivas, la inhibición es considerada por numerosos trabajos como uno de los componentes fundamentales de las mismas (Barkley, 2012; Miyake et al., 2000), así como un importante componente de la conducta humana (Albert, López-Martí y Carretié, 2010), de la inteligencia (Duan, Wei, Wang y Shi, 2010) y un mecanismo clave para la adaptación del ser humano, contribuyendo al éxito en la vida diaria (Petersen, Hoyniak, McQuillan, Bates y Staples, 2016). Más específicamente, y en referencia concreta a la infancia, la inhibición es el mejor predictor de la conducta infantil y de las competencias socioemocionales. Además, la inhibición evaluada en edades preescolares constituye un buen predictor del nivel de habilidades matemáticas y aspectos lectoescritores posteriores, pudiendo prevenir a partir de la optimización de la misma dificultades de aprendizaje posteriores (Stievano y Valeri, 2013).

Son muchos los autores que defienden que la inhibición constituye un constructo multidimensional. Se trata de una familia de procesos inhibitorios, separados pero relacionados (Brydges et al., 2013), siendo varias las propuestas existentes sobre los mismos (Dempster y Corkill, 1999; Friedman y Miyake, 2004; Harnishfeger, 1995; Howard, Johnson y Pascual-Leone, 2014; Nigg, 2000; Nee y Jonides, 2008). De entre estos procesos inhibitorios, este trabajo se centra en la denominada resistencia a la interferencia distractora (Dempster, 1993; Friedman y Miyake, 2004), menos estudiada que otros componentes inhibitorios y apenas abordada desde una perspectiva longitudinal. Supone la habilidad para resistir la interferencia o distracción generada por una información o estímulo externo que es irrelevante para resolver la tarea que se está llevando a cabo y que puede producir un decremento en su ejecución. Exige que la persona seleccione la información o estímulo adecuado para resolver la tarea, desatendiendo a otros estímulos distractores y competidores que aparecen junto a aquél (Mishra, Anguera, Ziegler y Gazzaley, 2014).

A pesar de las importantes implicaciones que tiene el desarrollo temprano de la resistencia a la interferencia, en particular, y de las FE en general, su emergencia de los 0 a los 3 años apenas ha sido estudiada (Hendry, Jones y Charman, 2016). Los trabajos que analizan las FE infantiles lo hacen fundamentalmente a partir de los años preescolares (es decir, a partir de los 3-6 años) y en periodos posteriores, pero son muy pocos aquellos que lo hacen en las edades más tempranas. Además, los estudios sobre las FE pocas veces adoptan una perspectiva longitudinal (Best y Miller, 2010; Willoughby, Holochwost, Blanton y

Blair, 2014; Willoughby, Wirth y Blair, 2011) siendo mucho más frecuentes los diseños *cross-seccionales* (Shanmugan y Satterthwaite, 2015). Así, nos encontramos que, a pesar de la abundante proliferación de trabajos en los últimos años sobre las FE, estas constituyen un área de estudio donde los aspectos conceptuales se encuentran en mejor estado de conocimiento que los aspectos referidos a su desarrollo y medición (Willoughby y Blair, 2016). Tal vez a ello contribuyan las dificultades propias que implica apresar el proceso de desarrollo y cambio en sí mismo (Isquith, Gioia y Espy, 2004), así como las dificultades que supone el estudio de los procesos mentales (García Molina, Tirapu Ustároz y Roig Rovira, 2007), pues dada su inaccesibilidad a la observación directa, son sus productos externos los que deben observarse directamente, para a partir de estos inferir aquellos (Willoughby et al., 2011). A ello se suma además la complejidad y las dificultades inherentes de trabajar con participantes de muy corta edad (Clark, Flewitt, Hammersley y Robb, 2013), como por ejemplo: la inestabilidad de la conducta infantil, sus periodos cortos de atención, la alta fluctuación de su motivación (Aslin y Fiser, 2005), su inmadurez para cooperar consistentemente (Field y Behrman, 2004) o evidentes dificultades de expresión verbal (Salley, Panneton y Colombo, 2013). Todo ello contribuye a la dificultad y complejidad que implica la obtención de datos fiables y válidos referidos al desarrollo y aprendizaje infantil y, por tanto, al hecho de que sean escasos los estudios acerca del desarrollo cognitivo infantil. Además, en estos estudios, nos encontramos con frecuencia que las muestras son de pequeño tamaño, pues a todas estas dificultades hay que añadir cuestiones éticas y legales, así como el hecho de que, en ocasiones, los padres pueden ser reacios a permitir la participación de su hijo en las investigaciones (Alderson, 2004; Shaw, Brady y Davey, 2011).

Estas características propias de la infancia más temprana hacen que el rango de procedimientos metodológicos que pueden utilizarse para su estudio y evaluación esté ampliamente limitado, siendo la metodología observacional una de las más adecuadas, e incluso con frecuencia, la única posible de aplicar (Anguera, 2001, 2010; Bryce y Whitebread, 2012; Herrero, 1992; Whitebread y Coltman, 2010). A pesar de los numerosos e importantes beneficios que implica su utilización, también es cierto que conlleva ciertos inconvenientes o dificultades, como es su coste temporal, reflejado por ejemplo, en la recogida y registro de datos, o en el proceso minucioso de formación que debe adquirir todo observador (Anguera, 2010). Estos aspectos, junto a los mencionados anteriormente, pueden estar contribuyendo asimismo al escaso estudio de los procesos cognitivos en los primeros años de vida.

Dado que una adecuada evaluación temprana de la actividad y competencias infantiles es esencial para potenciar un desarrollo y aprendizaje óptimos (pues permite identificar posibles dificultades y planificar las consiguientes acciones dirigidas a la prevención de futuras consecuencias adversas), este trabajo tiene como objetivo presentar las posibilidades y ventajas que conllevan nuevas técnicas de análisis de datos observacionales en el ámbito, tanto para controlar la calidad de los datos infantiles registrados a través de observación sistemática como para analizar su variabilidad. Más específicamente, este estudio se centra en nuevos procedimientos para calcular la fiabilidad (intra e inter-observadores) y la validez de los datos observacionales, así co-

mo la generalizabilidad de los resultados obtenidos a otra muestra de las mismas características, pero de tamaño distinto. Asimismo, se presentan las ventajas de otras técnicas (Modelo Lineal General y MIXED) para analizar la variabilidad de dichos datos.

Esperamos que este trabajo contribuya a la mejora de la medición del desarrollo y adquisición de las competencias infantiles, cuestión ampliamente reclamada desde la literatura más actual sobre el tema, dada la existencia todavía de ciertas limitaciones en el ámbito (Carlson, Faja y Beck, 2016; Escolano-Pérez y Blanco-Villaseñor, 2015; Willoughby, Wirth, Blair y Family Life Project Investigators, 2016).

Método

Diseño

Se siguió un diseño observacional nomotético, multidimensional y de seguimiento (Anguera, Blanco-Villaseñor, Hernández-Mendo y Losada, 2011; Anguera, Blanco-Villaseñor y Losada, 2001). Su justificación es la siguiente: nomotético, dado que se estudiaron diversos participantes; multidimensional, al observarse distintas dimensiones de la conducta infantil referidas al desarrollo de la correspondencia uno a uno y a la resistencia a la interferencia distractora; y de seguimiento, puesto que cada participante se estudió en tres momentos temporales (a los 18, 21 y 24 meses de edad).

Participantes

La muestra está compuesta por 48 participantes evaluados longitudinalmente a los 18, 21 y 24 meses. Se trataba de niños con un desarrollo normotípico. No presentaron factor de riesgo o patología al nacer, así como tampoco ninguna complicación durante el periodo pre, peri o postnatal (Grupo de Atención Temprana, 2000). Todos procedían de familias con un nivel socioeconómico medio-alto y asistían al mismo centro educativo de carácter privado.

La selección de la muestra fue realizada a través de un muestreo no probabilístico intencional entre todos aquellos participantes del centro educativo que cumplían las características arriba mencionadas (edades objeto de estudio; desarrollo normotípico; ausencia de factor de riesgo, patología y complicaciones pre, peri o postnatales) y cuyos padres firmaron el consentimiento informado para permitir la participación de su hijo en el estudio. Todos los participantes fueron tratados cumpliendo las normativas y principios éticos internacionales de investigación científica.

Instrumentos

Los instrumentos utilizados, de diferente tipología, fueron los siguientes:

1.- Como material estimular se diseñaron *ad hoc* tres tareas lúdicas de contenido no verbal cuyos elementos facilitan la generación, a través de la acción libre del niño, de la correspondencia uno a uno. Además, para la resolución de la última tarea, dada las propiedades de sus elementos, es necesario que el participante ponga en marcha asimismo su capacidad de resistencia a la interferencia distractora.

En concreto, la generación de la correspondencia uno a uno se ve facilitada porque el material que compone cada tarea constituye dos conjuntos de elementos que guardan entre sí relaciones de inserción. Más exactamente, se trata de un conjunto de 4 cubiletes de diferentes tamaños y un conjunto de 4 pelotas de los mismos tamaños que los cubiletes. Puesto que cada una de las pelotas puede ser introducida en el cubilete de su mismo tamaño (relación de inserción), esta propiedad de los elementos facilita el establecimiento de la correspondencia uno a uno. Además, de acuerdo a la literatura, el número de conjuntos presentados, dos, y el número de elementos que componen cada uno de ellos, cuatro, es adecuado para poder apresar el proceso que desarrollan los niños hasta llegar a alcanzar, la mayor parte de ellos, la correspondencia uno a uno a los 24 meses. Por otra parte, los cuatro cubiletes y las cuatro pelotas varían de color del siguiente modo: en la primera tarea, cada cubilete y cada pelota del mismo tamaño, además de compartir tamaño, comparten también color (cuestión que ayuda en su puesta en relación y correspondencia uno a uno); en la segunda tarea, todos los cubiletes y todas las pelotas son blancas, lo que hace más difícil que en la tarea anterior el establecimiento de la correspondencia uno a uno, dado que para ello ahora solo se dispone de la información referida al tamaño, y no del color; y en la tercera tarea, el cubilete y pelota del mismo tamaño presentan color diferente, pero estos colores son iguales a los de otros elementos de otro tamaño. Así, el color, que se percibe antes que el tamaño, "distra" la atención del niño y afecta su ejecución en la tarea, pues no permite poner cada una de las pelotas dentro de un cubilete (por ejemplo, la pelota más grande es roja, y el cubilete más pequeño también es rojo). El niño debe resistir la interferencia que genera el color para centrarse en el tamaño, característica que permite establecer la correspondencia uno a uno entre los cubiletes y las pelotas poniendo cada una de estas dentro del cubilete de su mismo tamaño.

2.- Además, atendiendo a los instrumentos propios de la observación sistemática, se utilizaron los expuestos a continuación.

a) Como instrumento tecnológico, una cámara de video digital.

b) Instrumento de observación: *ELEDA: Early Logical and Executive Development Assessment* (Escolano-Pérez y Sastre-Riba, 2010), que consiste en un instrumento mixto compuesto por formato de campos y categorías que permite apresar la actividad lógica y ejecutiva infantil, especialmente aquella referida a la construcción de la correspondencia uno a uno y de la resistencia a la interferencia distractora. A modo de ejemplo, algunos de los criterios y categorías que lo conforman son:

* *Contenido*: este criterio se refiere al tipo de actividad lógica que, siguiendo a Sinclair et al. (1984) y Langer (1986), constituye las actividades previas, aproximativas y necesarias en el desarrollo del establecimiento de correspondencias uno a uno hasta lograr ésta. Dado que la correspondencia uno a uno constituye un aspecto esencial para el estudio del desarrollo de la actividad lógica infantil, y consecuentemente también para esta investigación, es imprescindible incluir en el instrumento de observación actividades lógicas indicadoras de su curso o grado de desarrollo. Este criterio está constituido por 8 categorías exhaustivas y mutuamente excluyentes. Algunas de ellas son:

- «Agrupación»: Reunión de varios objetos pertenecientes a diferentes conjuntos (Sinclair et al., 1984).

- «Colección»: Reunión de varios objetos pertenecientes a un mismo conjunto (Sinclair et al., 1984).
- «Distribución uno a varios/todos»: Un elemento de un conjunto se relaciona secuencialmente con varios o con todos los elementos del otro conjunto.
- «Distribución varios/todos a varios/todos»: Varios o todos los elementos de un conjunto se relacionan individual y secuencialmente con varios o todos los elementos del otro conjunto.

* *Adecuación*: Este criterio informa acerca de la concordancia, o no, de las propiedades de tamaño y color entre los elementos relacionados de los dos conjuntos del material de estímulo. Por tanto, informa sobre el papel facilitador del color en la resolución de la tarea 1 y sobre el papel interferente del color en la resolución de la tarea 3. En éste último caso (tarea 3) evalúa la capacidad del participante para resistir, o no, la interferencia generada por el estímulo distractor (el color). Este criterio se compone de 6 categorías exhaustivas y mutuamente excluyentes. Algunas de estas categorías son:

- «Adecuación del tamaño y color»: En todas las puestas en relación existe concordancia entre el tamaño y color de un elemento de un conjunto y el tamaño y color del elemento del otro conjunto con el que se relaciona. Dada las características del material, solo es posible de aparición en la tarea 1. Esta categoría permite evaluar el papel facilitador del color en la resolución de la tarea 1.
- «Adecuación del tamaño pero no del color»: Todos los elementos de un conjunto relacionados con los elementos del otro conjunto comparten tamaño pero no color. Dada las características del material, solo es posible de aparición en la tarea 3. Implica la capacidad del niño para resistir la interferencia del color.
- «Adecuación del color pero no del tamaño»: Todos los elementos de un conjunto relacionados con los elementos del otro conjunto comparten color pero no tamaño. Dada las características del material estimular, únicamente es posible en la tarea 3 e implica la no resistencia a la interferencia generada por el color.

* *Extensión*: Este criterio hace referencia al número de elementos que participan en la acción infantil o en sus resultados. Está compuesto por dos categorías exhaustivas y mutuamente excluyentes:

- «Exhaustiva»: Participación en la acción infantil, o en sus resultados, de todos los elementos.
- «No exhaustiva»: Participación en la acción infantil, o en sus resultados, de solo algunos de los elementos.

Este instrumento de observación puede consultarse en toda su extensión en Escolano-Pérez y Sastre-Riba (2010).

c) Instrumento de registro: *Match Vision Studio v1.0* (Perea, Alday y Castellano, 2006).

3.- Por otra parte, para el análisis de los datos se utilizaron los programas *SAS 9.1.3* (SAS Institute Inc., 2004; Schlotzhauer y Littell, 1997) y *EduG 6.0-e* (Cardinet, Johnson y Pini, 2010).

Procedimiento

La actividad infantil fue grabada a los 18, 21 y 24 meses cuando cada participante realizaba individualmente cada una de las tres tareas, en el mismo orden en el que fueron presentadas

en el apartado anterior. Cada participante jugaba libremente, sin recibir instrucciones hasta que, voluntariamente, finalizaba su actividad, momento en el que se daba por acabada cada sesión de observación. Todas las sesiones se realizaron en el contexto escolar. El participante estaba acompañado de un adulto que presentaba los materiales de cada tarea, pero que no intervenía en la actividad infantil. En cada tarea, los cubiletes y pelotas eran presentados en el suelo en una distribución espacial aleatoria, tal y como caían al ser vaciados de la bolsa donde se guardaban, prestando especial atención a que nunca una pelota quedara dentro de un cubilete.

Posteriormente, la actividad infantil fue registrada a través del *Match Vision Studio v1.0*, haciendo uso del instrumento de observación creado *ad hoc*. Un mismo observador (persona experta tanto en metodología observacional como en el desarrollo lógico y ejecutivo infantil, autora del propio instrumento de observación y coautora de este manuscrito) llevó a cabo todos los registros codificados de cada una de las sesiones infantiles. Además, otro observador (experto en metodología observacional y en desarrollo y aprendizaje infantil) fue entrenado para la aplicación del ELEDA. Registró un total de 27 sesiones pertenecientes a participantes de las tres edades y en las tres tareas. Algunos de sus registros codificados fueron utilizados para el cálculo de la fiabilidad inter-observadores.

Análisis estadísticos

De acuerdo a las exigencias propias de la metodología observacional, se han realizado análisis estadísticos referidos al control de la calidad de los datos observacionales. Ello constituye un tema de suma importancia en cualquier estudio de metodología observacional y se lleva a cabo en cuatro vertientes: Fiabilidad, Precisión, Validez y Estimación del tamaño de la muestra (o bien conjuntamente a través de un estudio de Generalizabilidad). En este trabajo nos centramos en Fiabilidad, Validez y Estimación del tamaño de la muestra, así como en su análisis conjunto a través de la Generalizabilidad.

Se realizaron asimismo análisis de la variabilidad de los datos observacionales haciendo uso del Modelo Lineal General y del Modelo MIXED.

Control de la calidad de los datos observacionales: Fiabilidad, Validez, Estimación del tamaño de la muestra, Generalizabilidad

Hay diferentes formas de estimar la fiabilidad, como especificaremos a continuación, y cada una de ellas genera un coeficiente diferente. Podemos verificar las puntuaciones dadas por un mismo observador en dos momentos diferentes a la misma sesión de observación (*fiabilidad intra-observador*), o por diferentes observadores en el mismo o diferente período temporal (*fiabilidad inter-observadores*), o en diferentes ocasiones separadas por un corto intervalo de tiempo, o utilizar diferentes formas de una escala (*formas paralelas*) y así sucesivamente. Sin embargo, estas medidas *standard* no agotan todas las posibles fuentes de variación. El control de la calidad de los datos observacionales en el presente trabajo trata precisamente de utilizar una nueva vía de expresar esta variabilidad a través de los conceptos del análisis de la variancia. Ello fue desarrollado en el contexto de la Teoría de la Generalizabilidad (TG) por Cronbach, Gleser, Nanda y

Rajaratnam (1972). Su adecuación a los diseños observacionales se debe a los trabajos de Mitchell (1979), quien estableció claramente la inadecuación de la concordancia entre diferentes observadores como índices de fiabilidad.

Las diferencias entre concordancia (acuerdo) y fiabilidad (correlación) se basan en la forma en que se definen estos índices. Los coeficientes de fiabilidad dividen la variancia de un conjunto de puntuaciones en una puntuación verdadera (diferencias individuales) y un componente de error. Los porcentajes de concordancia inter-observadores, sin embargo, no aportan información sobre las diferencias individuales entre sujetos y tan solo contienen información de una sola de las posibles fuentes de error (diferencias entre observadores). Estas técnicas, por tanto, no permiten atribuir la variancia estimada a los observadores, a las formas diferentes, a las ocasiones, o no pueden considerar estas fuentes de error de forma simultánea. Es necesario, por tanto, una teoría multivariada que tenga en cuenta todas estas posibles fuentes de error, además de las generadas por las diferentes formas empíricas de validez (Blanco-Villaseñor, Castellano, Hernández Mendo, Sánchez-López y Usabiaga, 2014). Es esta precisamente nuestra particular forma de entender la fiabilidad de los registros observacionales y la que utilizamos en nuestro estudio.

En este sentido, una medición observacional puede presentar datos bajo la influencia de un cierto número de aspectos diferentes de una situación observacional (diferentes observadores, diferentes ocasiones, diferentes formas de registro, diferentes criterios del instrumento de observación), incluyendo las diferencias individuales entre participantes. Este punto de vista es la TG desarrollada por Cronbach, Gleser, Nanda y Rajaratnam (1972), que asume que hay otras fuentes de variación además de las diferencias individuales y que permite integrar en una estructura global no solo cada una de las fuentes de variación de los diferentes coeficientes de fiabilidad anteriores sino también las referidas a los criterios del instrumento de observación y a los participantes. En este sentido, si la faceta observador (intra-observador) u observadores (inter-observadores) son consideradas facetas de Instrumentación o de Generalización, estaremos hablando del coeficiente de correlación intraclase y, por tanto, de la fiabilidad de los registros. Si, por el contrario, estas facetas son consideradas facetas de Diferenciación, estamos hablando de la validez del instrumento de observación. Si la faceta participantes es la faceta de Instrumentación, podremos estimar si el tamaño de la muestra es suficiente para poder generalizar los resultados obtenidos a la población de referencia.

Análisis de la variabilidad

El análisis de datos que presentamos responde a la siguiente situación. En el ámbito de la Psicología del Desarrollo y de la Educación se producen habitualmente situaciones objeto de estudio que poco o nada tienen que ver con análisis de datos estandarizados en la literatura científica. Desde nuestro punto de vista, los análisis habituales que utiliza la investigación psicológica y educativa en muchas ocasiones no se adaptan a las ocasiones de registro en el ámbito del desarrollo y aprendizaje (más aún cuando se trata de desarrollo y aprendizaje infantil temprano), dado que ni los tamaños de muestra son los adecuados ni la representatividad de las mismas es totalmente aleatorizada.

Ello tiene que ver también con la pérdida de información (*missing data*) y, por tanto, los procedimientos habituales de cálculo del Modelo Lineal General (*General Linear Model -GLM-*) mediante mínimos cuadrados (*Least Squares -LS-*) tendrían que adaptarse a procedimientos MIXED de máxima verosimilitud (*Maximum Likelihood -ML-*).

En consecuencia, cuando trabajamos con modelos del desarrollo y aprendizaje individual es adecuado analizar datos de participantes en una investigación en los que la misma medida se repite a lo largo de dos ó más puntos de tiempo (denominados habitualmente estudios longitudinales de medidas repetidas). La medición, que siempre será la misma (aunque en diferentes ocasiones, situaciones, días, sesiones, edades...), nos permite contar con técnicas analíticas para las medidas repetidas que tienen en cuenta la co-variabilidad intra-sujeto a lo largo del tiempo.

Resultados

En lo que respecta a la fiabilidad, para el cálculo de los coeficientes de fiabilidad intra e inter-observadores, se han utilizado un total de 16 participantes de los 48 de la muestra (seleccionados aleatoriamente) en las tres edades diferentes (18, 21 y 24 meses) y en las distintas tareas. Ello supone un total de 16 sesiones de observación, de las que 10 han sido usadas para la fiabilidad intra-observador y 6 para la inter-observadores, según puede verse de forma individual en las Tablas 1 y 2, y un resumen de todos los participantes utilizados en las Tablas 3 y 4.

En todas las sesiones se han utilizado los coeficientes de generalizabilidad a través del programa de software EduG 6.0-e de un único Plan de Medida de tres facetas (Observadores -O-, criterios o Macro-categorías -M- y Categorías -C-), donde el observador (intra-observador) o los observadores (inter-observadores) siempre han sido la faceta de Instrumentación. Por tanto, el Plan de Medida $O \times M \times C$ tiene dos facetas de Diferenciación (MC) y una faceta de Instrumentación (O). En realidad, esta fórmula de cálculo coincide simultáneamente con el Coeficiente de Correlación Intraclase (CCI), dado que permite reducir el sesgo que pueden producir los observadores si uno de ellos registra valores o puntuaciones más altas o más bajas, pero de forma constante.

Podemos concluir, de forma general, que los resultados son muy satisfactorios en ambas modalidades (intra e inter), puesto que los CCI oscilan entre valores muy altos (0.96 – 0.99), sobre todo si tenemos en cuenta que no son coeficientes de fiabilidad basados en el error absoluto y que no se trata de porcentajes.

En el caso de las 10 sesiones de la Tabla 3 (fiabilidad intra-observador), los valores de los coeficientes casi rozan la unidad (0.98 – 0.99). Por tanto, se trata de coeficientes intra-observador excelentes, donde la misma sesión de observación ha sido visionada por el mismo observador en dos momentos diferentes.

Podemos observar, en este sentido, las diferentes tablas individualizadas donde el porcentaje de variabilidad de la faceta observadores es nula (0%) en 9 de las 10 sesiones. Ello indica que los resultados son casi similares en cada una de las sesiones visionadas por el mismo observador en dos ocasiones diferentes. La estimación de resultados en un número infinito de ocasiones nos permite afirmar que este observador apenas cometería errores en sus valoraciones.

Las diferencias entre una y otra sesión se deben a errores

producidos por el observador al codificar de forma general en las macro-categorías o criterios (O x M), aunque en ningún caso suponen una variabilidad mayor del 1%. Donde se producen algunos errores más es, lógicamente, en la codificación en una categoría determinada (O x C), aunque los valores también son nulos en algunas sesiones y de un 3% en otras.

En cualquier caso, todas las estructuras han permitido conocer que no hay otras causas diferentes al observador que hayan permitido graves sesgos o errores en la observación sistemática. En las 10 sesiones, el error residual (O x M x C) o parte de variabilidad que desconocemos es nula.

La Tabla 4 (fiabilidad inter-observadores: dos observadores, actuando independientemente, en el mismo o diferente momento temporal, codifican la misma sesión de observación) recoge también resultados excelentes, aunque los valores de los coeficientes son un poco inferiores (0.96 – 0.99). Es lógico que los valores sean relativamente inferiores a la fiabilidad intra-observador, dado que las percepciones de dos observadores diferentes pueden contener algún error más que la de un único observador. De hecho, no podría calcularse la fiabilidad inter-observadores si la intra-observador no tuviera valores cercanos a la unidad, como así ha sido en nuestro caso.

Podemos ver en las seis sesiones que los errores de los observadores al codificar en las macro-categorías o criterios (O x M) suponen valores similares a los obtenidos en la fiabilidad intra-observador. Sin embargo, sí que han supuesto una mayor variabilidad las codificaciones específicas en cada categoría, alcanzando en un único caso una variabilidad atribuida al observador del 5%, siendo también inferior en este caso el coeficiente de generalizabilidad. Exceptuando este caso, los valores son similares a la intra-observador.

Tabla 1. Ilustración de la Fiabilidad Intra-Observador mediante TG del Participante 28 Tarea 3 y sus Coeficientes de Generalizabilidad (CG), relativo y absoluto.

Fuentes de variación	SC	CM	%	CG (rel, abs)
Observador	0.6621	0.0008	0	
Macro	4278.2820	6.4213	46	
O x M	1.6920	0.0050	0	0.99
Categorías	5324.3125	7.3863	53	0.99
O x C	6.1458	0.0170	0	
M x C	0	0	0	
O x M x C	0	0	0	

Tabla 2. Ilustración de la Fiabilidad Inter-Observadores mediante TG del Participante 28 Tarea 2 y sus Coeficientes de Generalizabilidad (CG), relativo y absoluto.

Fuentes de variación	SC	CM	%	CG (rel, abs)
Observadores	3.4594	0.0053	0	
Macro	1048.2167	1.5629	47	
O x M	7.3113	0.0219	1	0.987
Categorías	1250.7291	1.7034	51	0.986
O x C	24.2291	0.0673	2	
M x C	0	0	0	
O x M x C	0	0	0	

Tabla 3. Resumen de 10 Sesiones de Fiabilidad Intra-Observador, con valores de Coeficiente de Generalizabilidad mediante TG y Coeficiente de Correlación Intraclase (CCI).

Sesiones Intra (%)	1	2	3	4	5	6	7	8	9	10
Observador	0	0	0	0	1	0	0	0	0	0
O x Macro	1	1	0	1	1	1	1	0	2	1
O x Categorías	0	2	0	3	3	0	3	3	3	3
O x M x C	0	0	0	0	0	0	0	0	0	0
CCI ó CG	.993	.987	.99	.98	.98	.99	.979	.984	.987	.982

Tabla 4. Resumen de 6 Sesiones de Fiabilidad Inter-Observador, con valores de Coeficiente de Generalizabilidad mediante TG y Coeficiente de Correlación Intraclase (CCI).

Sesiones Inter (%)	1	2	3	4	5	6
Observadores	0	0	0	0	0	0
O x Macro	3	2	1	3	0	1
O x Categorías	1	5	2	4	1	1
O x M x C	0	0	0	0	0	0
CCI ó CG	.981	.964	.983	.959	.996	.988

De otra parte, las estructuras de diseño utilizadas para el cálculo de la fiabilidad pueden servirnos simultáneamente para hacer una valoración de la validez de los criterios y las categorías que han conformado nuestro instrumento de observación (lo que habitualmente se denomina validez del instrumento de observación).

En este sentido, en lo que respecta a la validez, la Tabla 5 representa una máxima variabilidad (cercana al 96%) si tomamos conjuntamente las facetas macro-categorías y categorías, siendo el residual siempre de nula variabilidad. Ya sean las macro-categorías o criterios, ya sean las categorías, nos hace ver que su máxima variabilidad en todas las estructuras de diseño, nos permiten discernir adecuadamente lo que se ha codificado o valorado en una u otra categoría. Por tanto, si en cualquiera de las sesiones presentáramos mediante la Teoría de la Generalizabilidad (TG) un Plan de Medida donde la faceta de Generalización o Instrumentación fueran las macro-categorías o las categorías (ya sean individualmente, ya sean conjuntamente), en todos los casos el coeficiente tendría un valor cercano a 0. Se ha decidido obviar todos los resultados de estas tablas dado que los valores son casi iguales en todas ellas (alta variabilidad en macro-categorías y categorías y valor residual nulo).

Nuestra aseveración tiene aún mayor sentido con los resultados reales que revelamos en nuestras diferentes sesiones de observación, es decir, el instrumento de observación que se ha creado para esta investigación es válido para registrar aquello que habíamos definido previamente en el mismo (basándonos en la literatura científica y en sus constructos hipotéticos). Es evidente que las facetas macro o categorías tienen diferente significado en los Planes de Medida: si son facetas de Diferenciación obtenemos las más altas diferencias entre ellas con respecto a las otras facetas; si son facetas de Instrumentación los coeficientes tienen valores nulos casi siempre. Por tanto, nuestro instrumento de observación tiene validez.

Tabla 5. Resumen de 10 Sesiones para el cálculo de la Validez Intra-Observador.

Sesiones Intra (%)	1	2	3	4	5	6	7	8	9	10
Macro	61	47	46	37	46	39	28	52	29	33
Observador x M	1	1	0	1	1	1	1	0	2	1
Categorías	38	51	53	59	49	60	68	45	66	63
O x C	0	2	0	3	3	0	3	3	3	3
M x C	0	0	0	0	0	0	0	0	0	0
O x M x C	0	0	0	0	0	0	0	0	0	0
CG	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0

En suma, el análisis de las diferentes sesiones de observación mediante la TG nos permite contar con una misma estructura de diseño de tres facetas (observadores, macro-categorías y categorías, individualmente o en interacción) para analizar mediante filtros, que son clave en la obtención de una alta calidad de los datos (fiabilidad intra, fiabilidad inter y validez) la veracidad de los mismos. Así, este control de la calidad de los datos obtenidos mediante observación directa del comportamiento permite salvaguardar esta misma calidad en los análisis de datos que se presentan posteriormente. Dada la dificultad que entraña la recogida de datos longitudinales en estudios con bebés se han intensificado estos aspectos de control de la calidad. La información así registrada tiene calidad, una vez aplicados los controles previos (fiabilidad y validez), y la TG nos ha permitido estructurar todo ello de forma conjunta en una misma unidad de análisis.

En lo que respecta al tamaño más adecuado de la muestra, podemos observar en la Tabla 6 que el modelo contiene un error residual (Edad x Tarea x Participantes -E x T x P-) alto, 39%. Es decir, un 39% de la variabilidad de los resultados encontrados se deben a variables no contempladas en el modelo, y por tanto, desconocidas, por lo que, deberían introducirse más

Tabla 7. Plan de Optimización Edad Tarea/Participantes (ET/P) para el cálculo del Coeficiente de Generalizabilidad del tamaño más adecuado de muestra de participantes en nuestro estudio.

Faceta	Niveles	Tamaño	Opt. 1	Opt.2
Participantes	$n_p = 48$	$N_p = \infty$	60	70
Coeficiente de Generalizabilidad = 0.89			0.91	0.92

Una de las opciones más interesantes que pueden servir para amplificar nuestras investigaciones es la utilización de modelos de curvas de desarrollo individual, que han sido diseñados con el fin de explorar datos longitudinales de participantes en estudios donde la medida se toma repetidamente en diferentes momentos temporales (McArdle y Nesselroade, 2003). Estos modelos permiten al investigador explorar dos niveles de variabilidad de las variables de respuesta: intra e inter-sujetos (ver Tabla 8 para el tratamiento de datos univariados y Tabla 9 para el análisis multivariado donde se tiene en cuenta la variabilidad intra-sujeto a lo largo del tiempo, aunque con la desventaja en este caso de que no se tienen en cuenta los datos con pérdida de información -missing data-, algo habitual en investigaciones de Psicología del Desarrollo y de la Educación). En ambas tablas, todas las facetas y sus interacciones ofrecen resultados significati-

variables o facetas en el modelo para poder conocer a qué se deben los resultados encontrados y así, explicar en mayor medida las variables que contribuyen a la actividad lógica y ejecutiva infantil.

Tabla 6. Estimación del tamaño ideal de muestra. Plan de Estimación Edad x Tarea x Participantes (E x T x P) mediante la TG.

Fuentes de variación	Componentes de Variancia	
	Estimados	% Total Variancia
Edad	16.0436	0
Tarea	192.3609	9
Edad x Tarea	51.5028	4
Participantes	695.6675	11
Edad x Participantes	416.4789	12
Participantes x Tarea	646.7291	25
E x T x P	440.4757	39

La Tabla 7 ofrece el coeficiente de generalizabilidad de los resultados obtenidos con el total de participantes estudiados ($n_p = 48$) y los necesarios en el Plan de Optimización. Tal como se observa, el coeficiente de generalizabilidad de los resultados obtenidos es de .89, es decir, alto. Ello permite concluir que los resultados obtenidos en esta investigación pueden generalizarse con cierto grado de precisión a la población de pertenencia. Para conocer si esta precisión es la óptima o puede mejorarse, se ha realizado un Plan de Optimización con dos proyecciones diferentes: aumentando el tamaño de la muestra a $n = 60$ y a $n = 70$ participantes. Los valores obtenidos son de .91 y .92 respectivamente. Se considera que esta leve optimización del coeficiente de generalizabilidad no compensa el alto coste que supondría aumentar el número de participantes exigidos, teniendo en cuenta además su temprana edad.

vos, a excepción de la Edad (contrariamente a lo esperado) y de la interacción Participantes x Acciones.

Dado que los datos son longitudinales, las observaciones que tienen lugar en diferentes puntos temporales están anidadas intra-sujeto y así la población de interés tiene una estructura jerárquica de dos niveles: la variabilidad longitudinal intra-sujeto estaría en el nivel más bajo y la variabilidad inter-sujetos en el más alto. Las estimaciones de las matrices de covariancia nos expresarán los porcentajes correspondiente a la curva de desarrollo, teniendo en cuenta que los programas informáticos actuales permiten realizar ajustes más precisos al realizar estimaciones por máxima verosimilitud (ML). En esta situación presentamos una situación analítica más adecuada a este tipo de datos en la Tabla 10, mediante el procedimiento MIXED de SAS que nos ha permitido tener en cuenta los datos perdidos que se presentan en este estudio.

Tabla 8. Análisis de datos univariado mediante el Modelo Lineal General (GLM).

Fuentes de variación	GL	SC	CM	F-Valor	Pr > F
Edad	2	16.043689	8.021844	2.42	.0890
Tarea	2	192.360989	96.180495	29.03	< .0001
Participantes	47	694.293450	14.772201	4.46	< .0001
Edad x Tarea	4	52.876962	13.219240	3.99	.0031
Edad x Participantes	56	416.478932	7.437124	2.24	< .0001
Edad x Acciones	126	2945.405353	23.376233	7.06	< .0001
Participantes x Acciones	1974	6721.256565	3.404892	1.03	.2451

Tabla 9. Análisis de datos multivariado GLM de medidas repetidas donde no se han tenido en cuenta los datos perdidos: Test de hipótesis para efectos inter-sujetos y Test de hipótesis univariante para efectos intra-sujeto.

Fuentes de variación	GL	SC	CM	F-Valor	Pr > F
Participantes	20	1286.46167	64.32308	9.65	<.0001
Tarea	2	397.10068	198.55034	29.79	<.0001
Error	2923	19480.68273	6.66462		

Fuentes de variación	GL	SC	CM	F-Valor	Pr > F
Edad	2	13.78681	6.89341	2.25	.1060
Edad x Participantes	40	925.54614	23.13865	7.54	< .0001
Edad x Tarea	4	110.36546	27.59137	8.99	< .0001
Error (edad)	5846	17947.16869	3.06999		

Tabla 10. Análisis de datos multivariado mediante el procedimiento MIXED de SAS, test de efectos fijos, donde se han tenido en cuenta los datos perdidos.

Fuentes de variación	GL	F-Valor	Pr > F
Edad x Tarea	6	2.89	0.0082
Participantes	47	4.37	<.0001
Tarea	2	27.10	<.0001
Acciones	42	18.46	<.0001
Edad x Participantes	58	2.16	<.0001

Discusión

Algunas diferencias individuales cognitivas y conductuales se remontan ya a los primeros meses de vida (Bornstein, 2014), lo que supone un aumento potencial del riesgo de desarrollar trayectorias posteriores atípicas, y por tanto, incrementa la necesidad de llevar a cabo una adecuada y minuciosa evaluación del desarrollo y aprendizaje infantil lo más temprana posible, para posteriormente, en base a sus resultados, diseñar e implementar las intervenciones en este mismo período del desarrollo, cuando el cerebro es más maleable y respondiente (Karmiloff-Smith et al., 2014; Wass, 2015).

En participantes de tan corta edad, la metodología observacional es la única posible de aplicar para apresar su desarrollo y aprendizaje. A pesar de las numerosas e importantes ventajas que implica la observación sistemática para el estudio de la conducta y capacidades infantiles tempranas, ha sido escasamente utilizada. Así por ejemplo, a pesar del interés y expansión que ha tenido en los últimos años el estudio de las FE infantiles, la literatura científica revela que como instrumentos de medida se utilizan generalmente, o bien tareas ejecutadas puntualmente por los niños en un contexto clínico o de laboratorio, o bien inventarios o cuestionarios de calificación de la conducta cumplimentados por terceros informantes (padres y profesorado). Ambos tipos de medida conllevan limitaciones. Las tareas desarrolladas en laboratorio quedan lejos del contexto natural en el que se

desarrolla y aprende el niño, apresando su conducta en un contexto artificial y controlado, por lo que este tipo de medidas conllevan una baja validez ecológica (Miranda, Colomer, Mercader, Fernández y Presentación, 2016). Aunque la información procedente de terceros ofrece la ventaja de poder proporcionar información sobre la conducta de los niños en un mayor número de situaciones, su fiabilidad puede quedar en entredicho, especialmente debido a la influencia de la deseabilidad social, posibles fallos de recuerdo y memoria de los informantes e incluso falta de familiaridad y sensibilidad para percibir y detectar determinadas conductas infantiles (Wertz, 2014). Ambas limitaciones quedan superadas por el uso de la observación sistemática, caracterizada por apresar la conducta espontánea del individuo en su contexto natural y habitual (alta validez ecológica) y además por un observador, o varios, experto tanto en el propio ámbito objeto de estudio como en la propia metodología, es decir, tanto en el "qué" como en el "cómo" evaluar. En este sentido, el "observador no nace sino que se hace" (Anguera, 2010), debiendo seguir un proceso de minuciosa formación que, en lo deseable, no debiera finalizar en lo referido a la recogida de datos sino extenderse más allá de estas primeras fases del proceso de evaluación del desarrollo y aprendizaje infantil.

Las primeras fases de este proceso de evaluación determinan las siguientes, pudiendo dar lugar a decisiones determinantes para el desarrollo y aprendizaje del individuo, por lo que verificar el control de la calidad de los datos registrados es una de las primeras cuestiones a asegurar. Para ello debemos introducir las mejoras metodológicas que van apareciendo, y que no solo se limitan a estas primeras fases del proceso sino a toda su extensión. En este estudio hemos presentado algunas de ellas.

Los análisis de medidas repetidas que utilizamos en nuestro estudio han sido habituales en los paquetes de programas estadísticos más estandarizados como SAS y SPSS (Mushquash y O'Connor, 2006). Sin embargo, el procedimiento GLM del SAS conlleva sólo análisis tradicionales uni y multivariados. Entendemos que en próximos trabajos podríamos contar con nuevos procedimientos y estructuras desarrollados en SAS (MIXED) que utilizan una mejor aproximación a estos modelos a través de estructuras generales de covarianza (Castellano, Blanco-Villaseñor y Álvarez, 2011).

En este sentido, nuestras investigaciones futuras tendrían que tener en cuenta la solución a las particularidades del procedimiento GLM (que limita mucho los tipos de análisis, dada la dicotomía entre e intra-sujeto), y entonces analizar la estrategia de medidas repetidas mediante el procedimiento MIXED (SAS Institute Inc., 2004), que además permite analizar, en lugar de ignorar, aquellas ocasiones de medida en que se han producido una pérdida de datos. La razón de ello es que utiliza una estimación por el método de máxima verosimilitud (*Maximum Like-*

likelihood -ML-), en lugar de mínimos cuadrados (*Least Squares* -LS-), que utiliza el GLM y que requiere datos completos (Schlotzhauer y Littell, 1997). En este estudio hemos intentado una aproximación analítica de este tipo, aunque consideramos que la mejor opción sería similar a ésta pero a través de técnicas analíticas conocidas como modelos jerárquicos multinivel. Habitualmente son conocidos en la literatura científica con el nombre de estudios longitudinales de medidas repetidas (curvas de desarrollo, modelos de desarrollo latente) y tienden a comparar los procesos de estabilidad y cambio simultáneamente en las personas y en los grupos que conforman. De este modo podremos examinar más precisamente el conjunto de interacciones que conforman desarrollo cognitivo infantil.

Referencias

- Albert, J., López-Martín, S., y Carretié, L. (2013). Emotional context modulates response inhibition: Neural and behavioral data. *NeuroImage*, 49(1), 914-921.
- Alderson, P. (2004). Ethics. In S. Fraser, V. Lewis, S. Ding, M. Kellett, y C. Robinson (Eds.), *Doing research with children and young people* (pp. 97-111). London: Sage.
- Anguera, M. T. (2001). Cómo apresar las competencias del bebé mediante una aplicación de la metodología observacional. *Contextos Educativos*, 4, 13-34.
- Anguera, M. T. (2010). Posibilidades y relevancia de la observación sistemática por el profesional de la Psicología. *Papeles del Psicólogo*, 31(1), 122-130. Recuperado de <http://www.papelesdelpsicologo.es/pdf/1802.pdf>
- Anguera, M. T., Blanco-Villaseñor, A., Hernández-Mendo, A., y Losada, J. L. (2011). Diseños observacionales: ajuste y aplicación en psicología del deporte. *Cuadernos de Psicología del Deporte*, 11(2), 63-76.
- Anguera, M. T., Blanco-Villaseñor, A., y Losada, J. L. (2001). Diseños observacionales, cuestión clave en el proceso de la metodología observacional. *Metodología de las Ciencias del Comportamiento*, 3(2), 135-160.
- Aslin, R. N., y Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *TRENDS in Cognitive Sciences*, 9(3), 92-98.
- Barkley, R. A. (2012). *Executive Functions: What They Are, How They Work, and Why They Evolved*. New York: Guilford.
- Best, J. R., y Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development*, 81(6), 1641-1660. doi: 10.1111/j.1467-8624.2010.01499.x
- Blanco-Villaseñor, A., Castellano, J., Hernández Mendo, A., Sánchez-López, C.R., y Usabiaga, O. (2014). Aplicación de la TG en el deporte para el estudio de la fiabilidad, validez y estimación de la muestra. *Revista de Psicología del Deporte*, 23(1), 131-137.
- Bornstein, M. H. (2014). Human infancy ... and the rest of the lifespan. *Annual Review of Psychology*, 65, 121-158. doi: 10.1146/annurev-psych-120710-100359
- Bryce, D., y Whitebread, D. (2012). The development of metacognitive skills: evidence from observational analysis of young children's behaviour during problem-solving. *Metacognition and Learning*, 7(3), 197-217.
- Brydges, C. R., Anderson, M., Reid, C. L., y Fox, A. M. (2013). Maturation of cognitive control: delineating response inhibition and interference suppression. *PLoS ONE*, 8, e69826.
- Cardinet, J., Johnson, S., y Pini, G. (2010). *Applying Generalizability Theory using EduG*. Londres: Routledge.
- Carlson, S. M., Faja, S., y Beck, D. M. (2016). Incorporating early development into the measurement of executive function: The need for a continuum of measures across development. In J. A. Griffin, P. McCardle, L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 45-64). Washington, DC, US: American Psychological Association.
- Carlson, S. M., Zelazo, P. D., y Faja, S. (2013). Executive function. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology, Vol. 1: Body and mind* (pp. 706-743). New York: Oxford University Press.
- Castellano, J., Blanco-Villaseñor, A., y Álvarez, D. (2011). Contextual variables and time-motion analysis in soccer. *International Journal of Sports Medicine*, 32(6), 415-421. doi: 10.1055/s-0031-1271771
- Clark, A., Flewitt, R., Hammersley, M., y Robb, M. (2013). *Understanding Research with Children and Young People*. London: Sage.
- Cronbach, L. J., Gleser, G. C., Nanda, H., y Rajaratnam, N. (1972). *The dependability of behavioral measurement: theory of generalizability for scores and profiles*. New York: John Wiley and Sons.
- Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. L. Howe y R. Pasnak (Eds.), *Emerging themes in cognitive development* (Vol. 1, pp. 3-27). New York, NY: Springer-Verlag.
- Dempster, F. N., y Corkill, A. J. (1999). Interference and inhibition in cognition and behavior: unifying themes for education psychology. *Educational Psychology Review*, 11(2), 1-88.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64, 135-168. doi: 10.1146/annurev-psych-113011-143750
- Duan, X., Wei, S., Wang, G., y Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling*, 52, 419-431.
- Escolano-Pérez, E., y Blanco-Villaseñor, A. (2015). The longitudinal measurement of change: Intraindividual variability in behavior and interindividual differences observed in childhood. *Anales de Psicología*, 31(2), 545-551. doi: 10.6018/analesps.31.2.166361
- Escolano-Pérez, E., y Sastre-Riba, S. (2010). Early Infant Cognitive Assessment: Validity of an instrument. *Behavior Research Methods*, 42(4), 759-767.
- Field, M. J., y Behrman, R. E. (Eds.), (2004). *Ethical Conduct of Clinical Research Involving Children*. Washington (DC): National Academies Press.
- Flores-Lázaro, J. C., Castillo Preciado, R. E., y Jiménez-Miramonte, N. A. (2014). Desarrollo de funciones ejecutivas, de la niñez a la juventud. *Anales de Psicología*, 30(2), 463-473. doi: 10.6018/analesps.30.2.155471
- Friedman, N. P., y Miyake, A. (2004). The relations among inhibition and interference control function: a latent-variable analysis. *Journal of Experimental Psychology: General*, 133(1), 101-135.
- Funahashi, S., y Andreau, J. M. (2013). Prefrontal cortex and neural mechanisms of executive function. *Journal of Physiology-Paris*, 6, 471-482.
- García Molina, A., Tirapu Ustároz, J., y Roig Rovira, T. (2007). Validez ecológica en la exploración de las funciones ejecutivas. *Anales de Psicología*, 23(2), 289-299.
- García, T., Rodríguez, C., González-Castro, P., Álvarez-García, D., y González-Pianda, J. A. (2016). Metacognición y funcionamiento ejecutivo en Educación Primaria. *Anales de Psicología*, 32(2), 474-483. doi: 10.6018/analesps.32.2.202891
- Goswami, U. (2010). *The Blackwell Handbook of Childhood Cognitive Development*. New York: John Wiley and Sons.
- Grupo de Atención Temprana (2000). *Libro blanco de la Atención Temprana*. Madrid: Real Patronato de Prevención y de Atención a Personas con Minusvalía. Consultado el 28 de Septiembre, 2015, en <http://www.gat-atienciontemprana.org/>
- Guare, R. (2014). Context in the Development of Executive Functions in Children. *Applied Neuropsychology Child*, 3(3), 226-232. doi: 10.1080/21622965.2013.870015.
- Guralnick, M. J. (2015). Merging policy initiatives and developmental perspectives in early intervention. *Escritos de Psicología*, 8(2), 6-13. doi: 10.5231/psy.writ.2015.1004
- Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster y C. J. Brainerd (Eds.), *Interference and Inhibition in Cognition* (pp. 175-204). San Diego, CA: Academic Press.

- Hendry, A., Jones, E. J. H., y Charman, T. (2016). Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review*, 42, 1-33. doi: 10.1016/j.dr.2016.06.005
- Herrero, M. L. (1992). Posibilidades de la metodología observacional en el estudio analítico de conductas en el aula: Aplicación en escolares con problemas de comportamiento. *Anales de Psicología*, 8(1-2), 149-155.
- Howard, S. J., Johnson, J., y Pascual-Leone, J. (2014). Clarifying inhibitory control: Diversity and development of attentional inhibition. *Cognitive Development*, 31(1), 1-21. doi: 10.1016/j.cogdev.2014.03.
- Iglesias-Sarmiento, V., Carriedo, N., y Rodríguez, J. L. (2015). Updating executive function and performance in reading comprehension and problem solving. *Anales de Psicología*, 31(1), 298-309. doi: 10.6018.analesps.31.1.158111
- Isquith, P. K., Gioia, G. A., y Espy, K. A. (2004). Executive function in preschool children: examination through everyday behavior. *Developmental Neuropsychology*, 26(1), 403-422. doi: 10.1207/s15326942dn2601_3
- Izard, V., Streri, A., y Spelke, E. S. (2014). Toward Exact number: Young children use one-to-one correspondence to measure set identity but not numerical equality. *Cognitive Psychologist*, 72, 27-53. doi: 10.1016/j.cogpsch.2014.01.004
- Karmiloff-Smith, A., Casey, B. J., Massand, E., Tomalski, P., y Thomas, M. S. C. (2014). Environmental and genetic influences on neurocognitive development: The importance of multiple methodologies and time-dependent intervention. *Clinical Psychological Science*, 2, 628-632. doi: 10.1177/2167702614521188
- Langer, J. (1986). The origins of logic: One to two years. New York: Academic Press.
- Langer, J. (1990). Early cognitive development: Basic functions. In C.A. Hauert (Ed.), *Developmental Psychology: Cognitive, perceptuo-motor and neuro-psychological perspectives* (pp. 19-42). Amsterdam: North Holland.
- Massand, E., y Karmiloff-Smith, A. (2015). Cascading genetic and environmental effects on development: implications for intervention. In K. Mitchell (Ed.), *The Genetics of Neurodevelopmental Disorders* (pp. 275-288). Hoboken, U.S.: Wiley-Blackwell.
- McArdle, J. J., y Nesselroade, J. R. (2003). Growth curve analysis in contemporary psychological research. In J. Schinka y W. Velicer (Eds.), *Comprehensive handbook of psychology: Research methods in psychology* (Vol. 2, pp. 447-480). Nueva York: Wiley.
- Miranda, A., Colomer, C., Mercader, J., Fernández, I., y Presentación, M. J. (2016). Performance-based tests versus behavioral ratings in the assessment of executive functioning in preschoolers: associations with ADHD symptoms and reading achievement. *Frontiers in Psychology*, 29. doi: 10.3389/fpsyg.2015.00545
- Mishra, J., Anguera, J. A., Ziegler, D. A., y Gazzaley, A. (2014). A Cognitive Framework for Understanding and Improving Interference Resolution in the Brain. *Progress in Brain Research*, 207, 351-377. doi: 10.1016/B978-0-444-63327-9.00013-8
- Mitchell, S. K. (1979). The interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, 86, 376-390.
- Miyake, A., y Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Current Directions in Psychological Science*, 21(1), 8-14. doi: 10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., y Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'Frontal Lobe' tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Moriguchi, Y., Chevalier, N., y Zelazo, P. D. (2016). Editorial: Development of Executive Functions during Childhood. *Frontiers in Psychology*, 7(6). doi: 10.3389/fpsyg.2016.00006
- Muldoon, K., Lewis, C., y Towse, J. (2005). Because it's there! Why some children count, rather than infer numerical relationships. *Cognitive Development*, 20, 472-491.
- Mushquash, C., y O'Connor, B. P. (2006). SPSS and SAS programs for generalizability theory analyses. *Behavior Research Methods*, 38, 542-557.
- Nee, D. E., y Jonides, J. (2008). Dissociable interference-control processes in perception and memory. *Psychological Science*, 19(5), 490-500. doi: 10.1111/j.1467-9280.2008.02114.x
- Nelson, C. A., y Luciana, M. (2008). *Handbook of Developmental Cognitive Neuroscience* (2nd Ed.). London: The MIT Press.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, 126(2), 220-246.
- Perea, A. E., Alday, L., y Castellano, J. (2006). Registro de datos observacionales a partir del MATCH VISION STUDIO v1.0. In J. Castellano, L. M. Sautu, A. Blanco-Villaseñor, A. Hernández Mendo, A. Goñi, y F. Martínez (Eds.), *Socialización y Deporte: Revisión crítica* (pp. 135-152). Vitoria-Gasteiz: Arabako Foru Aldundia-Diputación Foral de Álava.
- Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., y Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review*, 40, 25-71. doi: 10.1016/j.dr.2016.02.001
- Salley, B., Panneton, R. K., y Colombo, J. (2013). Separable Attentional Predictors of Language Outcome. *Infancy*, 18(4), 462-489.
- SAS Institute Inc. (2004). *SAS 9.1.3 Help and documentation*. Cary, NC: SAS Institute Inc.
- Scharf, R. J., Scharf, G. J., y Stroustrup, A. (2016). Developmental Milestones. *Pediatrics in Review*, 37(1), 25-37. doi:10.1542/pir.376266
- Schlotzhauer, S. D., y Littell, R. C. (1997). *SAS system for elementary statistical analysis*. Cary, NC: SAS Institute Inc.
- Shanmugan, S., y Satterthwaite, T. (2015). Neural markers of the development of executive function: relevance for education. *Behavioral Sciences*, 10, 7-13. doi: 10.1016/j.cobeha.2016.04.007
- Shaw, C., Brady, L. M., y Davey, C. (2011). *Guidelines for Research with Children and Young People*. London: National Children's Bureau Research Centre.
- Sinclair, H., Stambak, M., Lezine, I., Rayna, S., y Verba, M. (1984). *Los bebés y las cosas*. Barcelona: Gedisa.
- Stievano, P., y Valeri, G. (2013). Executive functions in early childhood: Interrelations and structural development of inhibition, set-shifting and working memory. *Neuropsychological Trends*, 13(1), 27-45. doi: 10.7358/neur-2013-013-stie
- Viterbori, P., Usai, M. C., Traverso, L., y De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: longitudinal study. *Journal Experimental of Child Psychology*, 140, 38-55. doi: 10.1016/j.jecp.2015.06.014
- Wass, S. V. (2015). Applying cognitive training to target executive functions during early development. *Child Neuropsychology*, 21(2), 150-166. doi: 10.1080/09297049.2014.882888
- Wertz, F. J. (2014). Qualitative Inquiry in the History of Psychology. *Qualitative Psychology*, 1, 4-16.
- Whitebread, D., y Coltman, P. (2010). Aspects of pedagogy supporting meta-cognition and self-regulation in mathematical learning of young children: evidence from an observational study. *ZDM Mathematics Education*, 42(2), 163-178. doi:10.1007/s11858-009-0233-1
- Willoughby, M. T., y Blair, C. B. (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment*, 28(3), 319-330. doi: 10.1037/pas0000152
- Willoughby, M. T., Holochwost, S. J., Blanton, Z. E., y Blair, C. B. (2014). Executive Functions: Formative versus Reflective measurement. *Measurement: Interdisciplinary Research and Perspectives*, 12(3), 69-95. doi: 10.1080/15366367.2014.929.453
- Willoughby, M. T., Wirth, R. J., y Blair, C. B. (2011). Contributions of modern measurement theory to measuring executive functions in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology*, 108, 414-435. doi: 10.1016/j.jecp.2010.04.007
- Willoughby, M. T., Wirth, R. J., Blair, C. B., y Family Life Project Investigators (2016). Executive Function in Early Childhood: Longitudinal Measurement Invariance and Developmental Change. *Psychological Assessment*, 28(3), 319-330. doi: 10.1037/pas0000152
- Zelazo, P. D., y Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, 6(4), 354-360. doi: 10.1111/j.1750-8606.2012.00246.x

(Artículo recibido: 12-10-2016; revisado: 29-11-2016; aceptado: 28-02-2017)