How do children begin to engage executive functions in self-directed contexts? Modeling environmental, genetic, and cognitive processes supporting semantic verbal fluency

by

Jane E. Barker

B.A., Rice University, 2004

M.P.A., Indiana University Bloomington, 2008

M.A., University of Colorado Boulder, 2013

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Psychology and Neuroscience

2017

This thesis entitled:
How do children begin to engage executive functions in self-directed contexts?
Modeling environmental, genetic, and cognitive processes supporting semantic verbal fluency
written by Jane E. Barker
has been approved for the Department of Psychology and Neuroscience

Prof. Yuko Munakata

Prof. Eliana Colunga

Prof. Naomi Friedman

Prof. Akira Miyake

Prof. Mike C. Mozer

Prof. Randall C. O'Reilly

Date ____________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Jane E. Barker, (Ph.D., Cognitive Neuroscience)

How do children begin to engage executive functions in self-directed contexts?

Modeling environmental, genetic, and cognitive processes supporting semantic verbal fluency

Thesis directed by Prof. Yuko Munakata

Young children often struggle to accomplish their intended goals in a self-directed way, without instructions or reminders from adults. Although it is clear that the ability to meet goals without external direction emerges slowly across development, little is known about the cognitive processes that might support these improvements, and whether certain experiences might be more effective in facilitating emerging self-direction than others. Chapters 2 and 3 in this dissertation explore relationships between children's time in adult-structured activities, where they have fewer opportunities to decide what they will do, and their performance on a measure of self-direction in a task where few reminders are given, semantic verbal fluency (VF). Chapter 2 shows that 6- and 7-year-old children who spend more time in less-structured activities show better self-directed switching performance in VF, relative to children who spend more time in adult-structured activities. Structured activities, including adult-led lessons, homework, and chores, showed a trend-level negative association in the opposite direction, such that more time in structured activities predicted worse switching performance. These observed relationships were specific to self-directed forms of EF, as children's time spent in structured and less-structured activities did not relate to their performance in two standard, externally-driven measures of executive function. Chapter 3 replicates and extends findings from Chapter 2 by investigating relationships between VF switching performance and two measures of environmental structure in a genetically-informative longitudinal twin sample. In independent phenotypic models, twins who lived in more structured homes and participated in more structured activities at ages 3 and 4 showed worse, and marginally worse VF switching performance at age 7, respectively, controlling for earlier VF ability and concurrent levels of environmental structure. At the same time, children who showed better VF performance at age

4 were more likely to participate in structured activities at age 7. These relationships persisted in models controlling for general cognitive ability, vocabulary knowledge, and socioeconomic status. Subsequent tests of the etiology of twin-level structured activity participation reveal that associations between early time use and self-directed switching ability were mediated by environmental rather than genetic factors. Whereas nonshared environmental factors mediated links between early structure and later VF (consistent with causal environmental mechanisms, and ruling out potential genetic confounds), shared environmental factors partially explained links between early VF and later structure, suggesting that the observed phenotypic relationship may be driven by a passive gene-environmental correlation. Additionally, links between environmental structure and later VF switching performance disappeared or reversed after age 7, consistent with findings from previous studies showing potential benefits of structured activity participation to older children, and potentially reflecting changes in relative levels of adult structure across time. Chapter 4 focuses on the cognitive processes underlying production in semantic VF via the development of a computational model demonstrating how experience-dependent abstract representations aid children's word production. Chapter 5 concludes with a discussion of limitations, open questions, and future directions.

Acknowledgements

I am deeply thankful for Yuko Munakata's continual support and guidance, and for the immensely helpful (and fun!) conversations we had with committee members, including Naomi Friedman, who patiently guided the behavioral genetics approach in Chapter 3; Eliana Colunga, Mike Mozer, and Randy OReilly, who greatly improved the modeling approach in Chapter 4; and Akira Miyake, who made critical theoretical and methodological suggestions that supported this work at each stage.

I would also like to acknowledge the multiple coauthors who contributed to the published work in this dissertation, including Laura Michaelson, Lindsay Provan, Hannah Snyder, and Andrei Semenov. Many RAs volunteered to transcribe records analyzed in Chapter 3 over a period of years; special thanks goes to Kaylyn Van Deusen for her enthusiastic dedication to this work. Thanks also to the members of the Cognitive Development Center, Dan Gustavson, Alex Hatoum, and Wolfgang Pauli, for their feedback on early findings and later drafts, and to researchers at the Colorado Longitudinal Twin Study for generously making their data available and supporting this work.

Note: Experiment 1 was previously reported in *Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in children's daily lives predicts self-directed executive functioning. Frontiers in psychology, 5.* The projects reported in Chapters 3 and 4 were supported by a grant from the CU Boulder Innovative Seed Grant Program.

Contents

Tables

Table

Figures

Figure

Chapter 1

General Introduction

In the absence of external reminders and instructions, young children often struggle to accomplish their intended goals. Lapses are common even when intentions are clear: a child might fail to get a permission form signed prior to an eagerly awaited field trip, for example, or not grab a coat before going outside to play on a chilly day. The ability to meet goals in a self-directed way, without instructions or reminders from adults, emerges slowly across development. However, little is known about the cognitive processes that might support these improvements, and whether certain experiences might be more effective in facilitating emerging self-direction than others.

To accomplish goal-directed tasks, children must engage executive functions (EFs), the cognitive control processes that regulate thought and action in support of goal-directed behavior. EFs develop dramatically during childhood (e.g. Gathercole, Pickering, Ambridge, & Wearing, 2004; McAuley, Christ, & White, 2011; Munakata, Snyder, & Chatham, 2012), and support a number of higher-level cognitive processes, including planning and decision-making, maintenance and manipulation of information in memory, inhibition of unwanted thoughts, feelings, and actions, and flexible shifting from one task to another (Miyake et al., 2000). EFs are early predictors of success across a range of important outcomes, including school readiness in preschoolers (e.g. M. R. Miller, Müller, Giesbrecht, Carpendale, & Kerns, 2013), and subsequent academic performance (Best, Miller, & Naglieri, 2011; Blair & Razza, 2007; Cameron et al., 2012; Samuels, Tournaki, Blackman, & Zilinski, 2016; St Clair-Thompson & Gathercole, 2006). Children with worse EF go on to have poorer health, wealth, and social outcomes in adulthood than children with better EF, even after controlling for differences in general intelligence (Moffitt et al., 2011).

Given established links between early EFs and later life outcomes, there has been substantial interest in determining whether EF abilities can be changed through experience. Most of this work has focused on adult-led training or interventions, which allow children to practice EFs in an environment where adults provide guidance and support. For example, children's visuospatial working memory can be improved through short periods of targeted training, though training effects fade quickly, and show little transfer to other tasks (Bergman Nutley et al., 2011; Holmes, Gathercole, & Dunning, 2009; Melby-Lervåg & Hulme, 2013). Broader interventions implemented as part of preschool curricula have been shown to improve children's cognitive flexibility (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Diamond, Barnett, Thomas, & Munro, 2007; Lillard & Else-Quest, 2006; Röthlisberger, Neuenschwander, Cimeli, Michel, & Roebers, 2012). Relative to children in business-as-usual classrooms, children enrolled in such programs have subsequently shown better performance in tasks where they must flexibly shift from one rule (e.g., sorting cards by their shape) to another (e.g., switching to sorting the cards by color).

Although training and intervention studies have yielded improvements in children's EF, most benefits have been observed in EF tasks where instructions and reminders are provided by an experimenter. It is unclear how intervention and training experiences might influence the development of self-directed forms of executive functioning, where children must determine on their own what goal-directed actions to carry out and when. This topic is of substantial interest, since such behaviors may be especially predictive of subsequent life outcomes (e.g. Moffitt et al., 2011).

Investigations of the processes supporting EFs in self-directed contexts across development could also help to explain why laboratory assessments of EF abilities, which are traditionally externally-driven, often show low correspondence with rating scales of executive function. Frequently-used ratings scales of executive dysfunction in everyday activities show only small to modest associations with performance-based indices of EF (Toplak, West, & Stanovich, 2013). This is troubling, as such scales were developed to provide ecologically-valid assessments of competence during complex, real-world decision-making (Toplak et al., 2013). Current explanations of differences across ratings and performance-based measures have focused on differences in the time-scales probed by each measure (e.g., EF in daily life, versus EF in a single laboratory session) and differences in measured constructs. An alternative, unexplored explanation is that laboratory-based EF measures focus on how efficiently individuals achieve task goals when they are given, whereas ratings of executive function assess the extent to which the individuals achieving their own goals in situations where instructions and reminders may be absent.

Few studies have explicitly tested how children begin to engage EFs in self-directed contexts, in part because few laboratory measures are designed to investigate goal-directed behavior in the absence of clear instructions and reminders. Dissociations between performance in self-directed and externally-driven EF tasks have been evaluated more frequently in clinical patients. A longstanding observation is that frontal patients may perform well on cognitive batteries administered in structured testing environments, and quite poorly in less-structured, out-of-lab settings (Burgess, Alderman, Volle, Benoit, & Gilbert, 2009; Eslinger & Damasio, 1985; Reitan & Wolfson, 1994; Shallice & Burgess, 1991; Wilson, 1993). Patients who struggle in self-directed tasks frequently demonstrate intact IQ and long-term memory, and perform like neurotypical controls on standard, externally-driven measures of executive function such as Stroop and card sorting tasks (Alderman, Burgess, Knight, & Henman, 2003; Knight, Alderman, & Burgess, 2002; Shallice & Burgess, 1991; Tranel, Hathaway-Nepple, & Anderson, 2007). Clinical descriptions of patients with seemingly selective deficits in real-world decision-making highlight how those individuals struggle to select one action when confronted with competing options. When selecting a restaurant, or determining what to wear, patients make endless comparisons and contrasts, often being completely unable to come to a decision at all (Burgess et al., 2009, p. 495).

One task that is relatively well-studied in both clinical and developmental populations is semantic verbal fluency (VF). In semantic VF, participants are given a single categorical prompt (e.g., animals or foods), and attempt to produce as many exemplars from that category as possible in a given time interval, without additional reminders from an experimenter (Bechtoldt, Benton, & Fogel, 1962). Self-directed behavior on VF tasks is typically assessed via analysis of participant clustering and switching behaviors. Participants who produce many responses tend to cluster their responses by subcategory (Troyer, Moscovitch, & Winocur, 1997; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998), and switch between clusters when novel response options dwindle. Thus, fluency measures indexing the total number of switches between recognizable clusters (e.g., from zoo animals to farm animals) are considered a good index of self-guided, internally-directed search processes (Abwender, Swan, Bowerman, & Connolly, 2001; Snyder & Munakata, 2010; Troyer et al., 1997).

Children continue to improve in VF switching and production well into adolescence (Ardila, Rosselli, Matute, & Guajardo, 2005; Kavé, Kigel, & Kochva, 2008; Koren, Kofman, & Berger, 2005; Matute, Rosselli, Ardila, & Morales, 2004; Riva, Nichelli, & Devoti, 2000; Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004; Welsh, 1991), well after they show adult-like performance on externally-driven EF tasks (Hurks et al., 2010; Kavé et al., 2008; Klenberg, Korkman, & Lahti-Nuuttila, 2001; Koren et al., 2005; Sauzéon et al., 2004; Welsh, 1991). Improvements in switching ability persist after children produce clusters that are indistinguishable in size from those of adults (Hurks et al., 2010). These findings collectively suggest that the ability to engage EFs in self-directed contexts may draw on distinct, slow-developing cognitive processes or abilities.

If training and intervention experiences can influence children's executive functions, children's ability to engage EFs in self-directed contexts may show similar malleability. My dissertation explores this question by examining links between the amount of structure in children's daily environments and their developing self-directed switching in semantic verbal fluency. Chapter 2 tests for contemporaneous relationships between 7-year-old children's time use, as indexed by detailed daily and annual schedules, and their performance on semantic VF and other, externally-driven measures of EF. Chapter 3 builds on these initial findings by testing links between semantic VF, child structured activity participation, and household structure in a longitudinal, genetically-informative sample. Chapter 4 focuses on the cognitive processes hypothesized to support self-directed switching in semantic VF, via development of a novel computational framework.

1.1 Dissertation summary

Chapter 2 investigates whether 6- and 7-year-old children who spend more time in less-structured activities, where adults play a less-prominent role in determining what children will do, show better switching performance in VF, relative to children who spend more time in adult-structured activities. Information about children's activities was collected from parent reports of children's daily, annual, and typical schedules. Activities were then categorized as "structured" or "less-structured" based on categorization schemes from prior studies on child leisure time use (Eccles & Barber, 1999; Fletcher, Nickerson, & Wright, 2003; Hofferth & Sandberg, 2001b; Mahoney & Stattin, 2000; Meeks & Mauldin, 1990; Osgood, Anderson, & Shaffer, 2005). The more time that children spent in less-structured activities, including play, social outings, and enrichment, the better their switching performance on a self-directed verbal fluency task, controlling for differences in vocabulary, income, and age. Structured activities, including adult-led lessons, homework, and chores, showed a trend-level negative association in the opposite direction, such that more time in structured activities predicted worse switching performance. The observed relationships with time use were specific to self-directed forms of EF, as children's time spent in structured and less-structured activities did not relate to their performance two standard, externally-driven measures of executive function.

Chapter 3 builds on Chapter 2 by investigating relationships between children's exposure to adult-structured activities and their emerging ability to engage self-directed EFs within a broader, genetically informative twin sample. We find that twins who participated in more structured activities and lived in more structured households at age 4 showed marginally worse switching performance in VF at age 7, controlling for their age-4 ability and contemporaneous structured activity participation. Over the same period, children's early verbal fluency ability was positively associated with their subsequent structured activity participation. These opposing influences persisted after controlling for potential confounding influences, including general cognitive ability, vocabulary, and socioeconomic status. Follow-up genetic analyses suggest that observed temporal relationships were entirely driven by environmental, rather than genetic influences, ruling out genetic mediation as the source of observed relationships. Additionally, after age 7, relationships between environmental structure and VF switching ability trended in the opposite direction, consistent with a large body of work linking increased structured leisure time to positive academic and social outcomes in early adolescence. Finally, follow-up analyses reveal that measures of environmental structure show similar relationships with average cluster sizes in semantic VF, raising the possibility that observed relationships between environmental structure and VF may not be specific to self-directed switching. However, it is difficult to make clean distinctions across these highly-correlated processes, troubling interpretation. Taken together, these findings highlight the complexity of relationships between time use and child outcomes across development, and suggest that alternative mechanisms may play a role in observed links between environmental structure and developing EF, including the possibility that structured activities influence how children acquire semantic representations supporting production in VF.

Chapter 4 develops a novel computational framework modeling the cognitive processes underlying production in semantic verbal fluency, and explores whether recent models of VF task performance, which have posited a primary role for associative, non-executive processes, can be improved by incorporating selection processes that are guided by abstract representations (i.e., higher-order categorical representations, including subcategories like 'zoo animals' or 'pets' in response to an 'animal' prompt). Extant attempts to model such processes have relied on an incomplete set of unvalidated subcategories that likely do not reflect the full range of conceptual dimensions that individuals use to generate task-relevant responses; as such, it is perhaps unsurprising that they have been outperformed by associative models. To critically evaluate the role of abstract representations in guiding search during VF, models must incorporate valid representations. Therefore, this chapter explores the feasibility of developing experience-driven abstract representations using real-world child texts, and provides an initial model of how those representations may support word search during VF, building on theories positing that word production is aided by maintenance and selection of higher-order categorical representations that facilitate selection of lower-level exemplars (Hirshorn & Thompson-Schill, 2006; Troyer et al., 1997). Via a series of permutation tests informed by patterns of child production in VF, a simple associative model was compared to a model incorporating information from learned topical representations. Initial tests suggest the model incorporating topical representations outperforms a purely associative model, challenging previous accounts.

Chapter 5 describes patterns that emerge across studies, discusses limitations and open questions, and outlines future directions. Specifically, I consider experimental paradigms for testing causal directions between environmental structure and verbal fluency performance, consider experimental and computational methods for evaluating cognitive demands in semantic VF, and discuss methods for developing new measures of children's ability to engage EFs in self-directed contexts.

Chapter 2

Relationships between children's daily activities and their self-directed executive function

Note: This chapter is closely adapted from Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in childrens daily lives predicts self-directed executive functioning. *Frontiers in Psychology*, 5.

As a first step in examining the question of how children's experiences outside of formal schooling relate to executive functions, we conducted a naturalistic, correlational study, in which we measured the time that 6-year-old children spent in their daily lives in structured and less-structured activities and tested whether it predicted performance in the lab on well-established executive function tasks, both externally-driven and self-directed. At this age, children spend some time in both structured and less-structured activities (e.g. Hofferth & Sandberg, 2001b; Meeks & Mauldin, 1990) and show some ability in self-directed control tasks, without showing high levels of proficiency (e.g. Brocki & Bohlin, 2004; Kavé et al., 2008; Snyder & Munakata, 2010, 2013; Welsh, 1991).

We predicted that children's self-directed EFs might benefit from participation in less structured activities, where children, rather than adults, choose what they will do and when. Such experiences could support the practice of self-directed executive functioning, and lead to benefits. For example, children may practice engaging self-directed forms of EF by establishing goals and carrying them out across an afternoon ('first I'll read this book, then I'll make a drawing about the book, then I'll show everyone my drawing') or during a visit to a museum ('first I want to see the dinosaur exhibit, and then I want to learn about rocks').

To classify structured and less-structured activities, we relied on studies of child leisure time use (e.g. Fletcher et al., 2003; Hofferth & Sandberg, 2001b; Larson & Verma, 1999; Meeks & Mauldin, 1990; Osgood et al., 2005) which have attempted to discriminate between activities constituting structured, or constructive leisure, and "unstructured" leisure activities. "Unstructured" activities in this literature might be better thought of as "less-structured" activities, given that they can include some adult structuring, so we use the latter terminology throughout this paper. Most leisure time studies have identified structured leisure activities as those "supervised to some degree by a conventional adult, are highly structured, and provide [children] with a clear set of conventional activities in which to engage" (Agnew & Petersen, 1989, p. 335). Such activities "are [. . .] organized by adults around specific social or behavioral goals" (Fletcher et al., 2003, p. 641). Thus, structured time in the present study was defined to include any time outside of formal schooling.[1] spent in activities organized and supervised by adults (e.g., piano lessons, organized soccer practice, community service, homework). Less-structured activities have been described more loosely, and generally include voluntary leisure activities where adults provide fewer guidelines or direct instructions (e.g., activities that are "spontaneous, taking place without formal rules or direction from adult leaders, and featuring few goals related to skill development" (Mahoney & Stattin, 2000, p. 116). Our coding scheme follows existing coding schemes documented in Meeks and Mauldin (1990) and Hofferth and Sandberg (2001b). In cases where these coding schemes differed, we reviewed the literature to ensure that our coding was in accordance with the majority of other time use studies.[2] In the present study, less-structured activities included activities such as free play, family and social events, reading, drawing, and media time. While these classifications are imperfect (e.g., they do not capture the degree of structure within and across classifications - an

[1] We did not classify time spent in school as 'structured' because the degree of structure in school settings can vary a great deal, and parent reports are likely to be inaccurate (since parents often do not have direct knowledge of child activities during schooling hours) Our delineation of structured activities is also consistent with past studies of structured leisure time, which have excluded time spent in school (e.g. Meeks & Mauldin, 1990; Larson & Verma, 1999; Hofferth & Sandberg, 2001b; Mahoney & Stattin, 2000; Fletcher et al., 2003; Osgood et al., 2005)

[2] Hofferth and Sandberg (2001b) separately identify reading, studying, and television watching as learning activities. However, we have classified reading and television as less-structured time, and studying as structured time, in keeping with other studies (Meeks & Mauldin, 1990; Eccles & Barber, 1999; Fletcher et al., 2003).

issue we return to in the Discussion (Section 2.3)), they allow us to build on the existing literature, and serve as an important starting point for testing our predictions; further analyses allow us to test the importance of particular activities within these classifications.

We hypothesized that the amount of time children spent in less-structured activities would predict their self-directed EF, over and above any differences attributable to age, general vocabulary knowledge, and household income. We expected these effects to be specific, such that less-structured activities would not predict externally-driven EF and structured activities would not predict self-directed EF.

2.1 Method

2.1.1 Participants

Seventy children participated in the study (M_{age}= 6.58 years; range= [6.01– 7.00 years]; males = 37). All participants were recruited from a database of families who had volunteered to participate in research. During subject recruitment, parents were informed that they would be asked to document child activities during the week prior to the study visit. Three participants were excluded from analyses because detailed information on their weekly activities was unavailable, either because parents did not wish to provide this information (2), or because data were lost due to a technical error at the time of parent submission (1). Of the remaining participants, one child did not complete the Flanker task, one child did not complete the digit span task, and two children did not complete the verbal fluency task; each of these children was excluded from the analysis of only that task. All other participants completed all study tasks. Prior to their participation, parents gave informed consent, and children gave verbal assent. Children received small gifts (e.g., gliders, balls) throughout the project for their participation, and parents received $5 as compensation for travel.

2.1.2 Design and Procedure

Children were individually tested in a single session lasting approximately 1.5 hours, with breaks given as needed. All children completed tasks in the same order: AX-CPT, Flanker, forward digit span (for other purposes, not discussed further in this report,[3]) verbal fluency, and the Expressive Vocabulary Test. During the child tasks, parents provided demographic information and completed surveys of children's daily, annual, and typical schedules, as well as an exploratory 'helicopter parenting' scale (not discussed further in this report; from J. Obradovic; personal communication, October 26, 2011).

2.1.3 Parent Questionnaires

Parent Survey of Child Time Use. Parents reported all child activities during the week prior to the laboratory test session using a computer-based survey. At the time that the study visit was scheduled, parents were informed that they would complete a detailed child activity survey during their visit, and were encouraged to take notes on their child's activities throughout the week. Parents were allowed to consult notes as they completed the survey. The survey was formatted as a 36 x 7 grid, such that each cell represented a 30-minute time interval during the prior week (intervals occurring between 12:00 AM and 5:30 AM were excluded to reduce burden). In each cell, parents wrote short, open-ended description of their child's activities, excluding times where children were sleeping or in school (parents indicated sleep and school schedules in a separate section of the survey). Before completing the survey, parents were asked to indicate the extent to which their family's activities over the prior week reflected typical patterns of time use. Parents rated their level of agreement with the prompt, "Was your family's schedule last week unusual or atypical?" via a 7-point scale anchored by 'Strongly agree' and 'Strongly disagree'. Parents were then given verbal and written instructions, as follows:

[3] Forward digit span tasks (where children repeat numbers in the order they are presented by an experimenter) primarily index storage capacity, rather than combined storage and processing capacity, and therefore do not serve as a reliable measure of EF (Daneman & Merikle, 1996; Engle, Tuholski, Laughlin, & Conway, 1999).

"Be as specific as possible for every activity you report. For example, for time spent in the car during a commute, rather than writing, "Drove from ___ to ___," you could write, "Watched a DVD with his sister in the car while driving to the city for a research appointment."

Indicate who your child was interacting with during a given activity. For example, if your child had free time to play outside between dinner and bedtime, rather than writing "Free time outside", you could write, "Played tag outside with older sister and friends from the neighborhood." Or, if your child reads before bedtime, rather than writing, "Reading time," you could write, "Read aloud to mom before bed."

Indicate simultaneous activities. For example, if your child ate a snack after school or camp while he/she had some down time, rather than writing "Snack time," you could write, "Ate a snack while coloring."

,

As parents completed the survey, experimenters periodically reviewed responses and asked that parents modify entries that were difficult to interpret or insufficiently detailed. Experimenters were also available during breaks between tasks to respond to parent questions about specific responses.

Child activity data were coded by three independent raters who were blind to data on all other tasks during each stage of the coding process. Coders assigned a numeric code to each cell-based survey entry using an activity classification scheme (Table 2.1). To ensure consistency across raters and reduce procedural drift, all raters independently classified each cell for the first 35 participants. Then, coders met to discuss major discrepancies and to generate additional generalizable rules. Coders categorized responses from the final 32 participants using these agreed-upon criteria. The final 32 subjects were used to establish inter-rater reliability; reliabilities among pairs of coders ranged from .96 to .97, with coders agreeing on 7,942 to 8,021 cells out of 8,288 total (i.e., 2 cells per hour x 18.5 hours/day x 7 days a week x 32 participants). Excluding sleep and school cells (where there were no discrepancies between coders), reliabilities among pairs of coders were also high, ranging from .93 to .95. The three coders met to discuss discrepancies and generate a final, coded data set for each participant.

After the raters generated the final set of activity codes, each activity was further classified as either 'Structured' or 'Less-Structured' based on the coding scheme outlined in Table 2.1, following existing coding schemes (Meeks & Mauldin, 1990; Hofferth & Sandberg, 2001b; Eccles & Barber, 1999; Mahoney & Stattin, 2000; Fletcher et al., 2003; Osgood et al., 2005). All child-initiated activities (play, spontaneous practice, reading, watching television) and outings and events (museum or library visits, sporting events) were coded as 'Less-Structured'. Adult-led lessons and practices, homework and studying, religious activities, and organization meetings (e.g., community service) were coded as 'Structured'.

Parent Survey of Typical Child Time Spent in Less-structured Activities. In a separate survey, parents were asked to indicate how often their children engaged in typical play activities by using a 7-point scale ('Never', 'Less than once a month', 'Once a month', '2-3 times a month', 'Once a week', '2-3 times a week', 'Daily') to rate the following items: 'Surf the internet', 'Watch television, videos/DVD, or online media', 'Play video games (non-instructional)', 'Play interactive instructional or learning games', 'Play with toys alone', 'Play with toys with friends/siblings', 'Play physical games with friends/siblings', 'Play physical games alone', 'Play non-physical games alone', 'Play card or board games with family', 'Read', 'Help with housework or cooking', 'Play musical instrument', 'Listen to music.' Scores on each item (where 1= 'Never' and 7 = 'Daily') were summed to produce a typical less-structured activity score.

Parent Survey of Seasonal Child Activities. In a separate survey, parents were asked to indicate the number of hours their child spent in structured lessons during the past year. Parents responded to 18 common structured lessons (basketball, baseball, tennis, hockey, soccer, football, golf, swimming, dance, gymnastics, martial arts, skiing/snowboarding, ice skating, music, art, theater, and tutoring) and were asked to write in any structured lessons that did not fall into these categories (most commonly, religious activities and organizational meetings). To reduce burden,

Table 2.1: Classification of Child Time Use (Structured, Less-structured, and Other Activities).

Structured Activities

- Physical lessons (e.g. soccer practice, karate)
- Non-physical lessons (e.g. piano lessons, art class)
- Tutoring
- Homework and study
- Chores
- Religious activities
- Other formal organizational meetings and activities (e.g., community service)

Less-Structured Activities

- Unguided, child-initiated practice (e.g., playing piano or singing outside of scheduled practice times; shooting goals outside of soccer practice)
- Free play alone
- Free play with others
- Social outings
- Visits to family and friends
- Parties
- Camping
- Picnics
- Other group activities (e.g., walks, bike rides, skiing, swimming, bowling, golf)
- Enrichment activities
    - Sightseeing
    - Aquarium and zoo visits
    - Museums
    - Miscellaneous educational events (e.g., science fair)
- Other entertainment (e.g., live sporting events, performances, movies)
- Reading
- Media and screen time (e.g., TV, internet, video games)

Other Activities

- Sleeping
- Meals/eating
- School
- Care by others
- Personal care and hygiene item Child appointments
- Commuting and travel time
- Unknown/Unreported

Note: All entries that parents provided in the child time use survey were classified into these categories, following existing coding schemes (Eccles & Barber, 1999; Fletcher et al., 2003; Hofferth & Sandberg, 2001a; Mahoney & Stattin, 2000; Meeks & Mauldin, 1990; Osgood et al., 2005) parents provided seasonal time estimations for each activity (e.g., the typical hours per week a child spent participating in music lessons over the prior fall). Data were reviewed for accuracy to ensure that parent-reported structured activities adhered to the same coding guidelines used to evaluate the Parent Survey of Weekly Activities. Cumulative hours spent in structured activities across the year were summed to produce an annual structured hours score.

Household Income. Parents reported annual household income via an interval scale (median bracket: $100,000 - $124,999; range: from < $25,000 to > $150,000 USD). Fourteen parents chose not to disclose income information.

2.1.4 Child Endogenous Executive Function Measure

Verbal Fluency . In the verbal fluency task, children were asked to generate words in response to a categorical prompt. The task was presented as a game to make it more engaging for children (as in Snyder & Munakata, 2013). Children were told, "We're going to play a game where we think of lots and lots of words. I bet you're really good at thinking of words, aren't you? I'll tell you what kinds of words to think of, and every time you tell me one, I'll put a pom-pom in your cup. Let's see how many pom-poms you can get before all the sand is gone [experimenter pointed to a 1-minute sand timer children could use to estimate how much time was left]. I'll bet you can get a lot! And when we are all done thinking of words, you can trade the pom-poms for a prize." Before each category, the experimenter said, "This time I want you to tell me as many [category name] as you can think of. Can you think of lots and lots of [category name]? Ready, go!" The experimenter placed a pom-pom in a clear plastic cup in front of the child for each new exemplar. If children paused for 10 s or longer between items, they were encouraged to continue ("Good job, can you tell me some more [category name]?"). In the rare instance where a child stated that she/he had named all words, the experimenter double-checked with the child (e.g., "Are you sure? What other [category name] can you think of?") and waited with the child until the end of the block. Children completed three blocks using this procedure, each of 1-minute duration: a practice block (with the prompt 'household items'), and two test blocks (with the prompts 'animals' and 'foods', which were counterbalanced across participants).

Verbal fluency data were transcribed from audio recordings, and coded by the experimenter and two independent raters blind to data on all other tasks. Coders identified clusters of items that were semantically related (e.g., "cookies, pie, cake" when producing foods). Switches between clusters of related items were identified and summed to generate cumulative switch scores. Switch scores were weighted by cluster size (as in Snyder & Munakata, 2010, 2013), such that 1 point was awarded for a switch after a cluster of 2 related items, 2 points for a switch after 3 related items, 3 points for switch after 4 related items, and so on. Weighted switch scores were used because they reflect increasing confidence as cluster size increases that children are indeed clustering and switching. Unweighted scoring systems (e.g. Troyer et al., 1997), which count every transition between subcategories equally (including between single, unclustered items), have been criticized for confounding switching with a failure to cluster (e.g. Abwender et al., 2001). Inter-rater reliabilities were high between all pairs ($> 85\%$). To generate cumulative switch scores for each participant, weighted switch scores were averaged across coders within each prompt, and then summed.

2.1.5 Child Externally-driven Executive Function Measures

Flanker. Children completed a computerized flanker task (Eriksen & Schultz, 1979) assessing their ability to resolve conflicting visual information by appropriately responding to a central stimulus while ignoring flanking stimuli. The Flanker task is a commonly-used measure of externally-directed EF in 6-year-olds (McDermott, Pérez-Edgar, & Fox, 2007; Ridderinkhof & van der Molen, 1995; Röthlisberger et al., 2012; M. R. Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; M. Rueda et al., 2004) and has been shown to be sensitive to some interventions targeting EF in this age group (Fisher et al., 2011; Röthlisberger et al., 2012). During the task, children were instructed to indicate the orientation (left or right pointing) of a centrally-presented target stimulus, via a corresponding button press. In congruent trials, the target stimulus (the center fish) was surrounded by fish with the same orientation. In incongruent trials, the target image was surrounded by fish with an opposite orientation. In neutral trials, only the target image was presented and was not surrounded by any fish. Following a 10-trial practice block (4 congruent, 4 incongruent, 2 neutral), children completed three 32-trial blocks of the task: two incongruent blocks (for each block, incongruent trial N = 16; neutral trial N = 16), separated by one congruent block (congruent trial N = 16; neutral trial N = 16). Trials were presented in random order within blocks.

Reaction times were used to assess children's ability to resolve interference among conflicting stimuli, as in past work with this age group (e.g. M. R. Rueda et al., 2005; McDermott et al., 2007; Röthlisberger et al., 2012). Incongruent trials require children to attend to only the target middle fish and to ignore the surrounding fish. Therefore, the flanker task can be used to assess children's ability to filter out irrelevant information. Larger interference costs (i.e., the difference between average response time on incongruent trials and average response time on neutral trials) reflect greater difficulty filtering irrelevant information. To assess filtering ability, we first calculated participant mean response times for each trial type (neutral, incongruent/congruent) within each block across trimmed, correct trials (trials < 100 ms and > 3000 ms were excluded, as well as any trials three standard deviations outside that participant's mean for that trial type and block). To generate robust estimates of possible interference effects (as suggested by Lavie (1995) and implemented in (D'Ostilio & Garraux, 2012)) incongruent/congruent trial mean RTs were contrasted with neutral trial mean RTs from the same block, yielding one congruent-neutral contrast and two incongruent-neutral contrasts within each participant. Flanker conflict scores were generated by subtracting the congruent contrast from each incongruent contrast (yielding two conflict scores, one arising from each incongruent block). These conflict scores were averaged to generate a summary flanker conflict score.

AX-CPT. Children completed the AX Continuous Performance Task (AX-CPT, which provides a measure of proactive control, or the tendency to maintain goal-relevant information until it is needed (Braver, Gray, & Burgess, 2007). All procedures and analyses were conducted as in (Chatham, Frank, & Munakata, 2009). In this touchscreen-based, child-friendly version, children are allowed to prepare for future circumstances (the appearance of either "X" or "Y" image probes) based on previous experiences (the appearance of "A" or "B" image cues).

Children were instructed to respond with a target response whenever the "A" context cue was followed by an "X" probe. Children were instructed to provide a non-target response to all other cue-probe sequences (A-Y; B-X; B-Y). To improve child engagement during the task, popular cartoon characters were used as image stimuli, and the instructions took the form of character preferences. For example, children were told, "Spongebob likes watermelon, so press the happy face when you see Spongebob and then the watermelon," and, "Blue doesn't like the slinky, so press the sad face when you see Blue and then the slinky."

After the experimenter explained the task rules, children completed a "verification" phase to ensure that they understood the instructions and were capable of following rules. During this phase, each cue–probe pair was presented sequentially, and participants were asked to indicate the correct response for each pair. If subjects responded incorrectly to a cue-probe pair, the experimenter repeated the relevant rule ("Remember, when you see [A, B] and then you see [X, Y], tap this button [appropriate button blinks] as quickly as you can!") and subjects were allowed to try again. Participants then completed 7 practice trials. Cues were presented for 500 ms, followed by a 120 ms delay period, and a subsequent 6 s probe, as in test trials. Test trials were presented in four 30-trial blocks, where 70% of trials were target (A-X) trials, and 30% were non-target trials (A-Y; B-X; B-Y, appearing in equal proportion).

Proactive children show a characteristic behavioral profile that can be used to generate an RT-based measure of proactive control. Children who engage proactive control generate fast RTs in BX and BY trials, since maintenance of the "B" cue supports a non-target response to the subsequent "X" probe, and slower RTs on AY trials, since active maintenance of the "A" cue leads to anticipation of an 'X' probe (due to the expectancy generated by asymmetric trial type frequencies). Proactive control was thus calculated using the median of trimmed RTs on correct AY and BX trials, which were entered into the formula (AY-BX) / (AY+BX). All responses made < 200 ms after the presentation of the probe were removed from the analysis, resulting in the exclusion of $< 1\%$ of all trials.

Expressive Vocabulary Test. The EVT –(Pearson Assessments, Bloomington, MN) is a standardized, nationally normed, expressive vocabulary test, which we used (as in Snyder & Munakata, 2010) to control for differences in vocabulary that might have influenced verbal fluency performance (i.e., a child with a robust vocabulary might be capable of generating larger clusters than a child with a limited vocabulary, independent of either child's switching ability). On each trial of the EVT, children are shown a colored picture and are asked to name it or provide a synonym (e.g., "Can you tell me another word for father?"). Testing continues until children incorrectly answer five items in a row, and raw scores are then converted into a standardized score based on age.

2.2 Results

2.2.1 Preliminary Results and Analysis Approach

All analyses were conducted using standard linear regression. We included age, gender, and family income as factors in all models, given that they or related factors are often predictive of children's EF: age (e.g. Huizinga, Dolan, & van der Molen, 2006; Welsh, 1991), gender (e.g. Blair, Granger, & Peters Razza, 2005; Diamond et al., 2007), family income (Hughes, Ensor, Wilson, and Graham (2009); as a component of SES: (Farah et al., 2006; Noble, McCandliss, & Farah, 2007; Noble, Norman, & Farah, 2005; Raver, Blair, Willoughby, & Family Life Project Key Investigators, 2013). Child vocabulary, as indexed by EVT performance, was included as a covariate in all tests of verbal fluency performance. Descriptive statistics for executive function, vocabulary, and time use measures are given in Table 2.2. Individual EF measures were not correlated, before or after controlling for age (p's $> .4$). For all analyses, outlying observations were identified (Cook's D > 3 standard deviations above the mean) and removed. This resulted in the exclusion of no more than four cases from any analysis.

Table 2.2: Descriptive Statistics for Executive Function, Vocabulary, and Time Use Measures.

Measure	**Mean (SD)**
Self-directed EF	
Verbal fluency combined switch score	10.13 (4.1)
Externally-driven EF	
AX-CPT proactive control score	0.094 (.12)
Flanker conflict score	164.5 (168.7)
Vocabulary: EVT standardized score	112.9 (9.4)
Prior week child time use	
Structured hours	6.03 (5.9)
Less-structured hours	32.2 (14.2)
Typical child less-structured activities (combined score)	78.5 (8.8)
Seasonal child structured activities (annual hours)	91.5 (89.0)

2.2.1.1 Child Time Use and Self-Directed EF

Less-structured Time. As predicted, children who spent more time in less-structured activities demonstrated better self-directed EF, as indexed by verbal fluency performance (η_p^2= .07; F(1,44) = 4.46; $p < .05$; Table 2.3). In addition, older children and children with higher vocabulary scores demonstrated better verbal fluency performance (Age: η_p^2= .11; F(1,44) = 7.45; $p < .01$; EVT: η_p^2= .10; F(1,44) = 6.30; $p < .02$). In subsequent tests for interactions, we found an unexpected interaction between less-structured time and age (Less-structured time x Age: η_p^2= .08; F(1,43) = 5.48; $p < .03$). Post-hoc tests indicated that additional time in less-structured activities predicted better self-directed control in most but not all children, specifically, this finding held in both the youngest sample quartile (M_{age} = 6.38 yrs, Less-structured time: η_p^2= .07; F(1,43) = 10.37; $p < .003$) and at the median (M_{age}= 6.65 yrs, Less-structured time: η_p^2= .07; F(1,43) = 6.81; $p < .02$), but not in the oldest quartile (M_{age}= 6.86 yrs; $p > .8$). When the interaction between less-structured time and age was included in the model, children from higher-income households demonstrated marginally better verbal fluency performance (Income: η_p^2= .05; (F(1,43) = 3.36; $p < .08$)). Age, vocabulary, and time in less-structured activities also continued to predict self-directed EF (Age: η_p^2= .12; F(1,43) = 5.76; $p < .03$; Vocabulary: η_p^2=.07; F(1,43) = 4.80; $p < .04$; Less-structured time: η_p^2= .07; F(1,43) = 6.81; $p < .02$).

Exploratory Analyses. We next investigated whether specific kinds of less-structured activities were driving the observed relationship between less-structured time and self-directed control. Composite variables representing common less-structured activities were created by aggregating similar responses across prior-week and annual measures.[4] This procedure yielded seven broad categories of less-structured activities: unguided practice; play alone; play with others; social events with family (including parties, camping, picnics, and other group outings, such as hiking, biking, and swimming, [5] enrichment events (visits to the museum, library, aquarium, or zoo; sightseeing;

[4] Aggregate within-measure scores were z-scored, then summed to create cross-measure composites.

[5] Social and enrichment events included only prior-week reporting, as these were not adequately identified in the annual less-structured time measure, which included only general activities (e.g., playing outdoor with friends) that could occur in many contexts.

and miscellaneous educational events), other entertainment (movies, performances, and live sporting events); reading; and media and screen time. Enrichment activities (η_p^2 = .11; F(1,44)=6.95; p < .02) and social events (η_p^2 = .10; F(1,43)=7.26; p < .01) significantly predicted self-directed EF, and play with others was marginally predictive (η_p^2 = .05; F(1,44)=3.42; p<.072). Interactions with age were not significant in these models, and were therefore excluded (p's > .2). No other classes of less-structured activities predicted verbal fluency performance.

We then considered whether the relationship between less-structured time and self-directed EF persisted when we excluded from our less-structured time composite measure, in sequential analyses, media and screen time (which might reflect passive, rather than self-directed leisure activity), activities within the less-structured time classification that may have included more structure than other such activities, and enrichment activities that may have yielded benefits specific to verbal fluency performance (rather than self-directed control, per se). When media and screen time were excluded, less-structured time continued to demonstrate a positive relationship with self-directed EF (η_p^2 = .06; F(1,41)=5.23; p < .03). This finding persisted when we also excluded less-structured activities that may have included more structure than other such activities (e.g., board games played with a group; rule-based physical games such as golf and bowling; movies and performances; reading with others [6]) (η_p^2 = .06; F(1,43)=6.17; p < .02). As a final step, we also excluded visits to museums, aquariums, and zoos, which may have benefited organization of semantic clusters on the verbal fluency task (e.g., exposure to zoo animals may have helped to organize animal clusters, and thus yielded performance benefits). Using this fully-restricted measure of less-structured time, child time in less-structured activities continued to predict better self-directed EF (η_p^2 = .06; F(1,43)=6.23; p < .02). Interactions with age were significant and were included in each of these restricted analyses (all p's < .05).

We also explored whether participation in types of less-structured activities changed with age, and whether such changing patterns of time use could speak to the diminished link between

[6] Here and in the following analysis, we also excluded all reading from our typical-activities measure, because this measure did not discriminate between reading alone and reading with others.

less-structured time and self-directed control in the oldest quartile of children in our sample. Media and screen time use was more prevalent in older children (η_p^2 = .05; F(1,61)=5.15; p < .03). Time spent in other categories of less-structured activities did not vary with age (p's > .2).

Structured Time. Additional time in structured activities predicted marginally worse self-directed control (η_p^2= .06; F(1,43)= 3.57; p < .07; Figure 2B; Table 2.3). Again, self-directed EF was predicted by age (η_p^2= .13; F(1,43) = 4.43; p < .01), and vocabulary (η_p^2= .08; F(1,43) = 5.02; p < .04), and marginally predicted by household income (η_p^2= .05; F(1,43)=3.50; p < .07). [7]

Exploratory Analyses. We next examined whether the relationship between structured time and self-directed EF persisted when we excluded religious services and household chores, where children may have been supervised less often by adults, relative to other structured activities. Time in structured activities continued to predict worse self-directed EF when religious services and chores were excluded from the composite structured time measure (η_p^2 = .06; F(1,43)= 4.28; p < .05).

2.2.1.2 Child Time Use and Externally-driven EF

No measure of child time predicted any aspect of externally-driven EF (Figures 3A-3D). Specifically, child time spent in less-structured activities did not relate to performance on either externally-driven EF measure (Flanker conflict score: p > .2; AX-CPT proactive control score: p > .8). Similarly, time in structured activities was unrelated to externally-driven EF (Flanker conflict score: p > .6; AX-CPT proactive control score: p > .3). [8] Males demonstrated better Flanker conflict scores than females (η_p^2 = .10; F(1, 46) = 4.64; p < .04). No other variables predicted externally-directed EF.[9]

[7] This finding was not driven by a negative correlation between composite time in structured activities and time in less-structured activities. The less-structured and structured time composites were not significantly related (p > .8).

[8] Although it is not a targeted measure of conflict resolution, overall accuracy across all trials on the Flanker task has also been tested in prior intervention work with children (M. R. Rueda et al., 2005; Fisher et al., 2011; Röthlisberger et al., 2012), and is what improved in two prior intervention studies targeting EF in this age group (Fisher et al., 2011; Röthlisberger et al., 2012). Overall accuracy did improve with age in our sample (η_p^2 = .16; F(1, 47) = 4.67; p < .04), but was not predicted by any other variables (ps > .15).

[9] In separate analyses, we investigated whether the completeness of parent reporting of child time influenced observed relationships between child time use and EFs. For example, if parents who left fewer cells blank in the time use survey had children with higher self-directed EF, this could have contributed to the observed correlation between less-structured time and self-directed EF, since parents who left fewer cells blank might report more time

in less-structured activities. However, completeness of reported time use did not affect the results: it showed no relationship with any aspect of EF performance (verbal fluency, AX-CPT, and Flanker ps $> .3$), and controlling for it did not change whether or not any findings were significant.

Table 2.3: Models relating Child Verbal Fluency Performance to their Fender, Household Income, Vocabulary, and Time Use.

	Age, Income, & Gender			**EVT**, Age, Income, & Gender			**Less-structured Time** + Age, Income, Gender, & EVT			Less-structured Time x Age + Income, Gender, & EVT			**Structured Time** + Age, Income, Gender, & EVT		
Variable	β	t^b	p	β	t^b	p	β	t^b	p	β	t^b	p	β	t^b	p
(Intercept)	9.99	22.48	<.001⋆	9.95	22.13	<.001⋆	9.94	22.67	<.001⋆	10.09	23.86	<.001⋆	9.73	21.97	<.001⋆
Age (days)	.008	1.79	<.09	.013	2.95	<.01‡	.011	2.73	<.01‡	.010	2.40	<.05†	.012	2.90	<.01†
Gender (1 = female; -1 = male)	.008	0.02	>.9	-.194	-0.42	>.6	-.392	-0.82	>.4	-.375	-0.82	>.4	-.169	-0.36	>.7
Household income	.743	3.17	<.01‡	.301	1.13	>.2	.372	1.48	>.1	.442	1.83	<.08	.487	1.87	<.07
Vocabulary (EVT)	-	-	-	.177	2.78	<.01‡	.142	2.51	<.05†	.120	2.19	<.05†	.128	2.24	<.05†
Less-structured Time	-	-	-	-	-	-	.713	2.11	<.05†	.854	2.61	<.05†	-	-	-
Less-structured Time x Age	-	-	-	-	-	-	-	-	-	-.008	-2.34	<.05†	-	-	-
Structured Time	-	-	-	-	-	-	-	-	-	-	-	-	-.596	-1.89	<.07
Model F-Value	4.05			4.88			4.56			5.10			4.34		
Model Adjusted R^2	.16			.24			.27			.33			.26		

Note: Age, income, EVT scores, and less-structured and structured time composite scores are mean-centered. For each model, observations where Cook's D > 3 standard deviations above the mean were identified and removed. $N_{Model1} = 45$; $N_{Models2\text{-}3} = 44$; $N_{Models4\text{-}5} = 43$; † $p < .05$; ‡ $p < .01$; ⋆ $p < .001$

2.3 Discussion

Findings from Chapter 2 offer preliminary evidence of a relationship between the time children spend in less-structured and more-structured activities and the development of their self-directed executive function. When considering our entire participant sample, children who spent more time in less-structured activities displayed better self-directed control, even after controlling for age, verbal ability, and household income. By contrast, children who spent more time in structured activities exhibited poorer self-directed EF, controlling for the same factors. The observed relationships between time use and EF ability were specific to self-directed EF, as neither structured nor less-structured time related to performance on externally-driven EF measures. These findings represent the first demonstration that time spent in a broad range of less-structured activities outside of formal schooling predicts goal-directed behaviors not explicitly specified by an adult, and that more time spent in structured activities predicts poorer such goal-directed behavior. Consistent with Vygotskian developmental theory and programs that build on that theory, such as Tools of the Mind, less-structured time may uniquely support the development of self-directed control by affording children with additional practice in carrying out goal-directed actions using internal cues and reminders. That is, less-structured activities may give children more self-directed opportunities. From this perspective, structured time could slow the development of self-directed control, since adults in such scenarios can provide external cues and reminders about what should happen, and when.

Surprisingly, the relationship between less-structured time and self-directed control changed with age in our participant sample, such that less-structured time predicted self-directed control in all but the oldest quartile of participants. This interaction between less-structured time and age was reliably observed across increasingly restrictive measures of less-structured time. One interpretation is that most but not all age groups within our sample spent their less-structured time in activities that encourage the development of self-directed control. Indeed, despite a relatively limited age range, our sample demonstrated differences in the content of less-structured time across

6 to 7 years of age, with older children spending more time engaged in media and screen activities. However, time spent in unguided practice, enrichment outings, and some forms of play was the main driver of the relationship between less-structured time and self-directed control in our data, and time spent in such activities did not change as a function of age. Another possibility is that children who have less developed self-directed control are more likely to benefit from less-structured time (in the same way that some interventions show the greatest benefits to children who show the worst initial performance (Connor et al., 2010; Diamond & Lee, 2011)(cf. Bierman et al., 2008), such that the oldest and most advanced quartile of participants showed the least benefit.

While promising, it will be important for the present findings to be replicated and extended to address a number of limitations. For example, our sample came primarily from an affluent, suburban sample. This sample nonetheless included a broad enough range of incomes that income was predictive of self-directed EF, and the relationship between less-structured time and self-directed EF held even when controlling for income. However, less-structured time may be especially beneficial to children in safe, quiet, resource-rich environments, so it will be important to test whether it differentially relates to self-direction in more impoverished environments. In addition, although the current test of the relationship between less-structured time and self-directed EFs emerged from a targeted hypothesis, we conducted multiple post-hoc exploratory analyses to explore the relationship between specific activities and self-directed control, which are not ideal conditions for statistical inference.

Another limitation of the present study relates to our constructions of less-structured and structured time, which are imprecise, and most likely fail to capture important differences across activities. The broad, standardized definitions of structured and less-structured time adopted in this study (e.g. Meeks & Mauldin, 1990) ignore differences in the degree of independence that children experience within and across activities. In the present study, trips to museums, libraries, and sporting events are each classified as less-structured, but may vary in relative structure. That is, a typical library visit, where children may select their own sections to browse and books to check out, may involve much less structure (and more self-directed time) than a typical sporting event, where attention is largely directed toward the action on the field or court. Similarly, although any activity within the category of "media and screen time" counts as less-structured time, this category includes activities that range from passive movie-watching to self-directed internet searches to more structured video games. Even those activities that seem less-structured by definition, such as free play, can quickly become more structured when adults, older siblings, or peers impose additional rules or criteria. Indeed, many programmatic interventions have highlighted the importance of some structure to improve the quality of children's play and other learning experiences, and produce benefits (Diamond et al., 2007; Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010; Lillard, 2012; Lillard & Else-Quest, 2006).

We note however, that even though our classification system based on the existing literature does not capture these variations in exactly how structured various activities are, our primary finding of the relationship between less-structured time and self-directed EF holds across analyses dropping potentially more difficult-to-interpret classifications (e.g., media and screen time, various games, movies and performances, and visits to museums, aquariums and zoos). To generate a more precise estimate of the amount of time children spend pursuing activities in a self-directed way, one would ideally assess child time directly, possibly by supplementing parent-reported child time use data with direct observation. One possibility along these lines could be to employ experience sampling techniques (G. Miller, 2012), where parents are frequently queried (via cell phone or another mobile device) throughout the day and asked to provide specific detail about their child's activities in the moment. Such methods would also minimize the need to rely on a parent's memory for their child's daily activities and experiences. We view our work as providing an important starting point for this kind of more time-intensive study of children's time outside of formal schooling and its relationship to their self-directed EF.

In addition, although we have identified links between child time use and self-directed EF, we are unable to draw firm conclusions about whether the observed relationships were driven by activities occurring in the week preceding the test session (as has been observed in other domains, e.g. Berns, Blaine, Prietula, & Pye, 2013; Mackey, Miller Singley, & Bunge, 2013), activities occurring over a longer period, or some combination. –We used composite measures incorporating both recent and more distal/typical experiences, given that these measures were correlated and in an attempt to maximize the accuracy and reliability of parental estimates. We can test which one is more predictive of self-directed EF, recent or more distal/typical experiences, but it is difficult to make strong claims based on such analyses. For example, when examining less-structured activities and self-directed EF, we find that recent experiences predict self-directed EF ($F(1,60)=6.10$; $p < .02$), but typical experiences do not ($p > .6$). This finding could reflect the greater importance of recent experiences, or it could reflect the greater precision of the time-diary measure, which indexes recent experiences but is also representative of more distal/typical experiences.[10] Similarly, when examining structured activities and self-directed EF, we find that neither recent nor annual experiences alone predict self-directed EF (ps $> .2$). This finding could reflect the importance of the combination of recent and distal experiences, or simply the greater robustness of using a composite measure. Therefore, while we have posited that less-structured experiences allow children to practice self-directed, goal-oriented behavior, producing benefits over time, we cannot discount the possibility that observed linkages may have been driven by recent experiences which increased self-directed behavior. In either scenario, regular participation in less-structured activities would yield benefits.

Future investigations of the relationship between self-directed control and less-structured time would also benefit from the inclusion of additional measures of self-directed control, which more closely approximate real-world child behaviors. This process may benefit from the development and validation of new measures of self-directed control in children. Establishing effects using tasks tapping other forms of self-direction would also ensure generalizability. For instance, in the present

[10] Recent less-structured experiences also predict self-directed EF when controlling for parent-reported typicality of the prior week ($\eta_p^2 = .02$; $F(1,57)=9.12$; $p < .004$; $M_{typ} = 4$; SD = 2.05; range = 1-7), and there is no interaction between less-structured experiences and typicality in predicting self-directed EF ($p > .8$). These findings might suggest that the prior week's experience is predictive separate from the extent to which it reflects typical/distal experiences. However, this interpretation rests on the validity and sensitivity of the typicality measure, which is unknown. Parent-reported typicality is at least internally consistent with parent-reported time use. Specifically, recent less-structured experiences predicted typical/distal experiences when parent-reported typicality of the prior week was high ($M_{typ} = 6$; $\eta_p^2 = .03$; $F(1,60) = 5.81$; $p < .02$), but not when typicality of the prior week was low ($M_{typ} = 2$; $p > .9$), yielding a marginally significant interaction ($M_{typ} = 2$; $\eta_p^2 = .04$; $F(1,60)= 2.92$; $p < .093$).

study, time in less-structured activities such as family outings may have benefited verbal fluency performance in a specific way, by fostering the development of more well-organized semantic networks, rather than by more generally improving children's abilities to generate their own rules for how and when to employ executive functions to achieve their goals. This alternative account cannot explain the full pattern of results in the link between less-structured time and self-directed EF (e.g., the fact that this link persists when enrichment activities are excluded, and other less-structured categories such as unguided practice and play predict self-directed EF); however, a broader range of measures could provide a more robust and generalizable assessment of self-directed EF.

The findings of the current study are consistent with previous research in showing a link between children's experiences and EF (Bergman Nutley et al., 2011; Bierman et al., 2008; Diamond, 2012; Diamond et al., 2007; Holmes et al., 2009; Lillard & Else-Quest, 2006; Röthlisberger et al., 2012; Titz & Karbach, 2014; Zelazo & Lyons, 2012). However, while the current study found specific effects of time use on self-directed but not externally-driven EF, previous research found effects of training and preschool interventions on externally-driven EF (but did not evaluate self-directed EF; e.g., see discussion in Diamond (2012)). There are several possible reasons for this discrepancy. First, previous training studies that have shown benefits for externally-driven EF have specifically trained children on aspects of externally-driven EF (e.g., working memory span tasks; e.g., Holmes et al. (2009); Bergman Nutley et al. (2011)). Likewise, while preschool and other interventions include a wide variety of experiences, they likely include considerable practice with externally-driven EF. In contrast, we hypothesize that less-structured time primarily affords children practice with self-directed EF, and thus may not transfer to improving externally-driven EF. Second, it is possible that differences between the current versus previous studies could be accounted for by differences between the externally-driven EF tasks they employed. Many previous studies that have found effects of interventions on externally-driven EF used task-switching or working memory span tasks (Bergman Nutley et al., 2011; Bierman et al., 2008; Diamond et al., 2007; Holmes et al., 2010; Lillard & Else-Quest, 2006; Röthlisberger et al., 2012; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009), whereas the current study used tasks assessing proactive control (AX-CPT) and conflict resolution (Flanker). It may be that specific aspects of externally-driven EF are more sensitive to children's experiences, or that specific tasks are more sensitive to individual differences in general due to better psychometric properties.[11] Future research using a more comprehensive battery of EF tasks could address these possibilities.

Another key difference between our study and such prior research is the correlational nature of our study, which supports at least two alternatives to the interpretation that how children spend their leisure time shapes their EF. First, children with better self-directed EFs may engage in (or be encouraged to engage in) less-structured activities more often. Likewise, children with poorer self-directed control may be more likely to engage in structured activities. Alternatively, the observed relationship between less-structured time and self-directed control may be driven by a third, unmeasured variable. Although we have attempted to control for some characteristics that might influence both time spent in less-structured activities and verbal fluency, such as household income, we have not controlled for other possibilities, such as parent EF and child's fluid intelligence (which we did not assess). However, we did control for child vocabulary (an index of crystallized intelligence), which may serve as a proxy for fluid intelligence in testing relationships with EF, given that EF fully mediates the correlation between crystallized and fluid intelligence in 7-year-olds (Brydges, Reid, Fox, & Anderson, 2012).[12] Moreover, such factors might be expected to predict both children's self-directed EF and their externally-driven EF (Ardila et al., 2005; Kalkut, Han, Lansing, Holdnack, & Delis, 2009; Mahone et al., 2002), and so seem unlikely to explain why less-structured time predicts only the former. Similar issues have been raised in interpreting links observed between children's EF and pretend play: rather than reflecting a uniquely causal role for pretend play in EF, EF may instead play a causal role in supporting pretend play, or pretend play

[11] For example, some EF-interventions have not improved performance on the Flanker task in this age group (M. Rueda et al. (2004); M. R. Rueda, Checa, and Cómbita (2012); see also Diamond et al. (2007), which introduced switching demands that did show effects of intervention, and included only incongruent trials so that a standard conflict score could not be computed). The Flanker task can be sensitive to minor variations in stimulus parameters (Paquet, 2001) and intervention dosage in adults (Liu-Ambrose, Nagamatsu, Voss, Khan, & Handy, 2012). Failures to find effects of interventions have also been attributed in part to the task's sensitivity to practice effects in pre-post measure designs (as discussed in Rueda et al., 2012), which are not an issue in the present study.

[12] We also note that there is ongoing debate regarding the inappropriateness of IQ as a control in models of cognitive processes (Dennis et al., 2009).

may be one of many activities promoting EF development in young children (Lillard et al., 2013).

An important direction for future work lies in establishing the directionality of relationships between child time use and self-directed EF via experimental manipulations. Longitudinal studies could provide the first step toward establishing directionality. Specifically, if time spent in less-structured activities prospectively predicts change in self-directed EF, this would suggest that less-structured time may play a causal role in the development of self-directed EF. If, on the other hand, self-directed EF prospectively predicts changes in the amount of time children spend in less-structured activities, this would suggest that self-directed EF may play a causal role in children's time use (e.g., because parents might allow children with strong self-directed EF skills to play with less supervision). While such longitudinal studies could thus provide important information about temporal precedence, this information is not sufficient evidence of causality (e.g., additional unmeasured variables could actually be the causal factors). Thus, future research using experimental manipulations of time spent in less-structured activities is necessary to definitively test causality. One approach would be to attempt to randomly assign children to more structured or less structured environments, such as summer camps, where child activities could be carefully monitored via regular sampling of staff and/or on-site observation. Although this kind of work is ambitious, and poses challenges, it could be used to inform more targeted laboratory-based training studies.

Chapter 3

Etiology of longitudinal relationships between children's verbal fluency switching performance and exposure to structured environments

3.1 Background and Motivation

Although contemporaneous links between children's time use and their verbal fluency switching ability in early childhood are intriguing (Chapter 2), correlational designs cannot address questions of causality. We have hypothesized that children's time use may shape their developing EF, such that children who have more opportunities for self-directed behavior develop better skills for engaging such behaviors in the absence of external reminders. Alternatively, children who are more successful in engaging EFs in self-directed contexts may participate in (or be encouraged to participate in) less-structured activities more often; similarly, children who show worse self-direction may be more likely to participate in structured activities. A third option is that observed links are driven by an unknown, unmeasured variable, such as parent EF or intelligence. For example, parents with better EF may encourage their children, who have inherited better EF, to engage in less-structured activities.

Another question is whether observed links between time in less-structured activities and developing self-directed EF will generalize to other populations. The sample in Chapter 2 was primarily affluent and suburban. It is possible that less-structured time is especially beneficial to children in safe, quiet, resource-rich environments. Time in less-structured activities may be less beneficial – or even detrimental to developing EF – in environments were few structured learning opportunities with adults exist. Such interactions could be tested via a more representative sample.

To probe causal links between early time use and developing EF, the present study investigates relationships between children's verbal fluency switching performance, structured time participation, and household structure within a broader, genetically informative sample, the Colorado Longitudinal Twin Sample (LTS). [1] This approach affords several advantages over the correlational analyses conducted in Chapter 2. Initial relationships between exposure to structured environments and VF switching performance can be tested via cross-lagged longitudinal analyses, where two constructs with a hypothesized causal relationship, X and Y, are assessed at both ages, so that the effect of X at age 1 on Y at age 2 can be estimated independent of Y at age 1 and X at age 2, (Kenny, 1975). Additionally, because children's activity participation and verbal fluency performance were measured using several surveys and prompts in the LTS dataset, latent variables for both verbal fluency switching performance and structured time participation can be generated at multiple time points across development, providing more accurate estimation of relationships across variables within and across time points. Testing relationships between constructs at the latent variable level reduces measurement error by extracting variance common across each indicator, resulting in better estimates of each child's true score for each construct at each time point.

3.1.1 Relationships between Structured Activity Participation and Verbal Fluency Performance in a Genetically Informative Sample

Although cross-lagged panel analyses can be used to rule out causal pathways, they do not afford true causal inference, as even temporal relationships may reflect non-causal factors, including shared genetic influence across parents and children. We have proposed that children's early time in structured activities causally effects the development of self-directed forms of executive function. However, even if we were to observe temporal precedence in a cross-lagged panel model, such that earlier time in structured activities predicted later verbal fluency switching performance, we could

[1] We focus on children's structured activity participation because assessments of children's participation in less-structured activities were measured inconsistently across time, and across levels of analysis (at the household versus child level). Child-level indicators of less-structured activity participation were measured at year-4, but not year 7 or year 16; likewise, critical household-level indicators of tolerance for less-structured activities (i.e., familial independence) were measured only at year-7. Therefore, this Chapter focuses exclusively on relationships with measures of environmental structure.

not rule out the influence of common genetic factors. For example, parental genotypes might contribute to each construct via genetic influences on children's self-directed EF, and via choices that influence shared environmental experiences in the home (e.g., typical patterns of time use). Inherited traits may simultaneously influence how children shape their environments (e.g., children with better self-directed EF may seek out different experiences, shaping patterns of time use) or passively elicit changes in the environment (e.g., children with low self-directed forms of EF may be enrolled in more structured activities, changing patterns of time use). Each example is a form of gene by environment correlation, which independently and collectively constitute sources of genetic mediation. Such explanations are plausible in the case of verbal fluency ability, as other forms of executive function are highly heritable in childhood (Engelhardt, Briley, Mann, Harden, & Tucker-Drob, 2015) and in adults (Friedman et al., 2008).

One approach to evaluate evidence for genetic mediation of causal relationships is via multivariate twin methods. Distinct environmental and genetic contributions can be investigated using twin designs, which use intraclass correlations from monozygotic (MZ) and dizygotic (DZ) twins to estimate the extent to which variation in individual traits and abilities can be attributed to additive genetic (A; heritability), shared environmental (C), and nonshared environmental (E) sources of influence. Additive genetic influences include contributions from a large number of genes (or multiple alleles from a single gene) that additively contribute to a complex phenotype. Twin modeling techniques draw on three sources of information: MZ twins share all their genes; DZ twins share on average half their genes by descent, on average; and typically-reared twins share a common family environment. Given these conditions, phenotypes that show higher MZ within-pair correlations than DZ within-pair correlations suggest the presence of genetic influence on that phenotype. Shared environmental influences include all nongenetic factors that make twins similar to one another, including some aspects of the family environment, maternal hormone levels during gestation, and common peer groups. Nonshared environmental influences include all nongenetic factors that are independent across twins (e.g., distinct experiences outside the home, such as peer groups and teachers; distinct experiences in the home, such as differential treatment by family members), as well as measurement error. In practice, the proportion of variance explained by the influence of nonshared environment is measured as 1.0 minus the correlation between MZ twins, because MZ co-twins differences are driven by non-shared environmental experiences (Knopik, Neiderhiser, DeFries, & Plomin, 2017).

Genetic, shared environmental, and unshared environmental influences are traditionally modeled using the ACE framework. In ACE models, the correlation between the genetic effects (A) in the MZ twins is set to 1.0 (because they share all of their genes), and to 0.5 in DZ, who share on average half their genes by descent. The correlation between shared environmental influences (C) is set to 1.0 to reflect shared rearing across twins. Nonshared environmental influences (E) do not correlate across twins, by definition, and are set to zero for all twin pairs. To generate estimates of A, C, and E for a specific trait, separate MZ and DZ models are fit to covariance matrices relating performance across related twins.

Multivariate genetic methods extend the basic ACE model to analyze the covariance between variables and temporal tests of covariance across time; such tests can be used to explore how genetic and environmental factors contribute to observed relationships across traits or abilities. Relatively recent advances in multivariate techniques have allowed for more specific causal tests - specifically, to what extent cross-lagged paths between two variables at two time points is mediated by environmental or genetic influences (Luo, Haworth, & Plomin, 2010).

3.1.2 Genetic and Environmental Influences on Time Use and VF

Despite its frequent use in both clinical and experimental settings, the developmental trajectory of VF has not been investigated longitudinally in children, and studies have focused exclusively on total word production, rather than more specific switching indices. Multiple adult studies have shown both genetic and environmental contributions to performance on verbal fluency tasks (Hayiou-Thomas et al., 2006; Kavé, Shalmon, & Knafo, 2013; Kovas et al., 2005; Owens et al., 2011; Swan & Carmelli, 2002). Only one study has focused on a developmental sample (Kavé et al., 2013), and that work was limited to 4-year-olds, who show only nascent executive abilities. Modeling switch scores will provide a more robust estimate of developmental changes in self-directed forms of EF, since total production is influenced by both executive and non-executive abilities (e.g., vocabulary Unsworth, Spillers, & Brewer, 2011).

Similarly, investigations of genetic and environmental influences on child time use have largely focused on differences in physical activity and sedentary behaviors, and most have not explicitly modeled genetic contributions to structured versus less-structured leisure time. These studies generally show strong contributions to child time from shared environmental influences in young children, with increasing contributions from genetic influences in adolescence and beyond (Aaltonen, Ortega-Alonso, Kujala, & Kaprio, 2010; Stubbe, Boomsma, & De Geus, 2005). Only two studies have focused on pre-adolescent children, and both found no evidence of genetic influence on measures of physical activity. A study of physical activity expenditure in 4- 10 year old twins found that shared environmental influences accounted for the majority of familial resemblance across twins (Franks et al., 2005). In a second sample of 9 - 12 year-old twins, shared environment effects explained the majority (73%) of the variance in an accelerator-recorded measure of total physical activity over 7 days, with a smaller unshared environmental effect, and no significant genetic effect. In comparison, children's self-reported activity preferences showed a strong genetic contribution, and no significant contribution from shared environment (Fisher, Jaarsveld, Llewellyn, & Wardle, 2010).

In the present study, latent factors for twins' VF switching ability, participation in structured activities, and household characteristics theorized to support child autonomy were generated from multiple tasks and instruments at three time points across development (ages 4, 7, and 16 years). Phenotypic cross-lagged panel models were used to examine initial relationships across twin-level variables, and were subsequently modeled via genetic designs. Findings from this study will contribute to previous literature in three key ways: (1) by providing the first investigations of longitudinal relationships between children's structured activity participation and VF clustering and switching ability at the phenotypic level, using standard cross-lagged panel designs; (2) via investigations of genetic and environmental contributions to VF clustering and switching performance in childhood, using comparisons across monozygotic and dizygotic twins; and (3) via more specific tests of environmental and genetic mediation of those relationships.

3.2 Method

3.2.1 Sample

The sample included 936 individuals from 468 twin pairs recruited from the Colorado Longitudinal Twin Study. Of those twins, 506 were MZ (266 female, 240 male) and 430 were DZ (206 female, 224 male).[2] Families were located through birth records provided by the Division of Vital Statistics of the Colorado Department of Health from 1986 through 1990. Enrollment criteria included normal birth weight and gestation period, and a residence located within 2 hours of Boulder, Colorado; detailed information on sample characteristics is available in (Rhea, Gross, Haberstick, & Corley, 2006). Participants received compensation for each testing session. Zygosity was initially determined from parent and tester ratings on a zygosity questionnaire (Nichols & Bilbro, 1966) and subsequently confirmed via DNA genotyping for twins who remained enrolled in the study. For those twins, zygosity was confirmed via examination of twin concordance across a minimum of 11 highly information short tandem repeat (STR) polymorphisms.

3.2.2 Tasks and Measures

3.2.2.1 Semantic Verbal Fluency

Twins completed verbal fluency prompts at ages 4, 7, and 16 years. Parents reported activity participation for each twin at corresponding time points, and information about family household characteristics when twins were aged 3, 7, and 15. Details about tasks and measures at each age are provided in the following sections.

Twins independently completed three verbal fluency prompts at ages 4 and 7, and two at

[2] This number includes all twins who completed one or more study tasks or questionnaires. As detailed in Table 3.2, twin enrollment varied across the 12-year study interval, and not all twins completed all measures at each time point.

age 16 as part of a Specific Cognitive Abilities battery that included verbal, spatial, memory, and perceptual speed subtests.[3] Prompts were administered during a home visit at age 4, and in the laboratory thereafter.

Task prompts, administration and duration varied by age (Table 3.1). At age 4, the experimenter instructed the child, "Let's play a word game. I will tell you something that makes noise, and then you tell me something that makes noise that is different. How about a whistle. Now you tell me something that makes noise." If the child repeated the example (a whistle), the experimenter corrected the child, saying, "I told you about the whistle, now you tell me something different that makes noise." If the child did not appear to understand the task, the experimenter repeated the previous instructions, this time substituting 'radio' for whistle. After practice, the experimenter said, "Now tell me all the things that you know make noise." Children were then given 30 s to respond to each prompt. If the participant appeared to lose interest or stopped producing words during the test phase, the experimenter prompted, "Can you tell me some more things that make noise?" This process was repeated for each of 3 prompts (things that make noise, are soft, or are round), with primary and secondary examples varying by prompt (soft examples were 'blanket'/'pillows', round examples were 'ball'/'wheels'). At age 7, twins responded verbally to three oral prompts: 'Tell me all the animals you can think of', 'Tell me all the things you can think of that are round', '...that are made of metal'. Children were given 1 minute to respond to each prompt. At age 16, semantic VF was administered in a written format. Twins completed the prompts "List at the things you can think of [that are made of metal]/ [that are round]." Twins were given 3 minutes to respond to each prompt, and were not allowed to proceed to the second prompt until the full time for the first prompt had elapsed.

Responses to each prompt were transcribed and coded for switches between semantic subcategories.[4] Coders were trained to identify clusters of items that were semantically related (e.g., "dog", "cat", "gerbil") to demarcate between-cluster switches. As in past work, the final switch

[3] This battery was used to derive general cognitive ability (g-factor) in the present study.

[4] Children sometimes repeated experimenter-provided examples at age 4. These responses were removed from the dataset prior to the generation of production and switching indices.

score for each prompt was weighted by cluster size, reflecting the prediction that confidence in clustering and switching behavior should increase as cluster size increases (Barker et al., 2014; Snyder & Munakata, 2010, 2013). However, we modified this weighting procedure in the present study to address the concern that the simple weighting procedure used in past work (where each cluster contributed N-1 words to the total switch score for that prompt) could be biased in the presence of large cluster sizes. If children generate large clusters and switch infrequently, the switch score measure could be more likely to reflect factors that support production within clusters, such as semantic network density, rather than processes specifically supporting switching across clusters. We therefore modified the weighting procedure in the present study to incorporate a logarithmic adjustment that reduced the contribution of large clusters to the overall switch score. The switch score S of participant i was calculated according to:

$$S_i = \sum_{c}^{C} log(s_c) \tag{3.1}$$

where the switch score S_i was calculated as the sum over C log-adjusted clusters of size s_c, where s_c was defined as the number of words in the cluster. All prompts were scored by two coders, and both production and switching indices were averaged across coders. As in past work, coders demonstrated good inter-rater reliability for switch scores at each time point (all rs $> .85$).

3.2.2.2 Child structured activity participation

Indices reflecting children's participation in structured and less-structured activities were generated from items taken from the Child Behavior Checklist (CBCL Achenbach & Edelbrock, 1983) and an Interests and Activities Survey (IAS).

CBCL. The CBCL was administered to parents when twins were 4, 7, and 16 years old. Reporting format varied slightly at each time point. At age 4, parents indicated sports, clubs, and chores each twin participated in or completed, and rated the amount of time each twin spent in each activity relative to other children ("less than average", "average", "more than average", "don't know"). Reported activities were weighted according to parent-reported time spent in that

Table 3.1: LTS Structured Time and Verbal Fluency Measures at Ages 4, 7, and 16 Years.

Time	Child structured activity participation	Verbal fluency prompts
Age 4 strut	CBCL Activity Items: Structured Lessons & Activities Clubs/Organizations Chores	'Things that are round' 'Things that are soft' 'Things that make noise'
Age 7	CBCL Activity Items: Structured Lessons & Activities Clubs/Organizations Chores Interests and Activities Survey Structured Lessons & Activities	'Things that are round' 'Things made of metal' 'Animals'
Age 16	CBCL Activity Items: Clubs Chores	'Things that are round' 'Things made of metal'

Note: CBCL = Child Behavior Checklist; MFES: Moos Family Environment Scale; IAS = Interests and Activities Survey. † The MFES was independently administered to one or both parents at each time point. ‡ The IAS was independently administered to one or both parents at age 7. Ratings were averaged to form composites when information from both parents was available.

activity [5] and rated for relative level of structure using established criteria described in Chapter 2 (Meeks & Mauldin, 1990; Hofferth & Sandberg, 2001b; Eccles & Barber, 1999; Mahoney & Stattin, 2000; Fletcher et al., 2003; Osgood et al., 2005). As before, structured activities included adult-led lessons and practices (e.g, basketball, soccer, martial arts, music), organizational meetings or events (Girl/Boy Scouts, YMCA groups), and chores (housework, cleaning).[6] Activities were summed to form composite scores for time in structured activities. At age 7, parents completed just two items: an activity item (resulting in combined reporting of clubs and sports), and a chore item. At age 16, parents reported club involvement and chore completion. As in year 4, parents rated time spent in each activity at 7 and 16, and weighted composites were formed.

Interests and Activities Survey. At age 7, parents indicated twin participation in structured lessons and activities. This measure included both elicited and free-response items.

3.2.2.3 Household Structure

Moos Family Environment Survey (MFES). Household control and organization subscales were used to represent levels of household structure. Each subscale included items related to control over the home environment, e.g., "There are set ways of doing things at home", "Each person's duties are clearly defined in our family". Full scales at each age are given in Appendix A. Items were scored true/false (0/1) at age 3, and on an 5 point Likert scale at Y7 and Y15.

3.2.2.4 Covariates

General cognitive ability (g-factor). General cognitive ability (g) was calculated as the first principal component of child performance within a specific cognitive battery comprised of verbal, spatial, memory, and perceptual speed subtests. Subtests are described in detail in Kent & Plomin

[5] Parents did not report level of participation in chores; therefore, the chore variable represents the total number of chores reported.

[6] As in Chapter 2, these measures provide only coarse estimates of the amount of adult-structured time children were exposed to in each activity. The degree to which individual activities were adult-led likely varied across activities and ages (e.g., Boy Scout activities may have been largely organized by adults at age 4 and adolescent participants at age 16; basketball lessons may have been more structured than YMCA group activities at age 4, etc.). Thus, while we have used existing schema to classify activities as either 'structured' or not, these estimates likely encompass a range experiences both structured and unstructured, a point we will return to in the Discussion.

(1987) and Appendix C. Testers administered the battery during a home visit when twins were 4, and in the laboratory at years 7 and 16.

Vocabulary. Vocabulary was measured using the vocabulary subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) at age 4, the Wechsler Intelligence Scale for Children-Revised (WISC-R) at age 7, and the Wechsler Adult Intelligence Scale III (WAIS Wechsler, 1974) at age 16. Scaled scores were used at each age.

Parent occupational ratings. Occupational ratings for each parent were derived using the Revised National Opinion Research Center ratings of occupation status of the parents at the time of the 14-month testing (Hauser & Featherman, 1977). Ratings were z-transformed in mothers and fathers, then averaged to form a single indicator for each family.

3.3 Results

3.3.1 Analytical Approach

All analyses were conducted using Mplus (Muthén & Muthén, 1998, version 7.2). Models that included categorical (binned) indicators were estimated using weighted least squares, mean and variance adjusted (WLSMV); all other models were estimated using full-information maximum likelihood. Model fit was evaluated via overall chi-square (χ^2) tests, with $\chi^2 < 2 * df$ indicating good fit, as well as two tests insensitive to sample size: comparative fit index (CFI) and root-mean-square error of approximation (RMSEA). Established fit thresholds were used for each test, with RMSEA $< .06$ and CFI $> .95$ representing good fit (Hu & Bentler, 1999). Chi-squared difference tests were used to evaluate the significance of individual paths (χ^2_{diff}). In WLSMV models, significance was tested using the DIFFTEST option; in MLR models, significance was tested using Satorra-Bentler Scaled Chi-Square estimates (${\chi^{\mathbf{2SB}}}_{diff}$). In phenotypic analysis, the MPlus 'type = complex' option was used to generate chi-squares and standard errors adjusted for clustering within twin families. In genetic analyses, each twin was assigned a number (twin 1 or twin 2) using the same random assignments used in prior work (e.g. Friedman et al., 2008).

The distributions of all variables were examined for non-normality, since indicator non-normality can increase the standard errors of maximum likelihood parameter estimates, resulting in poorer estimates of model fit (West, Finch, & Curran, 1995). Variables that showed excess skewness or kurtosis were initially log-transformed. Because age-4 and age-7 VF prompts (apart from the age-7 'animals' prompt) continued to show censored, right-skewed distributions after transformation, problem indicators were binned into categorical variables. Year-3 prompts were coded such that switch scores of $0 = 1$; switch scores > 0 and $<= 1 = 2$; and scores $> 1 = 3$. Year-7 prompts (apart from 'animals', which showed acceptable skewness and kurtosis after transformation) were coded such that $1 =$ switch score of 0; $2 =$ switch score > 0 and < 1; $3 =$ switch score > 1 and < 2.5; and $4 =$ switch score > 2.5.

Finally, we note that we have foregone traditional tests of longitudinal measure invariance (which constrain factor loadings over time to test whether factors are interpreted similarly at different ages) because several measures were administered at some time points but not others. Inconsistency in measure availability results in differences in latent factor composition across time. As such, the present analyses cannot be used to investigate the stability of construct measurement across time and subpopulations (e.g., males versus females). Instead, these models can be interpreted as reflecting individual differences in each construct at each time point.

Descriptive statistics for structured activity participation, verbal fluency switching, and covariate variables are given in Table 3.2, along with twin 1/twin2 correlations for MZ and DZ twins. Categorical variable bins and twin 1/twin 2 correlations are provided in Table 2.3. Bivariate correlations for all measures are provided in Appendix A.

3.3.2 Phenotypic longitudinal panel models

To test relationships between child structured activity participation, household structure and children's VF switching performance *across time*, latent factors for verbal fluency switching performance and structured activity participation were fit to a series of phenotypic cross-lagged panel models. Each model tests whether age-4 indicators for each construct are predictive of corresponding indicators at later ages, controlling for earlier levels of each construct. For example, in Figure 3.1, paths B1 and B2 reflect autoregressive effects, or the effect of a construct on itself measured at a later time, reflecting the stability of individual differences in that construct from one time point to the next. Paths B3 and B4 represent cross-lagged effects of one construct on another at a later time point.

Children's structured activity participation and household-level structure were initially investigated in discrete panel models. This approach was taken for two reasons. First, indicators measuring family-level characteristics cannot be measured using standard twin methods, which capitalize on differences between twins in the same household; family-level reports do not capture such differences. Therefore, genetically-informative twin-level factors were developed independently of household-level factors to investigate the relative influence of genetic and environmental factors to observed relationships. Additionally, for theoretical reasons, combining indicators measuring traits at the level of the individual (structure time participation) with those measuring differences at the level of a family unit (household characteristics) likely results in a combined factor that captures variance across distinct constructs. Thus, while household levels of familial structure and twin-level measures of structured activity participation are correlated in the present study, they are analyzed in separate models.

We conducted several secondary models to test whether hypothesized third variables could explain any observed relationships between structured activity participation, household structure, and VF switching ability. Basic panel models relating each construct to verbal fluency performance were subsequently augmented with child general cognitive ability (*g factor*), vocabulary, and parent occupational ratings (as a proxy for SES). Such factors might relate to both VF and structured environments if, for example, children with low verbal ability in early childhood were encouraged to enroll in activities to improve language acquisition, or children from high-SES homes were given more opportunity to enroll in structured activities and also showed strong VF performance.

In addition to tests controlling for general cognitive ability and vocabulary, which evaluate whether relationships with VF were specific, or could instead be explained by other factors (general cognitive factors and/or crystallized intelligence), we conducted two tests exploring whether measures of semantic knowledge showed similar relationships with environmental structure. If exposure to structure affected affected children's vocabulary or semantic network development (or, alternatively, if such abilities affected their exposure to structure), we might predict subsequent changes in average VF cluster size, which shows stronger correlations with vocabulary than switching measures in adults (Unsworth et al., 2011; Lanting, Haugrud, & Crossley, 2009). If relationships between VF performance and environmental structure are driven by changes in semantic networks rather than changes in self-directed switching ability, we would predict vocabulary performance and cluster size measures might show similar patterns with measures of structure across time. We note, however, that cluster size and switching indices are particularly difficult to dissociate in young children who generate few switches and relatively small clusters; in such cases, clustering and switching measures are highly overlapping.

3.3.2.1 Structured activity participation and verbal fluency switching ability

To examine longitudinal associations between structured activity participation and VF switching ability, we first tested baseline models relating verbal fluency to structured activity participation across ages 4 and 7 (Table 3.1). Measurement residual correlations that significantly improved model fit were retained in all analyses.[7] A basic measurement model with no autoregressive paths showed poor model fit (χ^2 (48)= 84.06, p =.001; CFI = .882; RMSEA = .029). Inclusion of autoregressive paths from T1 to T2 and within-time correlations at T1 significantly improved model fit (M1; Table 3.4), and showed similar fit to a model including within-time correlations at T2 (M2). Adding cross-paths between traits at T1 and T2 improved fit relative to each model and yielded good overall fit (M3; χ^2 (46)= 61.78, p =.060; CFI = .950; RMSEA = .018).

The full cross-path model demonstrates that opposing influences may contribute to observed links between structured activity participation and VF switching ability (Figure 3.1). Children

[7] Residual correlations between chore indicators and structured activity measures significantly improved model fit, and were added to all models.

Table 3.2: Descriptive Statistics for Continuous Measures Used in the Longitudinal Analysis.

	N	Mean	SD	Range	Skewness	Kurtosis	*r*MZ	*r*DZ
Age 3 - 4 years								
Structured Activities								
CBCL Activities	591	1.15	.92	0.00-4.50	-.132	.865	**.99**	**.88**
CBCL Chores	591	1.37	1.01	0.00-3.00	.175	-1.17	**.99**	**.95**
CBCL Clubs	591	.303	.58	0.00-3.00	1.57	1.34	**.99**	**.97**
Household Structure								
FES Control	478	7.79	1.37	1.50-10.00	-.261	-.421	-	-
FES Organization	478	3.23	1.26	0.00-5.00	-.364	-.92	-	-
g-Factor	704	-.386	1.17	-5.39-3.28	.024	.520	-	-
Parental Occupation								
Vocabulary	708	10.14	5.24	0.00-26.00	-.304	.175	-	-
Age 7 years								
Structured Activities								
CBCL Activities	633	1.47	1.20	0.00-4.50	-.27	-1.21	**.96**	**.89**
CBCL Chores	633	2.19	.85	0.00-3.00	-.884	-.067	**.96**	**.86**
IAS Lessons	632	2.95	1.69	0.00-8.00	.17	-.36	**.95**	**.93**
Household Structure								
FES Control	478	29.52	3.69	14.00-41.00	-.261	-.421	-	-
FES Organization	478	30.94	5.01	16.00-43.00	-.364	-.920	-	-
Verbal Fluency								
Animal	696	4.69	1.86	0.00-10.23	.318	-.217	**.42**	**.16**
g-Factor	811	0.00	1.00	-4.22-2.87	-.165	.403	-	-
Vocabulary	820	10.41	2.89	1.00-19.00	.150	.201	-	-
Age 16 years								
Structured Activities								
CBCL Clubs	489	2.20	1.43	0.00-4.50	-.10	-1.11	-	-
CBCL Chores	489	2.34	1.18	0.00-4.50	-.30	-.69	-	-
Household Structure								
FES Control	346	17.15	3.46	6.00-25.00	-.358	.285	-	-
FES Organization	346	17.45	3.84	8.00-25.00	-.342	-.634	-	-
Verbal Fluency								
Round	808	3.77	2.32	0.00-14.23	.75	.55	-	-
Metal	806	4.35	2.72	0.00-14.95	.60	.03	-	-
g-Factor	571	.640	-3.27-3.78	-.324	.197	-	-	
Vocabulary	572	11.24	2.82	4.00-19.00	.094	-.146	-	-
Parental Occupational Ratings (*14 months*)	792	.010	.771	-1.63-2.31	.285	-.257	-	-

Note: All correlations significant at p < .05. Means, SDs, and ranges represent untransformed values, and reported skewness and kurtosis values are post-transformation.

Table 3.3: Bin Values for Categorical Variables.

			Age 4		*Age 7*
Bin Number	*Noise*	*Round*	*Soft*	*Round*	*Metal*
1	401	613	472	120	263
2	185	118	172	243	215
3	148	-	87	206	145
4	-	-	-	127	73
Total N	734	731	731	696	696
MZ Correlation	**.36**	**.31**	**.18**	**.11**	**.31**
DZ Correlation	**.29**	.01	.07	.04	**.22**

Note: Within-task participant attrition resulted in a change in N across the three VF prompts at age 4. Significant correlations signified by bold text ($p < .05$).

Table 3.4: Links between Child Structured Activity Participation and VF Switching Ability.

Model Comparison	χ^2	df	p	CFI	RMSEA	χ^2_{diff}	df
M1: Stability and Time 1 Correlation	74.05	48	.009	.915	.025		
M2: Correlated Change Model	74.78	47	.006	.909	.026	.149	1
M3: Addition of Cross-paths	62.95	45	.039	.941	.021	9.32^*	2

Note: * p < .01

who participated in more structured activities at age 4 showed marginally worse verbal fluency switching performance at age 7, controlling for earlier VF ability and contemporaneous structured time (path B_4; β = -.22; χ^2_{diff} (1) = 2.97; p = .085). This finding is consistent with the causal inference that early participation in structured activities contributes to worse self-directed EF later in childhood. Unexpectedly, the cross-path linking early VF to later structured activity participation was significant in the opposite direction, indicating that children with better VF switching ability at age 4 went on to participate in more structured activities at age 7, controlling for earlier participation in structured activities and contemporaneous VF switching ability (path B_3; β = .30; χ^2_{diff} (1) = 5.91; p = .015). This finding is consistent with a causal influence of children's early VF ability on subsequent structured activity participation. Such a pattern could indicate that children with better self-directed EF at age 4 seek out or are encouraged to enroll in more structured activities at age 7. These competing influences could explain why time in structured activities showed weak within-time correlations with verbal fluency switching performance (Chapter 2), and are explored further in subsequent sections.

3.3.2.2 Tests of SES and Vocabulary as Hidden Variables

To test whether VF switching ability serves as a proxy for an unknown variable that influences both time use and VF performance, we conducted focused tests controlling for vocabulary and SES. If either vocabulary or SES exert simultaneous causal influences on children's enrollment in structured activities in early childhood and verbal fluency performance, controlling for their influence will likely attenuate cross-lagged paths between earlier and later levels of each construct.

As shown in (Figure 3.2), controlling for children's vocabulary at each age did not attenuate the observed negative relationship between time in structured activities at age 4 and VF switching ability at age 7. The resulting model showed good fit ($\chi^{\mathbf{2}}$(65)= 92.26, p =.015; CFI = .946; RMSEA = .022), with vocabulary positively predicting VF switching performance at each age, and showing a weak positive association with structured time participation at age 7. After controlling for vocabulary knowledge, children who spent more time in structured activities continued to show

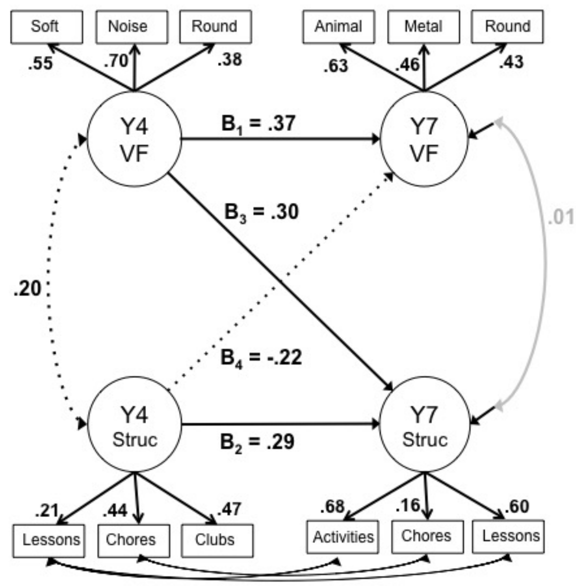

Figure 3.1: Phenotypic cross-lagged panel model relating structured activity participation and verbal fluency performance from age 4 to age 7. Children who showed better VF switching ability at age 4 spent more time in structured activities at age 7, controlling for age-4 structured activity participation and stability in VF switching ability from 4 to 7. The second cross-lag (B3) was also marginally significant, indicating that children who participated in more structured activities at age 4 showed worse performance in verbal fluency at 7. In all models, standardized parameters shown; trend-level paths are indicated by dashes; paths that do not meet statistical significance criteria are in gray. The model showed moderately good fit ($\chi^{\mathbf{2}}(45)= 62.95$, p =.039; CFI = .941; RMSEA = .021).

marginally worse VF switching ability at age 7 (β = -.23; $\chi^{\mathbf{2}}$ $_{diff}$ (1) = 3.31; p = .068). The positive cross-path between early VF performance and later structured activity participation also persisted (β = .27; $\chi^{\mathbf{2}}$ $_{diff}$ (1) = 5.10; p = .024), suggesting that early VF does not predict later structured activity participation simply because it is a proxy for vocabulary, which in turn predicts greater enrollment in structured activities early in childhood. These findings suggest that vocabulary development alone cannot explain observed links between structured activity participation and VF switching ability.

Similarly, controlling for SES did not alter paths between VF and structured activity participation. In a model including SES as a covariate loading on VF and structured activity participation at each age (($\chi^{\mathbf{2}}$(53)= 76.49, p =.019; CFI = .930; RMSEA = .022), cross-paths linking time use and VF were largely unaffected. Controlling for differences in SES, children who demonstrated better VF performance at age 4 participated in more structured activities at age 7 (β = .27; $\chi^{\mathbf{2}}$ $_{diff}$ (1) = 5.06; p = .025), accounting for earlier structured activity participation and contemporaneous VF. Similarly, children who participated in more structured activities at age 4 showed marginally worse performance in VF at age 7, controlling for earlier VF performance and year-7 structured time (β = -.23; $\chi^{\mathbf{2}}$ $_{diff}$ (1) = 3.08; p = .079)).

3.3.2.3 Specificity of relationships between structured activity participation and VF switching performance

We next conducted two sets of analyses to test whether the relationship between verbal fluency and participation in structured activities was specific to switching ability, or could instead be explained by other factors that do not reflect children's ability to engage EFs in self-directed contexts.

As an initial test, we examined whether the negative relationship between early structured activity participation and later VF switching ability persisted after controlling for general cognitive ability, or g-factor, at each age (Figure 3.3). General cognitive ability showed a strong positive association with VF switching performance and effectively eliminated the VF autoregressive path

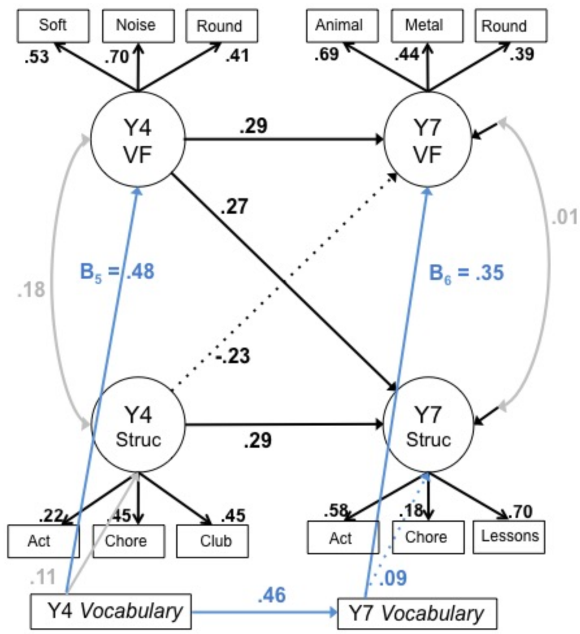

Figure 3.2: Controlling for vocabulary did not attenuate cross-lagged paths between VF switching performance and structured activity participation. The model showed moderately good fit (χ^2(65)= 92.26, p =.015; CFI = .946; RMSEA = .022).

between ages 4 and 7 ($\chi^{\mathbf{2}}{}_{diff}$ (1) = 2.61; p = .106). This finding suggests that stability in VF performance across early childhood at least partially reflects individual differences in general cognitive ability, insomuch that g-factor accounts for all common variance across year-4 and year-7 VF. However, relationships between time use and VF persisted after controlling for g, ruling out the possibility that VF relates to time use because it is a proxy for intelligence. Controlling for g-factor at each age, early structured activity participation continued to show a negative association with later verbal fluency performance (path B_4 r = -.22; $\chi^{\mathbf{2}}{}_{diff}$ (1) = 3.25; p = .071), and early VF ability continued to positively predict later structured time participation (β = .30;$\chi^{\mathbf{2}}{}_{diff}$ (1) = 5.33; p = .021).

Vocabulary. Children with better vocabulary knowledge typically produce more words in verbal fluency. Although we have attempted to control for differences in total word production by weighting switching measures to reduce bias from large clusters, it is possible that it is easier to detect switching ability in young children who know more words. If this is the case, replacing verbal fluency switching indices with an indicator of vocabulary performance should yield similar negative relationships between early time in structured activities and later vocabulary. To test the possibility that verbal fluency simply acted as a proxy for vocabulary knowledge at each age, we ran an additional cross-lagged panel model substituting verbal fluency factors with vocabulary performance at ages 4 and 7 (Figure 3.4). Structured activity participation at age 4 did not predict vocabulary at age-7 (β = .08, $\chi^{\mathbf{2SB}}{}_{diff}$ (1) = 1.42, p >.2), and early vocabulary did not significantly predict later structured activity participation (β = .08, $\chi^{\mathbf{2SB}}{}_{diff}$ (1) = 2.04, p = .153) suggesting that observed links are not driven by aggregate changes in vocabulary.

Cluster size. As an additional test of specificity, I evaluated whether observed relationships with structured variables were specific to self-directed switching processes in VF, as predicted by theoretical accounts suggesting that opportunities to complete activities independently of adults may benefit emerging self-directed EF. Latent variables were generated using participant average cluster sizes for each prompt at each age. The model showed good fit ($\chi^{\mathbf{2}}$ (45) = 47.02; p > .3; CFI = .995; RMSEA = .007; Figure 3.5). While VF cluster size at year-4 positively predicted

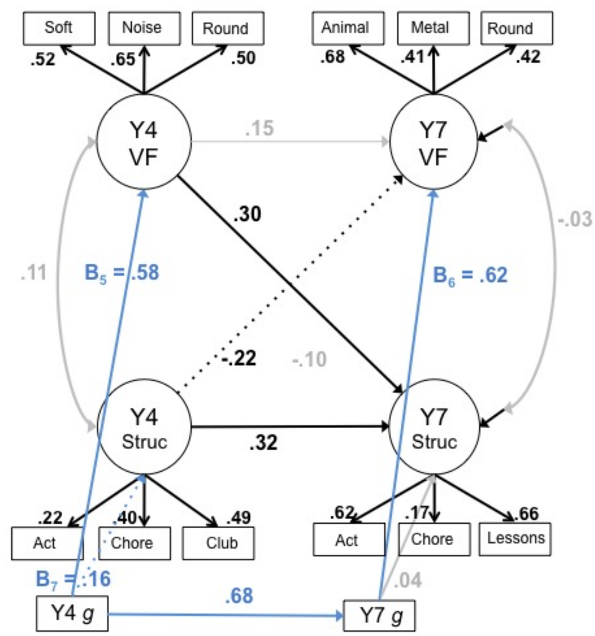

Figure 3.3: Controlling for general cognitive ability did not attenuate cross-lagged paths between VF switching performance and structured activity participation. The model showed moderately good fit ($\chi^2(65)$= 89.13, p =.025; CFI = .964; RMSEA = .020).

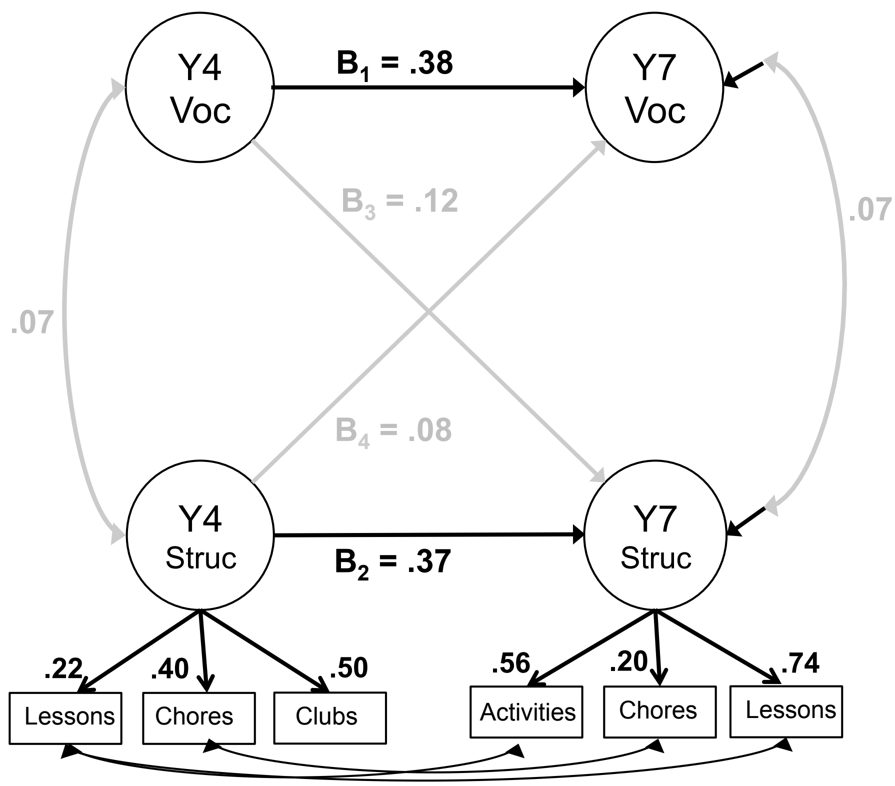

Figure 3.4: Child vocabulary and structured time participation show temporal stability, but levels of each construct at year-4 do not predict outcomes at year-7. Model fit was excellent ($\chi^2(13)=$ 19.36, p =.112; CFI = .972; RMSEA = .024).

structured activity participation at year-7 (β = .27, $\chi^{\mathbf{2SB}}{}_{diff}$ (1) = 4.96, p = .026), the negative path linking structured time to cluster size at year-7 did not meet significant criteria (β = -.31, $\chi^{\mathbf{2SB}}{}_{diff}$ (1) = 2.56, p = .110).

3.3.2.4 Relationships between structured activity participation and VF into adolescence

As a final step, we extended the model to test whether time in structured activities at age 7 continued to predict verbal fluency switching performance at age 16, controlling for temporal stability in each construct as in prior models. Because only two verbal fluency prompts and two measures of structured activity participation were collected at age 16 (chores and club participation), indicators for each construct were constrained to be equal. This constraint contributed to relatively poor overall fits across the base model and subsequent models testing for relationships with general cognitive ability, SES and vocabulary (general cognitive ability shown in (Figure 3.6). No relationship between structured activity participation and verbal fluency switching ability emerged between year 7 and year 16, which could suggest that observed relationships are specific to younger children, that the relationship does not hold using more restricted measures of structured activity participation, or that 'structured' activities show varying levels of adult involvement across age, a topic we return to in the Discussion.

3.3.2.5 Household structure and verbal fluency switching ability

We next evaluated whether household-level indicators of structure from the Moos Family Environment Survey showed similar relationships with children's verbal fluency, such that early exposure to a structured family environment was associated with worse self-directed switching in verbal fluency at age 7, and earlier VF ability predicted a more-structured environment at age-7. A cross-lagged panel model relating household structure and VF switching ability across ages 3 and 7 showed somewhat poor fit ($\chi^{\mathbf{2}}$ (29)= 49.15, p =.015; CFI = .938; RMSEA = .027; Figure 3.7). Controlling for earlier VF and concurrent household structure, children whose parents reported

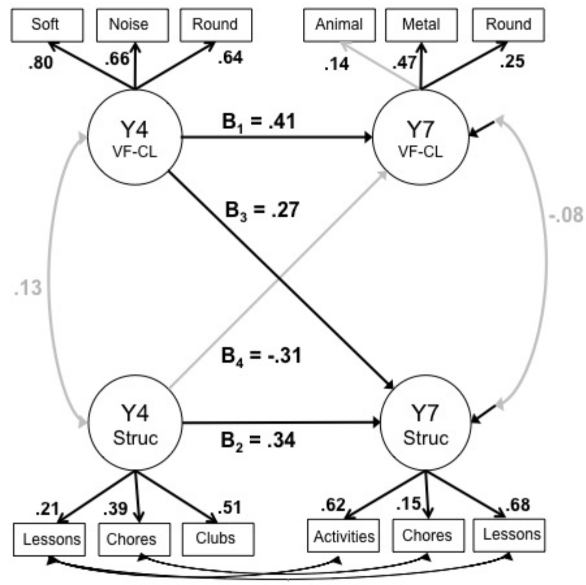

Figure 3.5: Children who produced larger average cluster sizes in VF at age-4 participated in more structured activities at age 7. The reverse cross-lagged path trended in the opposite direction, though this path did not meet significance criteria. The model showed good fit (χ^2 (45) = 47.02; p > .3; CFI = .995; RMSEA = .007).

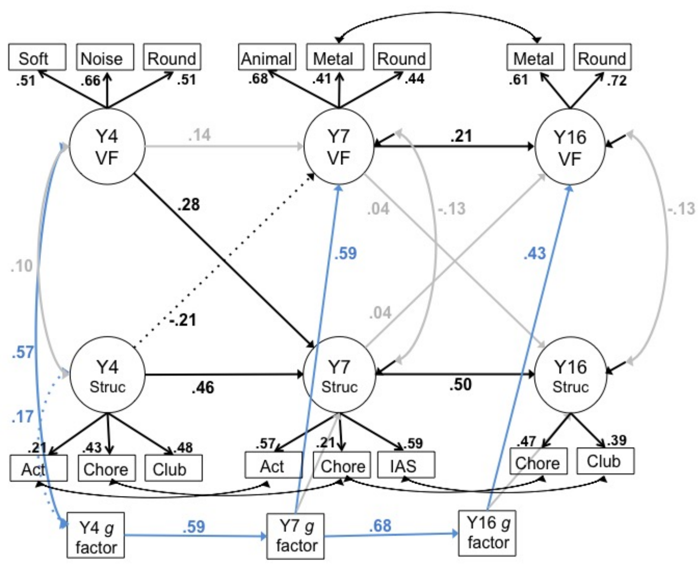

Figure 3.6: VF switching ability and structured activity participation did not show similar relationships across year-7 and year-16, such that each construct did not predict the other. The model produced poor fit (χ^2 (133) = 236.60; p < .0001; CFI = .910; RMSEA = .029), likely because unstandardized loadings for two-indicator factors were constrained at year-16.

high levels of household structure at age 3 showed worse VF switching ability at age 7 (β = -.23; χ^2_{diff} (1) = 4.68; p = .031). Children's VF performance at age 4 did not predict household structure at age 7 (path B_3; χ^2_{diff} (1) = 1.05; $p > .3$), challenging the reverse causal inference that children with worse self-directed switching at age 3 were exposed to more environmental structure at age 7, as would be expected if parents modified the household environment to accommodate poorer levels of child self-direction.

Tests of SES and vocabulary as hidden variables. As before, we tested whether the observed relationship between earlier exposure to household structure and later verbal fluency performance persisted after controlling for two potential confounding variables, vocabulary and SES. In each model, household structure at age 4 continued to show a negative relationship with subsequent verbal fluency switching performance, controlling for earlier VF ability and simultaneous structured activity participation (Vocabulary-augmented model path B_4 = -.22; χ^2_{diff} (1) = 4.46, p = .035; 3.8; SES-augmented model path BB_4 = -.28; χ^2_{diff} (1) = 6.01, p = .014. Additionally, as in the basic model relating household structure to VF switching performance, year-4 VF switching performance did not predict later household structure in either model, controlling for earlier structure and contemporaneous VF (ps > .3).

Specificity of relationships between household structure and VF switching performance

To test for specificity of relationships between household structure to self-directed switching indices, we evaluated whether relationships persisted in models controlling for general cognitive ability. Children exposed to more household structure at age 3 continued to generate fewer switches in VF at age 7 in models controlling for g at each age (path B_4 = -.23; χ^2_{diff} (1) = 5.60; $p > .018$, Figure 3.9). As a follow-up analysis, we again substituted VF switching ability for vocabulary performance at ages 4 and 7. Although this model demonstrated good fit ($\chi^2(6)$= 11.06, p =.087; CFI = .980; RMSEA = .031), neither cross-lagged path met significance criteria, indicating that earlier levels of vocabulary did not predict later household structure, and vice versa (p's > .3).

In a final test of specificity, we evaluated whether average VF cluster size demonstrated similar relationships with household structure across time. This model also demonstrated good fit

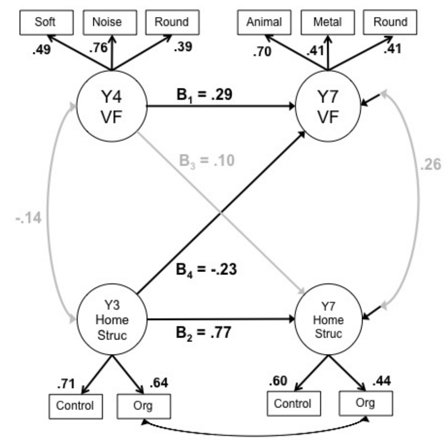

Figure 3.7: Children living in more structured household environments at age 3 showed worse verbal fluency switching performance at age 7, controlling for earlier VF ability and stability in household-level structure across time. The model showed somewhat poor fit, however, when loadings for household structure were constrained to equality: ($\chi^{\mathbf{2}}$ (29)= 49.15, p =.015; CFI = .938; RMSEA = .027.

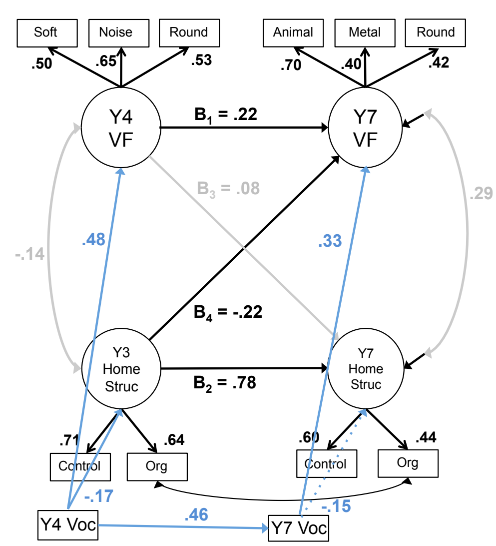

Figure 3.8: The relationship between children's earlier household structure and their later VF switching ability persisted after controlling for vocabulary at each age (path B4), though this model showed only moderately good fit ((χ^2(48)= 87.02, $p >$ .001; CFI = .920; RMSEA = .030).

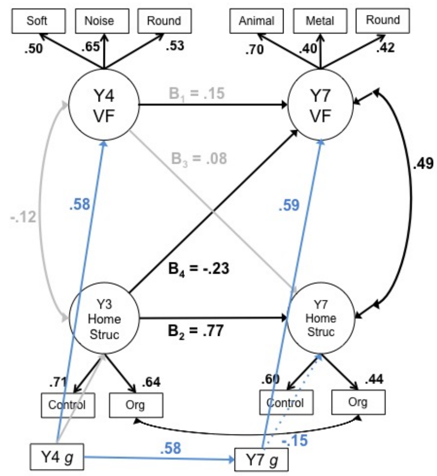

Figure 3.9: Cross-lagged panel model relating household structure to verbal fluency switching ability, ages 3-4 and 7, controlling for general cognitive ability at each age.

(χ^2(45)= 43.48, p =.053; CFI = .968; RMSEA = .023). Unexpectedly, children who experienced more household structure at age 3 produced smaller clusters at age 7 χ^2_{diff} (1) = 4.70; $p > .030$; Figure 3.10. However, the opposite cross-lagged path linking early VF cluster size to later household structure did not meet significance criteria ($p > .2$).

3.3.2.6

Relationships between household structure and VF into adolescence

Moos FES variables were only available for 346 twins at year 15, reducing power to detect effects. However, the relationship between household structure and later VF performance reversed from 7 to 15, such that twins in households with higher levels of structure at age 7 showed better VF performance at age 16 (χ^2_{diff} (1) = 4.08; $p = .044$). This path persisted in models controlling for general cognitive ability and SES; relationship with general cognitive ability shown in 3.11.

3.3.2.7

Composite Model

As a final step, I explored whether household structure and children's structured activity

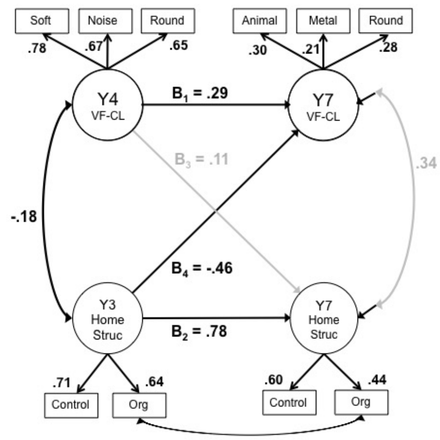

Figure 3.10: Cross-lagged panel model relating household structure to average cluster size in verbal fluency, ages 3-4 and 7.

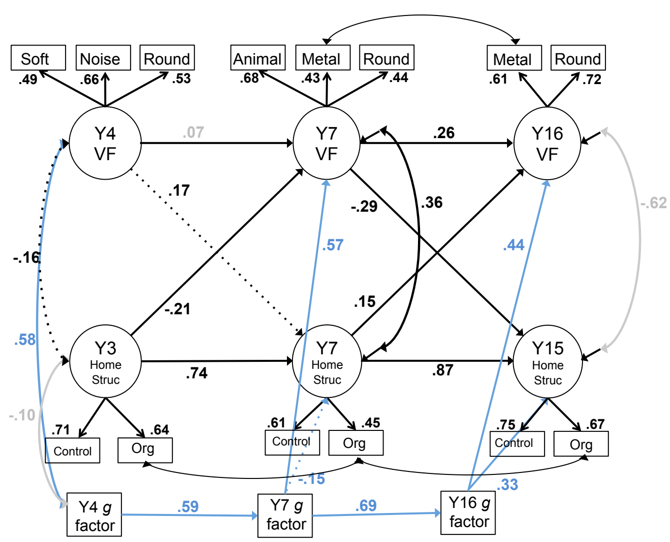

Figure 3.11: Cross-lagged panel model relating household structure to VF switching ability, ages 3-4, 7, and 15-16.

participation showed distinct relationships with VF switching ability, after controlling for their shared variance. This model demonstrated poor fit (χ^2 (92)= 126.62, $p > .009$; CFI = .917; RMSEA = .021). Although household structure and child structured activity participation were positively associated at ages 3 and 4 ($\beta = .15$; $p = .013$) and showed negative paths with year-7 VF switching, neither path reached significance (p's $> .3$). Modification indices suggested fit could be improved via correlations across family-level and twin-level structure factors, suggesting that these factors may capture different dimensions of a common construct (e.g., children's time in adult-structured environment).[8]

3.3.3 Genetic Analyses: Examining the Etiology of Links between Structured Environments and Self-directed Switching in Verbal Fluency

The longitudinal analyses in the preceding section are consistent with the hypothesis that time in more-structured environments causally affects children's ability to engage executive functions in self-directed contexts. Specifically, we found that twins who spent more time in structured environments at age 4 went on to show worse verbal fluency performance at age 7. Unexpectedly, we also observed that twins who switched more frequently in verbal fluency at age 4 went on to participate in more structured activities at age 7, though this relationship did not emerge in tests linking earlier VF to later household structure. One explanation for the latter relationship is that children who show the ability to engage EFs in self-directed contexts early in life are given more opportunities to participate in structured activities later in development, or are more likely to seek out such opportunities. Critically, however, controlling for temporal precedence does not definitively rule out hidden variables that may confound interpretation of links between verbal fluency and environmental structure.

One unexplored possibility is that observed links between environmental structure and self-

[8] This hypothesis could be tested via construction of formative, or latent composite (LC) variables at each time point composed of latent variables for household structure and structured activity participation. Because considerable debate remains over the interpretability of formative variables, particularly in their endogenous form (Bollen & Diamantopoulos, 2017; Cadogan & Lee, 2013), and questions of interpretability are further complicated in formative variables composed of latent variables, we have opted to present separate models for each construct.

directed switching ability in VF are driven by common genetic factors. If the same genetic factors that support developing self-directed EF also influence the structure of children's environments, hypothesized causal links between structured environments and VF switching ability could instead be driven by genetic mediation.

To test whether genetic factors simultaneously influenced children's VF ability and their time use, we used a cross-lagged genetic model to test for environmental mediation. In this model, variance decomposition isolates confounding influences so that the etiology of specific cross-lagged paths can be investigated. As in the standard cross-lagged panel design, genetic models adjust for shared variance across constructs at T1, the temporal stability of each construct between T1 and T2, and the reverse cross-lag (e.g., between earlier verbal fluency performance and later time use).

As a preliminary step, we generated common pathway models analyzing genetic, shared environmental, and nonshared environmental influences on VF switching ability and structured activity participation at ages 4 and 7 (Table 3.5). ACE models for verbal fluency switching ability show mixed genetic and environmental influences on abilities across time. At age 4, most variance was associated with the E component, representing strong effects of nonshared environment. Genetic and shared environmental factors also showed influence at age 4, though neither parameter met significance criteria. At age 7, unshared environmental and genetic factors influenced performance, though the A contribution only met significance criteria after the shared environmental (C) component was dropped from the model.

By contrast, structured activity latent variables showed no genetic influence (A) in year 7, and minimal, non-significant genetic influence at year 4. At each age, shared environmental variance contributed to structured activity performance. The influence of shared environmental factors is clear from comparison of MZ and DZ twin correlations, which illustrate this minimal genetic influence; as shown in Table 3.3.2, MZ and DZ twins showed similar correlations in both years, and correlations were high overall. These findings suggest that structured activity participation is driven by environmental rather than genetic influences in early childhood, and are consistent with past studies of child time use (e.g. Fisher et al., 2010; Franks et al., 2005).

Table 3.5: Genetic and Environmental (ACE) Contributions to VF Switching and Structured Time Participation at Ages 4 and 7.

	A_c	C_c	E_c	Model Fit
Year 4				
VF Switching	.29	.52	**.80**	χ^2 (36)= 55.86, p = .018; CFI = .812
Structured Activities	.21	**.97**	.12	χ^2 (37)= 43.59, p >.2; CFI = .997
Year 7				
VF Switching	.85	.16	**.50**	χ^2 (47)= 43.89, p = .602; CFI = 1.00
Structured Activities	.00	**1.00**	.09	χ^2 (37)= 55.55, p = .026; CFI = .990

Note: To improve fit, VF switching models were estimated with indicator-specific residuals instead of indicator-specific ACE estimates. Significant estimates (tested using difference tests) shown in bold.

We next tested the etiology of relationships between verbal fluency switching ability and structured time participation in two cross-lagged genetic models estimating individual contributions from ACE components via Cholesky decomposition. In a Cholesky decomposition, an observed variable is regressed on the latent variance components (A, C, and E) of the variables preceding it. Thus, in Figure 3.12, A1, C1, and E1 explains the total variance in verbal fluency switching ability at year 4. Latent variables representing year-4 structured time, year-7 structured time, and year-7 VF ability are subsequently regressed on A1, C1, and E1. Latent variables A2, C2, and E2 thus explain residual variance in year-4 structured time that is not correlated with A1, C1, and E1. A2, C2, and E3 explain additional variance in year-7 structured time and year-7 VF ability that is uncorrelated with A1, C1, and E1, and A2, C2, and E2. In this way, shared ACE contributions to temporally-lagged variables can be investigated by reordering variables, such that the preceding variable is in position 2 (e.g., year-4 structured time in 3.12), and the outcome variable is in position 4 (e.g., year-7 VF ability in 3.12).

Because the year-4 verbal fluency common path model showed relatively poor fit (Table 3.5), contributing to fit issues in the combined Cholesky model, we fit genetic cross-lagged models to latent variable factor scores extracted at the phenotypic level. Figure 3.12 represents decomposition of the age-4 structured time to age-7 verbal fluency path; the corresponding model for the age-4 VF to age-7 structured time is shown in Figure 3.13. In each model, path loadings sum to denote relationships at the phenotypic level. For example, the contribution of combined A, C, and E sources of influence to the association between structured activity participation and verbal fluency at age 4 is equivalent to $(a_{11} * a_{12}) + (c_{11} * c_{12}) + (e_{11} * e_{12})$. Thus, ACE estimates for the phenotypic cross-lag between year-4 structured activity participation and year-7 VF, controlling for contemporaneous correlation at age 4, stability in VF and structured time participation from 4 to 7, and the reverse cross-lag, is equivalent to the sum of the products of ACE paths linking structured activity participation at 4 and VF switching performance at 7, via latent variables A_2, C_2, and E_2. These paths sum to the phenotypic correlation, as follows: $(a_{22} = .14 * a_{24} = .03) + (c_{22} = .70 * c_{24} = .06) + (e_{22} = .24 * e_{24} = .43) = .15$.

As shown in Figure 3.12, the phenotypic cross-lag between year-4 structured activity participation and year-7 verbal fluency switching ability can be explained entirely by environmental contributions. Approximately 68 percent of the variance in the phenotypic cross-lag can be explained by the influence of nonshared environmental factors, as calculated by dividing the product of related path estimates by the phenotypic cross-lag: $(e_{22} * e_{24}/(a_{22} * a_{24} + c_{22} * c_{24} + e_{22} * e_{24})$, equivalent to $(.24 * .43)/.15$. The remainder of the phenotypic association is explained by shared environmental influence (though these paths do not reach significance), with additive genetic influences contributing <1 percent of the observed association.

The corresponding cross-lag between year-4 verbal fluency performance and year-7 structured activity participation can also be explained by environmental rather than genetic influences (Figure 3.13). Approximately 69 percent of the variance in the phenotypic cross-lag can be explained by the influence of shared environmental factors, and 28 percent is attributable to non-shared environmental influences, with negligible contributions from genetic factors.

3.4 Discussion

These results collectively support the causal inference that children who are exposed to more structured environments at age 4 go on to show worse verbal fluency performance at age 7, providing the first evidence of longitudinal links between structured environments and VF performance.

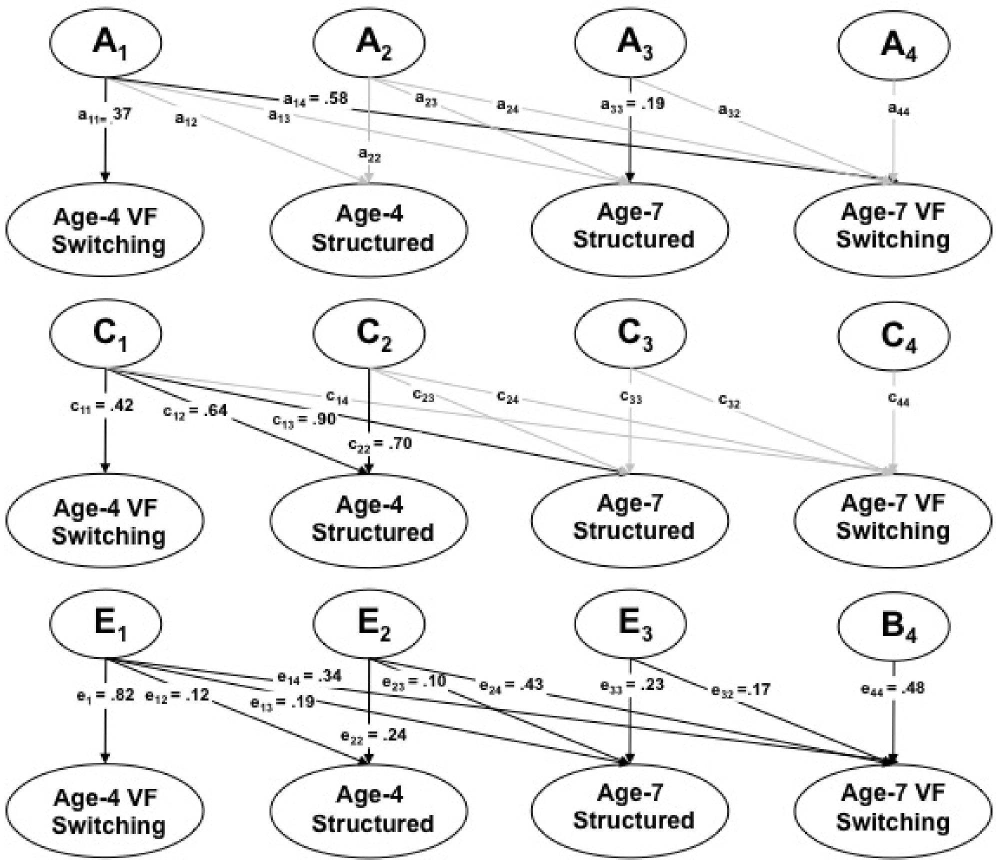

Figure 3.12: This model evaluates ACE contributions to the cross-lag from structured activity participation at age 4 to verbal fluency switching ability at age 7. Factor scores were used to evaluate relationships between VF performance and structured time (here represented as latent variables). Latent variables indicated contributions from A (additive genetic influences), C (shared environmental influences), and E (nonshared environmental influences). Path estimates are standardized; only one twin shown. Significance of mediating paths between year-4 structure and year-7 VF confirmed by difference tests; non-significant paths indicated in gray.

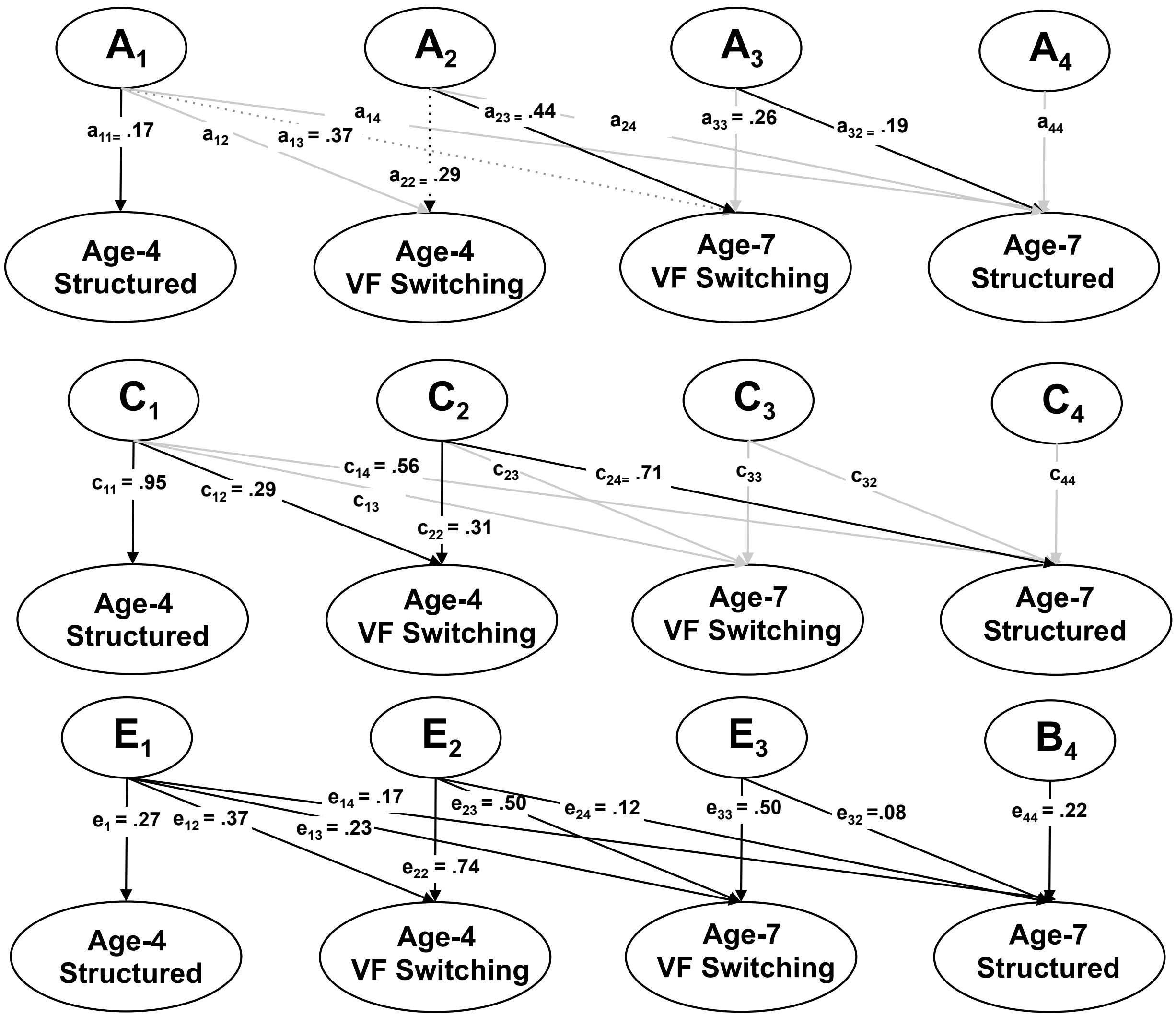

Figure 3.13: This model evaluates ACE contributions to the cross-lag from verbal fluency switching ability at age 4 to structured activity participation at age 7. Factor scores were used to evaluate relationships between VF performance and structured time (here represented as latent variables). Latent variables indicated contributions from A (additive genetic influences), C (shared environmental influences), and E (nonshared environmental influences). All path estimates are standardized; only one twin shown. Significance of mediating paths between year-4 VF and year-7 structure confirmed by difference tests; non-significant paths indicated in gray.

Negative relationships between early structure and later VF switching ability were observed for two measures of environmental structure, structured activity participation and household level organization and control, and persisted after controlling for multiple potentially confounding variables, including general cognitive ability, vocabulary, and a measure of socioeconomic status. Follow-up genetic analyses relating children's early time in structured activities to their later VF switching performance did not show significant genetic mediation. Instead, this relationship was largely explained by nonshared environmental factors, consistent with the hypothesis that that early structured time may causally affect later VF performance.

Unexpectedly, we also found that children who switched more often in verbal fluency at age 4 went on to participate in more structured activities at age 7. Although this relationship persisted in models controlling for SES, vocabulary, and general cognitive ability, it did not extend to our measure of environmental structure in the home. One explanation for this finding is that parent EF predicts both child EF and household structured activity preferences, resulting in a passive gene-environment correlation (e.g., parents pass on high EF to their children, and are also more likely to enroll their children in structured activities). This explanation is potentially supported by the significant shared environmental mediation (C) in the present model, which can indicate passive G-E mechanisms. The potential influence of such mechanisms could be investigated via alternative twin designs (e.g., Price and Jaffee, 2008).

Our findings also raise the possibility that verbal fluency tasks tap distinct cognitive processes across 4 and 7 year olds. In models controlling for general cognitive ability, the autoregressive path between year-4 and year-7 VF switching performance was attenuated, suggesting that covariance across these measures could be explained by general cognitive ability rather than individual differences in VF-specific processes. One interpretation of this finding is that 4-year-olds draw on distinct processes in verbal fluency, relative to 7-year-olds, and any remaining common variance across 4 and 7 can be attributed to general cognitive ability. To definitely test whether VF draws on distinct cognitive abilities across development, future tests could systematically explore other factors that might contribute to performance at 4 and observed links with structured time participation at 7, including confidence and behavioral inhibition.

Additionally, analyses of cluster size and environmental structure partially replicated patterns observed with switching-specific indices in VF, such that more time in structured activities at age 4 predicted lower average cluster sizes in VF at age 7. In adults, average cluster sizes are more likely to reflect differences in underlying semantic network robustness and density, suggesting that links between time use and VF may not be driven by switching–specific factors. If this is true, structured environments might show stronger influences on verbal fluency than other executive function tasks because they influence developing semantic networks. Though this possibility is intriguing, the present sample does not allow us to test this question definitively, as average cluster sizes were highly correlated with switch scores in our sample, particularly at age 4, where most children produced few words. Thus, coincident relationships between VF cluster size, switching ability, and environmental structure may reflect our inability to disambiguate clustering and switching processes in young children.

As in Chapter 2, conclusions about links between children's time use and their developing self-directed EF are limited by our available measures of time use, which show significant limitations. We have assumed that structured activity participation (i.e., the number of structured activities children participate in) serves as a reliable proxy for 'time in structured activities', but measures of participation cannot be used to draw inferences linking time in structured activities to outcomes. It is possible - and likely - that some children in this study spent a great deal of time engaging in a single structured activity, and others spent very little time in multiple structured activities. These differences likely introduced noise in the present analyses. Studies measuring response-to-dosage (e.g., how time in specific structured activities relates to specific outcomes) could help to clarify how structured experiences relate to the ability to engage EFs in self-directed contexts. Additionally, differences in factor structure across time driven by lack of measure availability may have attenuated effects, particularly in analyses relating year-16 to year-7. Additionally, because the present data set did not include reliable longitudinal measures of time in less-structured activities, we cannot rule out the possibility that observed relationships between structured activity participation and verbal fluency switching ability were driven by opportunity costs, such that time in structured activities reduced children's time in less-structured activities that are beneficial to verbal fluency switching performance. For example, although reading did not appear to drive links between VF switching performance and time use in Chapter 2, it is possible that activities such as co-reading (with adults) have stronger effects on verbal fluency switching performance in very young children.

More sensitive measures of time in structured activities could facilitate better analyses of how individual differences in experiences relate to developing VF ability. The present modeling approach estimates the rank order stability of children's verbal fluency switching ability and environmental structure over time. Although cross-lagged parameters in this model are often interpreted as between-individual effects of $X_{\text{t-1}}$ on Y_{t}, controlling for $Y_{\text{t-1}}$, such descriptions are imprecise: observed estimates from cross-lagged panel models pool within-subject and between-subject effects (Curran & Bauer, 2011; Berry & Willoughby, 2017). Under optimal circumstances, such effects would be disaggregated. For example, one could test whether an increase in an individual child's typical level of structure (above her typical trend, or baseline structure) predicted a subsequent increase in her self-directed switching performance, after adjusting for earlier VF performance and time-invariant covariates. Although techniques to capture within-subject variation in observed variables have been developed (e.g., structured residual approaches), limitations in the present data set (ordinal variables, limited time points) prevent their use (Curran, Howard, Bainter, Lane, & McGinley, 2014).

This study also raises the possibility that links between environmental structure and emerging self-direction may vary across development. Specifically, children living in more structured households at age 7 showed better verbal fluency performance at 16, controlling for potential confounding factors such as SES and general cognitive ability. One explanation for this finding is that structured environments begin to confer more benefits than costs in late middle childhood or early adolescence. This interpretation is consistent with a large body of work finding associations between structured leisure activity participation and several positive outcomes in older children and adolescents. For example, a study of 10$^{\text{th}}$ graders found that student participation in academic clubs, religious activities, volunteer organizations, and sports teams was associated with better academic performance (Eccles & Barber, 1999). These relationships persisted after controlling for verbal and quantitative skills. Some forms of adolescent structured activity participation have also been linked to improved emotional well-being (Bohnert, Kane, & Garber, 2008; Fredricks & Eccles, 2008) and decreased substance abuse (Eccles & Barber, 1999). Thus, adolescent participation in structured leisure activities benefits factors that protect against typical patterns of academic decline in adolescence, such as confidence or the desire to achieve status among similarly-inclined peers. Interestingly, unstructured leisure time predicts several *negative* outcomes in adolescence, including delinquency and low classroom engagement (Osgood et al., 2005). Unstructured leisure time may provide adolescents more time to engage in risk-seeking behaviors, highlighting the possibility that links between time use and emerging self-direction may vary across development. A second explanation is that 'structured' activities vary in relative levels of adult-structured time across development, such that 'structured' activities for older children and adolescents offer more opportunities for self-direction than similar such activities for younger children. Whereas adults play a central role in structuring organizational meetings and lessons for younger children, adolescents and older children may take on an increasingly central role in structuring organizational meetings and lessons. If relative levels of structure vary with age, broad classifications of time use may be particularly misleading in longitudinal samples, obscuring true relationships between environmental structure and emerging self-direction.

Chapter 4

What makes semantic verbal fluency self-directed? Exploring the computational mechanisms underlying associative versus controlled word retrieval processes

4.1 Background and motivation

Although we have posited that semantic verbal fluency draws on self-directed forms of executive function, recent theoretical accounts have challenged this account, instead suggesting that word production is largely supported by low-level associative memory processes (Abbott, Austerweil, & Griffiths, 2015; Hills, Jones, & Todd, 2012). This view has been supported by computational models of VF production in adults, where indices of word semantic similarity have been fit to sequential verbal fluency output. Using focused comparative tests, these accounts have argued that word sequences and response latencies are better explained by purely associative memory processes, rather than a combination of associative processes and goal-directed retrieval guided by subcategories. For example, Abbott and colleagues (2015) successfully reproduced adult-like VF data by modeling memory search as a random-walk across semantic memory, where production in verbal fluency is stochastically determined by the similarity of words in semantic memory. According to this account, production sequences that appear to reflect controlled clustering and switching between higher-order subcategories are in fact an epiphenomenal consequence of the underlying semantic structure of memory, wherein more semantically-similar words are more closely related in representational space. If VF production is largely driven by the strength of underlying associations across exemplars, differences in VF clustering and switching ability could simply reflect differences in network density and cohesion.

A competing associative account posits that memory search in VF follows an optimal foraging process, characterized by switches between two stages: global search, in which memory is probed for a new item to list based on frequency-based global cues, and local search, in which search is guided based on semantic relatedness to the previously-produced item (Hills et al., 2012). Under conditions of optimal foraging, participants select an initial exemplar on the basis of frequency (i.e., an exemplar that is frequently encountered in daily life), and produce subsequent items on the basis of their representational similarity to that high-frequency item and subsequently generated items (i.e., a potential response option's local similarity to other words in the current cluster). This process is repeated until there are no items found with sufficiently high similarity to the previously produced, at which point the individual moves to a new item that is frequently associated with the cue. While this dynamic search process yields 'patches' of similar words, and 'switches' between patches, optimal foraging accounts hold that item generation is not supported by strategic selection of subcategorical representations; instead, selection of the first word in a cluster is driven by a single, basic strategy (selection of a high frequency exemplar), and local item production draws on associative processes. More recent versions of optimal foraging models have suggested that search follows an associative Markov process, where only the most recently retrieved item is activated and supports search for a new exemplar (Hill, Bordes, Chopra, & Weston, 2015). Although proponents of this account have suggested that executive processes may support maintenance of global cues (Hill et al., 2015), these mechanisms are left unspecified.

Although both random-walk and optimal foraging models appear to provide a more parsimonious explanation for production patterns than competing models incorporating a role for subcategorical representations, they have been evaluated against topic-selection models that rely on hand-generated, unvalidated categories, which may account for their relatively poor fit. Both Abbott et al. (2015) and Hills et al. (2012) developed comparative models incorporating switches across subcategories, but neither model outperformed associative alternatives. In both instances, category membership was generated from a list of 22 category labels that was not rated for accuracy or completeness by external raters (Troyer et al., 1997). The list excludes some subcategories based on common thematic conceptual relations (e.g., 'zoo animals') and assigns membership inconsistently ('gorillas', for example, were classified as 'primates', but not 'African animals'). Other categories are likely incomplete, as they fail to capture idiosyncratic clusters; when asked to explain how they categorize animals in free response tasks, for example, adults have explained clusters as reflecting "unpleasant" or "human-like" dimensions, which are not represented in the Troyer coding scheme (Barsalou, 1983; Montez, Thompson, & Kello, 2015). Thus, it is unlikely that extant models testing the influence of category maintenance on production reflect the full range of representations individuals use to guide and constrain search.

Additionally, purely associative computational models of VF are difficult to reconcile with a large body of empirical findings suggesting that individuals who have intact semantic memory and compromised EF produce fewer words than neurotypical or mature controls. For example, clinical patients and children with observed deficits in EF reliably generate normal clusters, but fail to switch between clusters as frequently as neurotypical adults (e.g. Hurks et al., 2010; Koren et al., 2005; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). Such deficits are thought to reflect specific difficulties in engaging EFs in self-directed contexts. For example, many patients with frontal lesions perform well on cognitive batteries administered in structured testing environments, and quite poorly in less-structured, out-of-lab settings (Burgess et al., 2009; Eslinger & Damasio, 1985; Reitan & Wolfson, 1994; Shallice & Burgess, 1991; Wilson, 1993) despite demonstrating intact IQ and long-term memory, and performing like neurotypical controls on standard, externally-driven measures of executive function (e.g., digit span, Stroop, and card-sorting tasks (Alderman et al., 2003; Knight et al., 2002; Shallice & Burgess, 1991; Tranel et al., 2007)). Although proponents of associative theories have speculated that executive functions may support maintenance VF prompts or monitoring for production within clusters (Hills et al., 2012), the specific contributions of EF to VF switching ability have not been systematically modeled or tested.

Nevertheless, associative theoretical perspectives suggest an alternative interpretation of observed links between children's environmental experiences and their developing VF ability. If clustering and switching indices broadly reflect the strength of children's underlying semantic networks, it is possible that environments shape and support children's knowledge acquisition, rather than their executive function. Improvements in network density and cohesion may benefit production in VF even if the task draws on largely associative processes. For example, children who spend more time in less-structured activities may acquire more information about relationships between entities (e.g., animals, round objects), develop more robust semantic networks, produce larger clusters on VF, and switch between clusters more efficiently. If VF draws on associative rather than controlled forms of memory retrieval, developmental improvements in production may be driven by increases in the density and cohesion of children's semantic networks, rather than improvements in self-directed forms of EF.

To critically evaluate the role of abstract representations in guiding search during VF, models must incorporate valid representations. Therefore, the aim of the current chapter is to explore the feasibility of developing experience-driven abstract representations using real-world child texts, and to provide an initial test of how those representations support word search during VF. The architecture of the model builds on theories positing that word production is aided by maintenance and selection of higher-order categorical representations that facilitate selection of lower-level exemplars (e.g. Hirshorn & Thompson-Schill, 2006; Troyer et al., 1997). Higher-order representations reduce selection demands during word retrieval by reducing competition from words that do not fall in the currently-maintained category (Snyder & Munakata, 2008, 2010). Laboratory interventions designed to support children's production in verbal fluency support this view. For example, children who are primed with subcategorical labels in verbal fluency (e.g., zoo animals, farm animals) generate more words and switch more often on the task; these benefits extend even to unprimed subcategories (Snyder & Munakata, 2010). Abstract representations may help to reduce selection demands in endogenous control tasks where there are many options, by reducing the number of competing alternatives that children must choose from. Interestingly, priming young children with abstract representations yields more benefits to their VF production than training them to use a clustering and switching strategy; in young children, training in strategy can actually impair performance (Hurks, 2013).

Finally, the present model develops an approach to consider how intrusions from invalid VF responses could influence word production. Extant models of semantic VF have artificially constrained the search space by limiting model inputs to words that are typically produced in verbal fluency tasks. This optimization is problematic for two reasons. First, it overlooks a behavioral pattern that could yield insights into the processes individuals draw on to produce new words. Specifically, young children often respond to categorical prompts with words that are not obvious exemplars of that prompt (e.g., 'crayon' in response to the prompt 'things that are made of metal'). These failures suggest that models which constrain the search space to valid prompt responses do not incorporate a basic task demand: participants must maintain a representation of the task prompt ('things that are made of metal') to guide search and selection during VF. A second issue with constraining the search space is that invalid responses may facilitate production by guiding the participant to related exemplars (e.g., 'crayon' may lead the child to recall a metal desk where crayons are kept in the home). Because invalid responses may influence both associative and strategic forms of recall, the present modeling framework explicitly considers how they might affect production.

4.2 Overview of Modeling Approach

The model was designed to replicate typical patterns of child production in a standard verbal fluency task. To evaluate whether maintenance of abstract representations supported word production in verbal fluency, we abstracted topics from child-texts, rather than using unvalidated adult-generated categories, and incorporated those into the model. We then compared two models: a model where word selection was made on the basis of simple semantic similarity, reflecting associative processes (the *Associative* model); and a model where word selection was guided with equal contributions from both an index of semantic similarity and abstract topic representations (the *Topic-Support* model).

We developed the Topic-Support and comparison Associative model iteratively. First, to estimate semantic similarities across words typically found in a child lexicon, we evaluated word co-occurrence patterns within a corpus of children's texts. The Associative model operated over these simply estimates of word semantic relatedness. To generate the Topic-Support model, we derived probabalistic, higher-order topics from the same measures of word similarity informing the Associative model, using an unsupervised, data-driven process that capitalizes on the tendency of semantically-similar words to appear in similar textual contexts. The resulting higher-order representations of category structure are thus more representative of typical child clusters than the unvalidated, adult-generated categorization scheme used in previous modeling approaches. We subsequently evaluated how well the Associative and Topic-Support models reproduced child patterns of verbal fluency production during an 'animal' prompt.

4.3 Training Corpus

The training corpus was made up of 182 freely-available child texts. Of those texts, 98 were taken from Project Gutenberg.[1] This sample was augmented with a selection of books from freekidsbooks.org, which offers a range of texts categorized by subject. To develop a robust training corpus, we oversampled texts labeled with animal-relevant categories by selecting all books from categories that were likely to include descriptions of animals and their typical contexts ('nature', 'animals', and 'science' subcategories). All words in the corpus were used to train initial word embeddings to ensure that indices of semantic similarity capitalized on typical word contexts. Subsequently, standard stop-words (such as 'but', 'should', and 'or') were removed from the model data set. This process yielded 6349 unique words.

[1] This corpus is available at https://research.fb.com/downloads/babi/.

4.3.1 Generation of Higher Order Representations of Typical Word Contexts from Training Texts

Clusters of semantically related words were identified using a two stage training process. Semantic relationships across words were initially estimated as word embeddings, or distributed representations of word similarity, generated using the Word2vec algorithm. Word2vec uses a continuous-bag-of-words (CBOW) architecture to predict a given word from the words surrounding it (Mikolov, Chen, Corrado, & Dean, 2013). This approach produces a dense vector representation for each word, such that a word embedding for word n from document d in the training corpus is represented by a normalized M-dimensional vector x_{dn}.

Similarities among word embeddings can be exploited to identify clusters of words, representing abstract topics. Vectors of semantically-similar words are more similar, resulting in clustering of related words in the n-dimensional vector space. The similarity between words in the training corpus can therefore be quantified by the cosine distance between the corresponding word vectors. To identify clusters of semantically related words from the distribution of word embeddings generated via Word2vec, we used a topic model that assumes that the occurrences of words in a document is generated according to a spherical Hierarchical Dirichlet Process (sHDP; Batmanghelich, Saeedi, Narasimhan, & Gershman, 2016; Teh, Jordan, Beal, & Blei, 2006), a nonparametric process that infers how data are clustered within groups when the number of clusters (or latent factors) is unknown. sHDP was used to identify latent topics distributed across documents in the training corpus, where topics are delineated as a function of the density of words over a unit sphere, by identifying clusters of related words using the directional information represented in Word2vec word embeddings. sHDP capitalizes on the tendency of words that are more similar to be grouped in coordinate space, by drawing spheres around groups of semantically-related words that denote probabilistic topic membership (Batmanghelich et al., 2016). Cosine similarities between word vectors were modeled using a von Mises-Fisher distribution to identify topic 'centers'. In sHDP, topic generation is unsupervised, meaning a potentially infinite number of latent clusters can be used for prediction. As such, topics are automatically identified from word semantic information and co-occurrence patterns extracted from the training corpus. This method has been shown to generate more coherent topics (reflecting higher average similarity across member words as indicated via automated approaches and human raters) relative to alternative approaches, such as Gaussian Latent Dirichlet Allocation (Batmanghelich et al., 2016).

To derive the probability that a given word from a training corpus fell within a particular topic, we used sHDP to generate word-topic probabilities (i.e., the probability that a word fell within a topic in a single document/book from the training corpus). These probabilities were then averaged across documents to generate a probability density distribution for each word in the training corpus.

4.3.2 Model Evaluation and Selection

We generated 100-dimensional word embeddings for each word in the training corpus using Word2vec. These vectors were normalized using the l^2 norm and submitted to the sHDP process as in (Batmanghelich et al., 2016). To optimize sHDP model fit, we tested a range of parameter values, and evaluated each model for topic coherence using methods established in (Batmanghelich et al., 2016). Topic coherence was evaluated using Pointwise Mutual Information (PMI; Newman et al., 2010), using the training corpus as reference corpus.[2] This measure of topic coherence is an established approach that has been shown to effectively correlate with human judgments of topic coherence (Lau, Newman, & Baldwin, 2014). PMI is calculated over co-occurrence statistics for pairs of words (u_i, u_j) within 20-word sliding windows, such that:

$$PMI(u_i, u_j) = log\frac{p(u_i, u_j)}{p(u_i) \cdot p(u_j)} \tag{4.1}$$

[2] Previous tests of sHDP have used a larger external corpus such as Wikipedia to generate PMI metrics for evaluation. We have chosen to use the same training text to generate this metric, as it is likely word co-occurence patterns in child texts do not reflect patterns in adults texts. Future tests of the model would benefit from an independent evaluation corpus.

4.3.3 Evaluation of Verbal Fluency Production Models

The Associative model operated directly on the $l^2 - normalized$ Word2vec word embeddings x_{dn}. Successive words were selected according to semantic cosine distances. The model maintained a temporarily evolving semantic representation $\bar{x}$ of the history of l^2-normalized word embeddings x_{dn}, according to:

$$\bar{x}_{t+1} = \frac{1}{2}(\bar{x}_t + x_t) \tag{4.2}$$

, where x_t is the word produced in trial t. This model produced new exemplars by selecting the nearest neighbor in cosine distance to $\bar{x}$. Thus, the influence of previously selected words on the current $\bar{x}$, which informs selection of the next word, decays over time, such that the word immediately preceding the current trial has more influence in selecting the next representation.

The Topic Support model builds on this architecture by selecting words on the basis of two sources of information: semantic similarity (derived from the same $l^2 - normalized$ Word2vec word embeddings used in the Associative model) and abstract topical representations t_{dn}, derived from word-topic loadings estimated by sHDP. Like the Associative model, the Topic-Support model maintained a history $\bar{x}$ of word embeddings. This history was supplemented with the history of topic probabilistic distributions $\bar{t}$, so that each new word was selected according to its similarity to $\bar{x}$ (as in the Associative model) and $\bar{t}$. These two sources contributed equally to selection of each new word.

4.4 Results

4.4.1 Quality of Training Data

As a first step, we examined how well the text corpus captured typical child responses to an 'animal' prompt. Words retained in the training sample (i.e., words that appeared in the text corpus at least 5 times) were compared to 'animal' prompt production data from a sample of 400

Table 4.1: Closest Associates for Sample Animal Exemplars After Word2vec Training

'cat'	'whale'	'giraffe'	'owl'	'wolf'
dog	seal	python	farmer	man
mouse	dolphins	leafy	elf	Mowgli
elephant	penguin	magpie	pigeon	growl
bush	storm	picture	scene	tiger
cats	Killer	kookaburra	eagle	lair
puppy	Cassie	Giraffine	Buddha	bird
dad	shark	magnet	accent	cub
dormouse	captain	Sloth	Olmal	maiden
Samira	girl	kid	Dragon	afterward
Tefnut	Whale	foreign	Hunter	snake

7-year-olds.[3] Despite our use of a targeted, sizeable corpus of child texts, the training corpus did a poor job of capturing the typical vocabulary elicited from 7-year-olds during semantic VF. Of the 275 unique 'animal' responses extracted from the Chapter 3 response sample, only 142 were represented in the Word2vec word embedding sample.[4] Thus, the training corpus of 182 child texts yielded only ˜52% of the animal words generated in a typical semantic VF task by children in the target age-range.

Word2vec training generated reasonable estimates of semantic similarity (Table 4.1, though words appearing more frequently in the corpus ('cat' and 'whale') showed more sensible close associates than words appearing less frequently ('gecko'). As evident from the table, inclusion of children's fiction in the training sample resulted in uninterpretable or low probability associations for several words (e.g., the given name 'Samira'; the Egyptian cat-headed goddess Tefnut) in raw Word2vec output.

Since topic association strengths were derived from Word2Vec similarity data, they showed similar idiosyncrasies. Visual inspection of topic clustering in sHDP models showed that few animals clustered into obvious subcategories (Figure 4.1).

[3] These data were selected from a random sample of Chapter 3 participants.
[4] All unique words in this sample were judged valid responses to the animal prompt by raters.

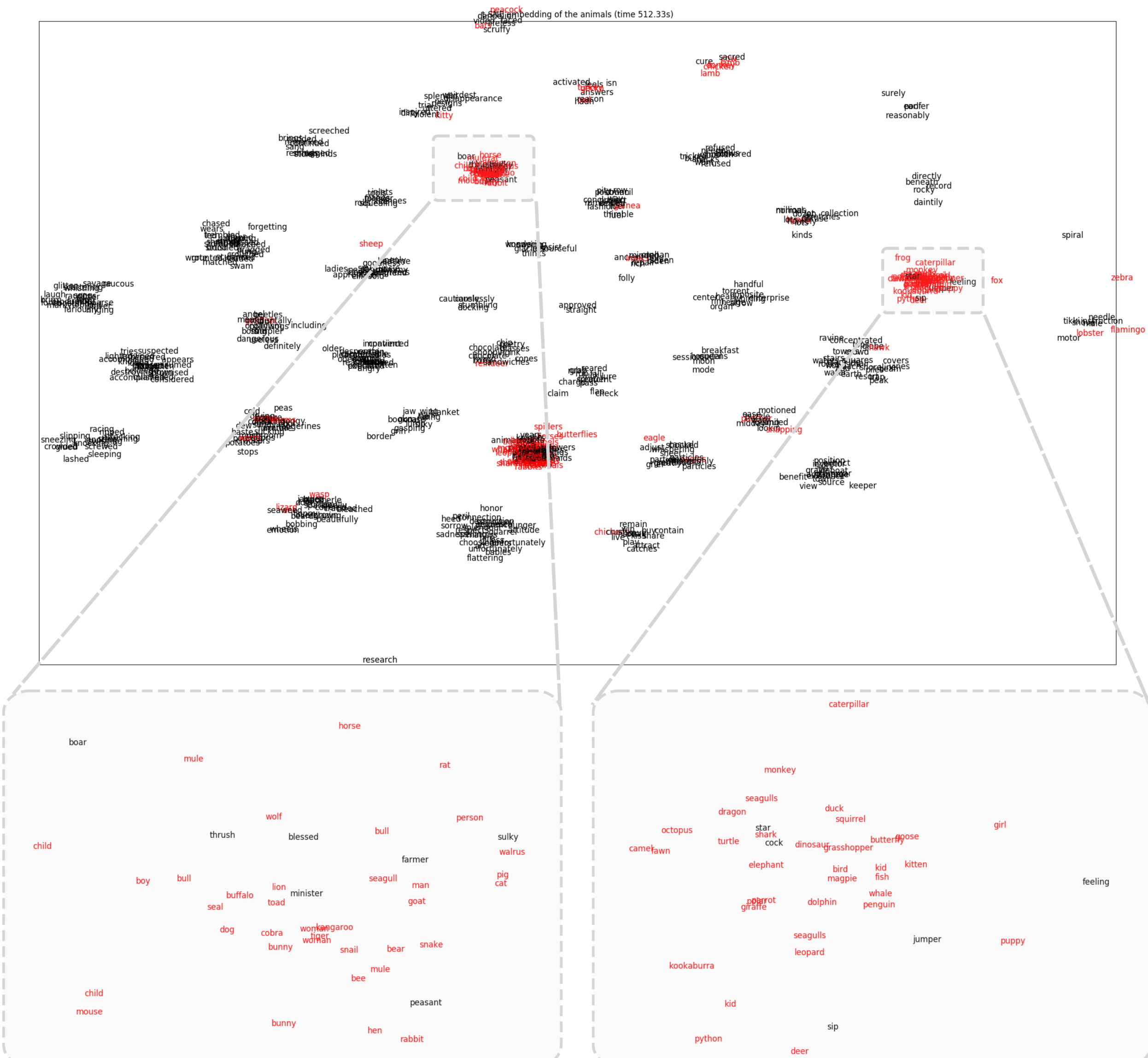

Figure 4.1: This figure illustrates clustering of animal exemplars derived from sHDP, with semantic distances across words corresponding to physical placement of words in space. Although not all word placements are interpretable, the red inset square shows a cluster of insects.

4.4.2 Model Evaluation Results

We evaluated differences in model performance by testing how accurately they reproduced child responses to an 'animal' prompt sub-sampled from Chapter 2. Each model was yoked to data from individual child response sets (N = 400, resulting in 400 runs per model), such that the first word spoken by a given child subject initialized the model on that run (i.e., by providing the starting word). Yoking initial words to participant data standardized starting contexts across models. As a further control, we limited total word production on each run to the total number of words by that participant.

To investigate differences in performance across the Topic-Support and Associative models, we calculated each models' percentage overlap with child-produced responses for each iteration, yielding 400 trials. The percentage of words from each trial that matched child responses were averaged to generate an overall performance metric. For example, if a participant produced 8 words, including the words 'bat', and 'cow', and those words were also produced by the model for that run, the model received a score of .25 for that trial.

To determine whether each model performed above chance, we compared model accuracy in each trial with 100 permutations (100 * 400 = 40,000 permutations). In each permutation, N-1 random words were drawn from the set of all words spoken by all participants during the task, with N indicating the number of words generated by a given child participant on a given trial. (One word was eliminated from this draw since the first word was initialized based on child data in model runs.)

In a linear mixed effects model comparing performance of each model to chance (i.e., permuted data), only the Topic-Support model showed above-chance reproduction of child responses on each trial (Figure 4.2). In this model, model accuracies relative to average permutation test accuracy (i.e., percentage of child responses successfully reproduced) were modeled as within-subject fixed-effects, and participants were modeled as random effects. The Topic Support model showed an above chance fit to child data (improvement relative to permutation test: 5.2 % (SE = 0.38), df

= 399, t = 13.53, $p < 0.001$)), with an average overlap with participants' VF data of .19%. In contrast, the Associative model did not show a fit above chance (improvement: -0.7% (SE = 0.28), df = 399, t = -2.07, p = 0.039; average overlap = 13%). In a follow-up direct comparison, the Topic-Support model showed better fit than the Associative model (improvement: 5.8% (SE = 0.52), df = 399, t = 11.32, $p < 0.001$).

4.5 Discussion

In this chapter, we explored the feasibility of developing experience-driven abstract categorical representations from a corpus of children's books, and provided an initial test of how those representations could support child-like word production patterns in a VF task. The architecture of the resulting model builds on theories positing that word production is aided by maintenance and selection of higher-order categorical representations that facilitate selection of lower-level exemplars (e.g. Hirshorn & Thompson-Schill, 2006; Troyer et al., 1997). Modeling how acquisition and maintenance of typical word contexts influences production dynamics in verbal fluency will ideally extend and challenge findings from recent associative models, and help to explain the failure of such models to explain longstanding empirical findings. The present model offers a data-driven way to derive probabilistic relationships between words and categories, and thus offers a valid alternative to previous approaches, which relied on unvalidated lists. Additionally, simple demonstration models suggested that incorporating a metric of topic similarity can aid selection over and above basic semantic similarity metrics, raising the possibility that previous comparisons favoring associative over category-based models of VF production may have been biased by crude measures of category membership.

Although direct comparisons with alternative models present an ideal method for evaluating the explanatory power of the proposed model, we have focused on replicating child patterns of production rather than explaining adult data. In future iterations, we plan to investigate whether the model reproduces behavioral patterns observed in adults, and seemingly explained by both random-walk and optimal foraging accounts: specifically, that the first item in a cluster is associated

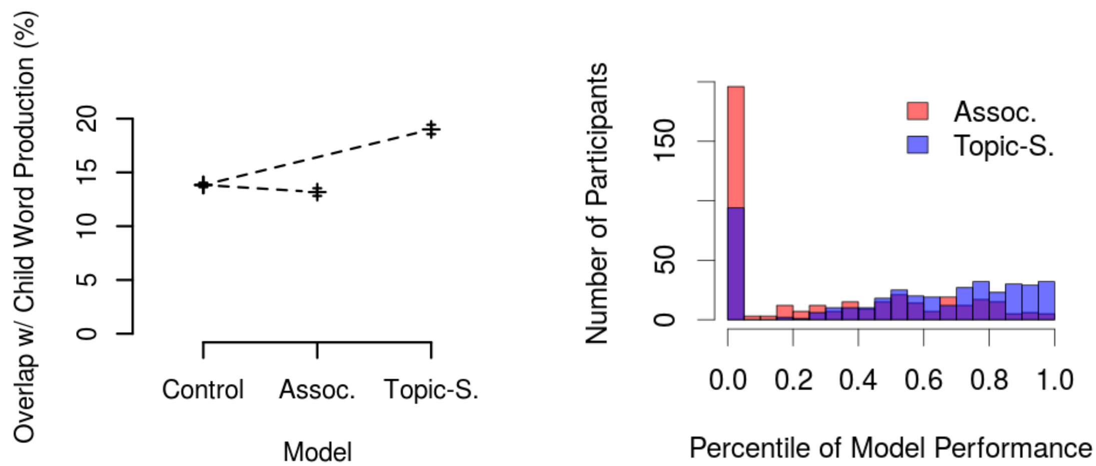

Figure 4.2: The Topic-Support model more accurately predicted child responses to an 'animal' prompt than the Associative model. **Left:** The Topic Support model, but not the Associative model, fit child VF data better than chance level (represented by randomly-permuted child data in the Control condition). Means represent average percentage overlap with child response sets (i.e., X out of X words correctly reproduced for a given child). Error bars indicate standard error (SE). **Right:** In comparison to random sampling during permutation testing, the Topic-Support model (blue) predicted VF data of children better than the Associative model (orange; overlap with Topic-Support model shown in purple). A score of 1.0 indicates that the model out-performed random sampling for a given child in 100 out of 100 permutation results. A score of 0.0 indicates that the model never out-performed random sampling in any permutation. The Topic-Support model shows more right-skew, reflecting greater accuracy in predicting individual responses from a given child's set.

with a relatively longer search latency, relative to other produced words (as indexed by interitem response time, or IRT), and the second item in a cluster is associated with a relatively shorter search latency. Such comparisons will provide more direct tests of existing theoretical claims, and help to elucidate differences in behavior across associative and category-based models.

A primary trade-off associated with generating higher-order representations from probabilistic word-topic loadings is that resulting categories sometimes show little similarity to human-generated categories. Development of a larger training corpus would likely improve measures of semantic similarity and yield more coherent topics. However, even if unlimited texts were available, it is unlikely that such an approach would fully capture the richness of human semantic associations, which also reflect physical attributes (e.g., shape, color). Such similarities may contribute to other classes of abstract representations that guide selection, particularly when categorical prompts are abstract (e.g., things that are round, things that are soft). Because text-based models are trained in a relatively impoverished environment, they are unlikely to replicate the full repertoire of representations guiding selection in VF, and may be particularly ill-suited to model processes supporting selection in more abstract prompts.

Future tests of these models could explore how learning and maintenance of higher-order representations independently support production, by determining how adjustments to the size and composition of the training corpus selectively influence production. This may yield new insights into how executive processes guide solution-finding in open-ended tasks: for example, are children, who have less robust semantic networks than adults, more likely to use executive strategies (e.g., switching from search guided by one abstract dimension to another) to generate words? Subsequent iterations of this model could be used to test other theoretically-relevant questions, including whether reducing the capacity of the model to support abstract representations results in intact associative clustering and impaired switching replicating previous behavioral findings. Likewise, if maintenance of a given VF prompt (e.g., animals) supports selective activation of related exemplars, weakening maintenance of that prompt should induce ‘stickiness’, such that exemplars that were appropriate responses for a previous VF prompt are selected for the current prompt. This behavior is sometimes observed in children who complete sequential VF tasks (e.g., responding 'sandwich' to an 'animal' prompt, some minutes after completing a 'food' prompt). We have thus far treated maintenance as a constant (with information from previously-activated topic representations decaying constantly across time), but future models could investigate this as a free parameter that varies across subjects, potentially contributing to individual differences in clustering and switching.

Finally, future extensions of this model could also be used to explore new, testable hypotheses. One possibility is that individuals may be less likely to draw on executive processes (e.g., activating and maintaining an abstract representation of typical word context to guide search through semantic space) when they are responding to a VF prompt that is well-organized in semantic memory along a small set of contextual dimensions. Like extant associative computational models, we have evaluated model performance on the basis of a single prompt ('animals') that may make relatively low demands on EF relative to more abstract prompts, particularly in adults (van der Elst, van Boxtel, van Breukelen, & Jolles, 2006). Although the idea that fluency prompts may vary in their executive demands has not yet been systematically tested, this hypothesis could easily be modeled via the proposed model. For example, the taxonomic organization of animals may make it easier to draw on largely associative processes during production, such that adults require relatively few switches across distinct contextual dimensions to produce many words. By contrast, other prompts may activate more distributed contextual representations, such that effective production requires search across a number of topics (e.g., 'round objects'). Such accounts would predict that prompts associated with a broader array of possible search dimensions should be associated with longer between-cluster search latencies than prompts where search dimensions are more restricted. It is not clear that an optimal foraging account, where search strategies are informed by environmental frequencies, would yield a similar prediction. For example, to support a similar finding, the frequency with which round objects are encountered would have to vary substantially from the frequency with which animals are encountered.

Chapter 5

General Discussion

How do children improve in their ability to achieve goals endogenously, without external instructions and reminders? Young children who can successfully respond to external cues signaling what they should do often fail to reach goals when they have to generate their own strategies and reminders. The ability to engage EFs in self-directed contexts emerges gradually across development, raising the possibility that endogenous forms of control may be more sensitive to environmental experiences than externally-driven forms of EF. This dissertation presents two studies examining how children's early exposure to adult-structured environments relate to their developing ability to engage executive functions in self-directed contexts, as measured via indices of self-directed switching in semantic verbal fluency, and a computational model exploring how the development of higher-order abstract representations of semantic structure could facilitate word retrieval in VF.

Chapters 2 and 3 are the first studies to explore how the structure of children's daily experiences relate to developing EF in early childhood. In Chapter 2, children's daily time in less-structured activities at age 6 and 7 predicted their verbal fluency switching performance after controlling for potential confounding factors, including household income and vocabulary. Children's time in structured activities showed a trend relationship in the opposite direction, such that more time in structured activities predicted worse verbal fluency switching ability. These effects did not extend to two other measures where cues and reminders signaling when children should engage control were provided by the experimenter.

Chapter 3 tested longitudinal associations between children's early time in structured activities and structured household environments and their later verbal fluency performance within a genetically informative twin sample. In independent analyses, children's structured activity participation and household ratings of structure at age 4 predicted children's verbal fluency performance at age 7, controlling for earlier verbal fluency performance and concurrent structured time/household structure. Additionally, children who showed better early VF switching performance at age 4 went on to spend more time in structured activities at age 7. This opposing relationship could explain why links between VF switching ability and time in structured activities were somewhat tenuous in Chapter 2. Follow-up genetic analyses relating children's early time in structured activities to their later VF indicated that this relationship was mediated by nonshared environmental factors, consistent with the hypothesis that that early structured time may causally affect later VF performance. The corresponding relationship between early VF performance and later structured activity participation was mediated by shared environmental factors, consistent with the interpretation that links between early VF and later structure reflect a passive gene-environment correlation (e.g., parents pass genes contributing to higher EF to their children, and are also more likely to enroll their children in structured activities).

Chapter 4 explores the feasibility of developing a computational model informed by real-world child texts. Although Studies 1 and 2 are both consistent with the possible influence of environmental structure on self-directed processes within verbal fluency, some debate remains about whether observed patterns of clustering and switching in VF are in fact epiphenomenal. These accounts have argued that word production patterns instead reflect automatic, associative memory processes. We therefore developed a simple model exploring how goal-oriented maintenance of abstract representations might improve search for new exemplars. The resulting model showed a better fit to child VF data than a comparison, purely associative model, demonstrating how such processes could facilitate word production, and challenging accounts suggesting that word production can be parsimoniously explained via simple associative processes. Future iterations of this model could explore how different training environments influence the development of representations that constrain and guide search (e.g., via exposure to different texts), and whether self-directed monitoring for performance within clusters improves performance. Such extensions would complement investigations into links between environmental structure and VF by exploring how self-directed processes contribute to word production.

5.1 Future Directions

5.1.1 Direct tests of directionality via experimental manipulations

Although findings from Chapter 3 are consistent with a causal role for structured time on developing semantic EF, and suggest that observed links are driven by shared environmental rather than genetic factors, definitive tests of causal claims demand experimental manipulation. Thus, an important direction for future work is to develop experimental designs that manipulate children's exposure to environmental structure. This work could also help to distinguish whether observed relationships between structured activity participation and emerging VF performance are driven by opportunity costs to time in less-structured activities.

Developing such an intervention will pose challenges. If effects of structure accrue over time, short-term laboratory studies are unlikely to replicate the conditions necessary to produce predicted outcomes on self-directed EF. Thus, a key component of any intervention will be the development of a well controlled experimental intervention wherein exposure to environmental structure is longitudinal, amenable to external monitoring, and consistent across both intervention and control conditions. One option is to develop a virtual environment or game that allows for direct manipulation of child activities. In such an environment, more-structured intervention conditions could provide children with detailed guidance about how they should achieve given goals, reducing the necessity of planning relevant actions in the absence of external supports. In less-structured intervention conditions, children would be given more latitude about how to reach task-relevant goals, facilitating their self-directed choices and actions.

5.1.2 Unique Demands in Semantic Verbal Fluency?

In this dissertation, I have investigated task dynamics using computational approaches that explicitly represent theorized distinctions between internally- and externally-driven cognitive processes. An alternative, complementary approach is to investigate overlapping and unique sources of variance associated with switching-specific VF measures within a broader executive model that dissociates demands in VF from other EF abilities. Extant tests of links between EFs and VF performance have shown that individuals with higher working memory capacity generate more words in verbal fluency tasks, on average (Fisk & Sharp, 2004; Fournier-Vicente, Larigauderie, & Gaonac'h, 2008; Hedden, Lautenschlager, & Park, 2005; Kane & Engle, 2000; Unsworth & Engle, 2007), and show greater impairments in production under high-load conditions when a second task is introduced (Rosen & Engle, 1997). Additionally, correspondence between total word production in VF and complex WM has also been established via latent factor models (Unsworth et al., 2011). However, extant explorations of shared variance across VF performance and other EFs have focused on total word production, rather than switching-specific indices.

As a complementary approach to task-based analyses, computational modeling of VF task dynamics could also be extended to explore how executive processes guide production in semantic VF. For example, it is unclear from current accounts whether executive processes are instrumental in (a) monitoring within-cluster performance (e.g., rate of return), (b) maintaining current subcategorical cues, (c) inhibiting return to previously generated clusters or exemplars, or (d) some combination of these processes.

A second question is whether EFs are more important to some forms of search than others. As discussed in Chapter 4, different task prompts may place different demands on executive processes in VF. It may be that individuals are more likely to use simple, frequency-based cues in well-integrated networks (e.g., foods, personal acquaintances), where there are many connections between high-frequency options. By contrast, when responding to VF prompts that require search over networks with low clustering coefficients (where fewer links connect one network node to another), and/or networks with fewer high-frequency responses, participants may be more likely to fall back on alternative generation strategies (e.g., 'where have I encountered things that are round?'; 'what round things do I use in the morning/at the office/during cooking?'). Such a model would also predict differences in search processes across development. Children with less robust semantic networks might be forced to engage alternative strategies more frequently than adults with more mature networks. Consistent with this prediction, some studies of ideational fluency, or unusual uses tasks, have suggested that strategy generation and selection (e.g., coming up with unusual uses for a common object) are a key process contributing to individual differences in performance in fluency tasks (Nusbaum & Silvia, 2011; Silvia, Beaty, & Nusbaum, 2013).

5.1.3 Development of Additional Measures Testing EF in Self-directed Contexts

These studies test how environmental structure and opportunities for self-directed behavior relate to performance in a single task, semantic verbal fluency. To develop more nuanced, mechanistic accounts of how environments might shape the processes underlying self-directed behavior, it would be helpful to test whether environmental structure and opportunities for self-direction relate to performance on other tasks drawing on endogenous forms of control, as well as behaviors that may draw on other abilities that facilitate self-direction, including planning ability, curiosity, and information-seeking.

Although several tasks have been developed to assess how well individuals can carry out goal-directed behaviors in complex environments with competing task demands, most were designed to maximize clinical relevance (e.g., ability to discriminate poor real-world planning and decision-making), rather than support focused tests of endogenous control mechanisms. As such, many self-directed tasks tax both executive and non-executive abilities, making it difficult to isolate the source or sources of observed deficits. Real-world endogenous control tasks, for example, may place greater demands on non-executive cognitive processes than standard laboratory-based measures of executive function (e.g. Shallice & Burgess, 1991; Mackinlay, Charman, & Karmiloff-Smith, 2006). Thus, a clear direction for future work is to explore commonalities across tasks taxing self-directed control processes, and to test whether these tasks place unique demands on cognitive processes that do not contribute to performance in externally-driven tasks.

5.1.4 Conclusion

Historical shifts in child time use have inspired broad, societal questions about how environmental structure might affect children's self-directed behavior. Modern children may have fewer opportunities to engage in child-directed activities (determining on their own what they will do, and how they will do it) than their predecessors. Growing societal emphasis on early skill acquisition and heightened parental vigilance have contributed to reductions in the time children spend in unsupervised activities, including independent travel and play (Clements, Sarama, & Wolfe, 2011). Activities such as outdoor play have increasingly been replaced by media activities, including video game play, computing, and television watching (Bavelier, Green, & Dye, 2010; Johnson, 2010; Vandewater et al., 2007), which typically offer fewer opportunities for child decision-making than other forms of leisure. Children also spend more time in structured, adult-led activities (Hofferth & Sandberg, 2001a; Larson, 2001). Longitudinal studies of child time use and developing EF offer one way of testing how shifts in time use might affect children's goal-directed behavior, an approach we have adopted in the present dissertation. Future investigations would benefit from exploration of relationships across societies and cultures,using controlled, cross-sectional comparisons. Better understanding of the sensitivity of self-directed EF to specific childhood experiences may ultimately inform, extend, and improve extant EF-focused interventions across the lifespan.

References

Aaltonen, S., Ortega-Alonso, A., Kujala, U. M., & Kaprio, J. (2010, October). A Longitudinal Study on Genetic and Environmental Influences on Leisure Time Physical Activity in the Finnish Twin Cohort. Twin Research and Human Genetics, 13(5), 475–481. doi: 10.1375/twin.13.5.475

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015, July). Random walks on semantic networks can resemble optimal foraging. Psychological Review, 122(3), 558–569. doi: 10.1037/a0038693

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001, September). Qualitative Analysis of Verbal Fluency Output: Review and Comparison of Several Scoring Methods. Assessment, 8(3), 323–338. doi: 10.1177/107319110100800308

Achenbach, T. M., & Edelbrock, C. S. (1983). Manual for the child behavior checklist and revised child behavior profile. Department of Psychiatry of the University of Vermont.

Agnew, R., & Petersen, D. M. (1989, October). Leisure and Delinquency. Social Problems, 36(4), 332–350. doi: 10.2307/800819

Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003, January). Ecological validity of a simplified version of the multiple errands shopping test. Journal of the International Neuropsychological Society, 9(01), 31–44. doi: 10.1017/S1355617703910046

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005, August). The Influence of the Parents' Educational Level on the Development of Executive Functions. Developmental Neuropsychology, 28(1), 539–560. doi: 10.1207/s15326942dn2801_5

Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in children's daily lives predicts self-directed executive functioning. Developmental Psychology, 5, 593. doi: 10.3389/fpsyg.2014.00593

Barsalou, L. W. (1983). Ad hoc categories. Memory & Cognition, 11(3), 211–227. doi: 10.3758/BF03196968

Batmanghelich, K., Saeedi, A., Narasimhan, K., & Gershman, S. (2016, April). Nonparametric Spherical Topic Modeling with Word Embeddings. arXiv:1604.00126 [cs, stat]. (arXiv: 1604.00126)

Bavelier, D., Green, C. S., & Dye, M. W. (2010, September). Children, Wired: For Better and for Worse. Neuron, 67(5), 692–701. doi: 10.1016/j.neuron.2010.08.035

Bechtoldt, H. P., Benton, A. L., & Fogel, M. L. (1962). An application of factor analysis in neuropsychology. The Psychological Record, 12(2), 147–156.

Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011, May). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: a controlled, randomized study: Fluid intelligence gains after training non-verbal reasoning. Developmental Science, 14(3), 591–601. doi: 10.1111/j.1467-7687.2010.01022.x

Berns, G. S., Blaine, K., Prietula, M. J., & Pye, B. E. (2013, December). Short- and Long-Term Effects of a Novel on Connectivity in the Brain. Brain Connectivity, 3(6), 590–600. doi: 10.1089/brain.2013.0166

Berry, D., & Willoughby, M. T. (2017, July). On the Practical Interpretability of Cross-Lagged Panel Models: Rethinking a Developmental Workhorse. Child Development, 88(4), 1186–1206. doi: 10.1111/cdev.12660

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011, August). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. Learning and Individual Differences, 21(4), 327–336. doi: 10.1016/j.lindif.2011.01.007

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008, June). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. Development and Psychopathology, 20(03). doi: 10.1017/S0954579408000394

Blair, C., Granger, D., & Peters Razza, R. (2005, May). Cortisol Reactivity Is Positively Related to Executive Function in Preschool Children Attending Head Start. Child Development, 76(3), 554–567. doi: 10.1111/j.1467-8624.2005.00863.x

Blair, C., & Razza, R. P. (2007, March). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. Child Development, 78(2), 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Bohnert, A. M., Kane, P., & Garber, J. (2008, February). Organized Activity Participation and Internalizing and Externalizing Symptoms: Reciprocal Relations during Adolescence. Journal of Youth and Adolescence, 37(2), 239–250. doi: 10.1007/s10964-007-9195-1

Bollen, K. A., & Diamantopoulos, A. (2017, September). In defense of causal-formative indicators: A minority report. Psychological Methods, 22(3), 581–596. doi: http://dx.doi.org/10.1037/met0000056

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. Variation in working memory, 76–106.

Brocki, K. C., & Bohlin, G. (2004, October). Executive Functions in Children Aged 6 to 13: A Dimensional and Developmental Study. Developmental Neuropsychology, 26(2), 571–593. doi: 10.1207/s15326942dn2602_3

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012, September). A unitary executive function predicts intelligence in children. Intelligence, 40(5), 458–469. doi: 10.1016/j.intell.2012.05.006

Burgess, P. W., Alderman, N., Volle, E., Benoit, R. G., & Gilbert, S. J. (2009, October). Mesulam's frontal lobe mystery re-examined. Restorative Neurology & Neuroscience, 27(5), 493–506. doi: 10.3233/RNN-2009-0511

Cadogan, J. W., & Lee, N. (2013, February). Improper use of endogenous formative variables. Journal of Business Research, 66(2), 233–241. doi: 10.1016/j.jbusres.2012.08.006

Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., & Morrison, F. J. (2012, July). Fine Motor Skills and Executive Function Both Contribute to Kindergarten Achievement: Fine Motor and Kindergarten Achievement. Child Development, 83(4), 1229–1244. doi: 10.1111/j.1467-8624.2012.01768.x

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009, April). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. Proceedings of the National Academy of Sciences, 106(14), 5529–5533. doi: 10.1073/pnas.0810002106

Clements, D. H., Sarama, J., & Wolfe, C. B. (2011). Tools for early assessment in math (TEAM). Teachers Guide. McGraw Hill Education Series.

Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison, F. J. (2010, October). First graders' literacy and self-regulation gains: The effect of individualizing student instruction. Journal of School Psychology, 48(5), 433–455. doi: 10.1016/j.jsp.2010.06.003

Curran, P. J., & Bauer, D. J. (2011). The Disaggregation of Within-Person and Between-Person Effects in Longitudinal Models of Change. Annual Review of Psychology, 62(1), 583–619. doi: 10.1146/annurev.psych.093008.100356

Curran, P. J., Howard, A. L., Bainter, S., Lane, S. T., & McGinley, J. S. (2014, October). The Separation of Between-person and Within-person Components of Individual Change Over Time: A Latent Curve Model with Structured Residuals. Journal of consulting and clinical psychology, 82(5), 879–894. doi: 10.1037/a0035297

Daneman, M., & Merikle, P. M. (1996, December). Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review, 3(4), 422–433. doi: 10.3758/BF03214546

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009, May). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. Journal of the International Neuropsychological Society, 15(03), 331. doi: 10.1017/S1355617709090481

Diamond, A. (2012, October). Activities and Programs That Improve Children's Executive Functions. Current Directions in Psychological Science, 21(5), 335–341. doi: 10.1177/0963721412453722

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007, November). Preschool Program Improves Cognitive Control. Science, 318(5855), 1387–1388. doi: 10.1126/science.1151148

Diamond, A., & Lee, K. (2011, August). Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old. Science, 333(6045), 959–964. doi: 10.1126/science.1204529

D'Ostilio, K., & Garraux, G. (2012, January). Dissociation between unconscious motor response facilitation and conflict in medial frontal areas: Unconscious motor facilitation and conflict. European Journal of Neuroscience, 35(2), 332–340. doi: 10.1111/j.1460-9568.2011.07941.x

Eccles, J. S., & Barber, B. L. (1999, January). Student Council, Volunteering, Basketball, or Marching Band: What Kind of Extracurricular Involvement Matters? Journal of Adolescent Research, 14(1), 10–43. doi: 10.1177/0743558499141003

Engelhardt, L. E., Briley, D. A., Mann, F. D., Harden, K. P., & Tucker-Drob, E. M. (2015, August). Genes Unite Executive Functions in Childhood. Psychological Science, 26(8), 1151–1163. doi: 10.1177/0956797615577209

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General, 128(3), 309–331. doi: 10.1037/0096-3445.128.3.309

Eriksen, C. W., & Schultz, D. W. (1979, July). Information processing in visual search: A continuous flow conception and experimental results. Perception & Psychophysics, 25(4), 249–263. doi: 10.3758/BF03198804

Eslinger, P. J., & Damasio, A. R. (1985, December). Severe disturbance of higher cognition after bilateral frontal lobe ablation Patient EVR. Neurology, 35(12), 1731–1731. doi: 10.1212/WNL.35.12.1731

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., ... Hurt, H. (2006, September). Childhood poverty: Specific associations with neurocognitive development. Brain Research, 1110(1), 166–174. doi: 10.1016/j.brainres.2006.06.072

Fisher, A., Boyle, J. M., Paton, J. Y., Tomporowski, P., Watson, C., McColl, J. H., & Reilly, J. J. (2011, December). Effects of a physical education intervention on cognitive function in young children: randomized controlled pilot study. BMC Pediatrics, 11(1). doi: 10.1186/1471-2431-11-97

Fisher, A., Jaarsveld, C. H. M. v., Llewellyn, C. H., & Wardle, J. (2010, April). Environmental Influences on Children's Physical Activity: Quantitative Estimates Using a Twin Design. PLOS ONE, 5(4), e10110. doi: 10.1371/journal.pone.0010110

Fisk, J. E., & Sharp, C. A. (2004, October). Age-Related Impairment in Executive Functioning: Updating, Inhibition, Shifting, and Access. Journal of Clinical and Experimental Neuropsychology, 26(7), 874–890. doi: 10.1080/13803390490510680

Fletcher, A. C., Nickerson, P., & Wright, K. L. (2003, November). Structured leisure activities in middle childhood: Links to well-being. Journal of Community Psychology, 31(6), 641–659. doi: 10.1002/jcop.10075

Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008, September). More dissociations and interactions within central executive functioning: A comprehensive latent-variable analysis. Acta Psychologica, 129(1), 32–48. doi: 10.1016/j.actpsy.2008.04.004

Franks, P. W., Ravussin, E., Hanson, R. L., Harper, I. T., Allison, D. B., Knowler, W. C., ... Salbe, A. D. (2005, October). Habitual physical activity in children: the role of genes and the environment. The American Journal of Clinical Nutrition, 82(4), 901–908.

Fredricks, J. A., & Eccles, J. S. (2008, October). Participation in Extracurricular Activities in the Middle School Years: Are There Developmental Benefits for African American and European American Youth? Journal of Youth and Adolescence, 37(9), 1029–1043. doi: 10.1007/s10964-008-9309-4

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. Journal of Experimental Psychology: General, 137(2), 201.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. Developmental Psychology, 40(2), 177–190. doi: 10.1037/0012-1649.40.2.177

Hauser, R. M., & Featherman, D. L. (1977). The process of stratification: Trends and analyses. New York: Academic Press.

Hayiou-Thomas, M. E., Kovas, Y., Harlaar, N., Plomin, R., Bishop, D. V. M., & Dale, P. S. (2006, May). Common aetiology for diverse language skills in 41/2-year-old twins. Journal of Child Language, 33(2), 339–368. doi: 10.1017/S0305000906007331

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010, February). The rate of return to the HighScope Perry Preschool Program. Journal of Public Economics, 94(1-2), 114–128. doi: 10.1016/j.jpubeco.2009.11.001

Hedden, T., Lautenschlager, G., & Park, D. C. (2005, January). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. The Quarterly Journal of Experimental Psychology Section A, 58(1), 169–190. doi: 10.1080/02724980443000179

Hill, F., Bordes, A., Chopra, S., & Weston, J. (2015, November). The Goldilocks Principle: Reading Children's Books with Explicit Memory Representations. arXiv:1511.02301 [cs]. (arXiv: 1511.02301)

Hills, T. T., Jones, M. N., & Todd, P. M. (2012, April). Optimal foraging in semantic memory. Psychological Review, 119(2), 431–440. doi: 10.1037/a0027373

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. Neuropsychologia, 44(12), 2547–2557. doi: 10.1016/j.neuropsychologia.2006.03.035

Hofferth, S. L., & Sandberg, J. F. (2001a, January). Changes in American children's time, 1981–1997. Advances in Life Course Research, 6, 193–229. doi: 10.1016/S1040-2608(01)80011-3

Hofferth, S. L., & Sandberg, J. F. (2001b, May). How American Children Spend Their Time. Journal of Marriage and Family, 63(2), 295–308. doi: 10.1111/j.1741-3737.2001.00295.x

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009, July). Adaptive training leads to sustained enhancement of poor working memory in children. Developmental Science, 12(4), F9–F15. doi: 10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010, September). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. Applied Cognitive Psychology, 24(6), 827–836. doi: 10.1002/acp.1589

Hu, L.-t., & Bentler, P. M. (1999, January). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. doi: 10.1080/10705519909540118

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009, December). Tracking Executive Function Across the Transition to School: A Latent Variable Approach. Developmental Neuropsychology, 35(1), 20–36. doi: 10.1080/87565640903325691

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006, January). Age-related change in executive function: Developmental trends and a latent variable analysis. Neuropsychologia, 44(11), 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Hurks, P. P. M. (2013, October). Administering Design Fluency Tests in School-aged Children: Analyses of Design Productivity over Time, Clustering, and Switching. The Clinical Neuropsychologist, 27(7), 1131–1149. doi: 10.1080/13854046.2013.821170

Hurks, P. P. M., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010, July). Developmental Changes in Semantic Verbal Fluency: Analyses of Word Productivity as a Function of Time, Clustering, and Switching. Child Neuropsychology, 16(4), 366–387. doi: 10.1080/09297041003671184

Johnson, G. M. (2010, October). Young children's Internet use at home and school: Patterns and profiles. Journal of Early Childhood Research, 8(3), 282–293. doi: 10.1177/1476718X10379783

Kalkut, E. L., Han, S. D., Lansing, A. E., Holdnack, J. A., & Delis, D. C. (2009, September). Development of Set-Shifting Ability from Late Childhood Through Early Adulthood. Archives of Clinical Neuropsychology, 24(6), 565–574. doi: 10.1093/arclin/acp048

Kane, M. J., & Engle, R. W. (2000, March). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(2), 336–358. doi: 10.1037/0278-7393.26.2.336

Kavé, G., Kigel, S., & Kochva, R. (2008, March). Switching and clustering in verbal fluency tasks throughout childhood. Journal of Clinical and Experimental Neuropsychology, 30(3), 349–359. doi: 10.1080/13803390701416197

Kavé, G., Shalmon, M., & Knafo, A. (2013, January). Environmental contributions to preschoolers' semantic fluency. Developmental Science, 16(1), 124–135. doi: 10.1111/desc.12010

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001, August). Differential Development of Attention and Executive Functions in 3- to 12-Year-Old Finnish Children. Developmental Neuropsychology, 20(1), 407–428. doi: 10.1207/S15326942DN2001_6

Knight, C., Alderman, N., & Burgess, P. W. (2002, June). Development of a simplified version of the multiple errands test for use in hospital settings. Neuropsychological Rehabilitation, 12(3), 231–255. doi: 10.1080/09602010244000039

Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2017). Behavioral Genetics (7th ed.). New York: Worth Publishers.

Koren, R., Kofman, O., & Berger, A. (2005, December). Analysis of word clustering in verbal fluency of school-aged children. Archives of Clinical Neuropsychology, 20(8), 1087–1104. doi: 10.1016/j.acn.2005.06.012

Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V. M., & Plomin, R. (2005, May). Genetic Influences in Different Aspects of Language Development: The Etiology of Language Skills in 4.5-Year-Old Twins. Child Development, 76(3), 632–651. doi: 10.1111/j.1467-8624.2005.00868.x

Lanting, S., Haugrud, N., & Crossley, M. (2009, March). The effect of age and sex on clustering and switching during speeded verbal fluency tasks. Journal of the International Neuropsychological Society, 15(02), 196–204. doi: 10.1017/S1355617709090237

Larson, R. W. (2001, October). How U.S. Children and Adolescents Spend Time: What It Does (and Doesn't) Tell Us About Their Development. Current Directions in Psychological Science, 10(5), 160–164. doi: 10.1111/1467-8721.00139

Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: Work, play, and developmental opportunities. Psychological Bulletin, 125(6), 701–736. doi: 10.1037/0033-2909.125.6.701

Lau, J. H., Newman, D., & Baldwin, T. (2014). Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality. In EACL (pp. 530–539).

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. Journal of Experimental Psychology: Human Perception and Performance, 21(3), 451–468. doi: 10.1037/0096-1523.21.3.451

Lillard, A. S. (2012, June). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. Journal of School Psychology, 50(3), 379–401. doi: 10.1016/j.jsp.2012.01.001

Lillard, A. S., & Else-Quest, N. (2006, September). The Early Years: Evaluating Montessori Education. Science, 313(5795), 1893–1894. doi: 10.1126/science.1132362

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013, January). The impact of pretend play on children's development: A review of the evidence. Psychological Bulletin, 139(1), 1–34. doi: http://0-dx.doi.org.libraries.colorado.edu/10.1037/a0029321

Luo, Y. L. L., Haworth, C. M. A., & Plomin, R. (2010, October). A Novel Approach to Genetic and Environmental Analysis of Cross-Lagged Associations Over Time: The Cross-Lagged Relationship Between Self-Perceived Abilities and School Achievement is Mediated by Genes as Well as the Environment. Twin research and human genetics : the official journal of the International Society for Twin Studies, 13(5). doi: 10.1375/twin.13.5.426

Mackey, A. P., Miller Singley, A. T., & Bunge, S. A. (2013, March). Intensive Reasoning Training Alters Patterns of Brain Connectivity at Rest. Journal of Neuroscience, 33(11), 4796–4803. doi: 10.1523/JNEUROSCI.4141-12.2013

Mackinlay, R., Charman, T., & Karmiloff-Smith, A. (2006, June). High functioning children with autism spectrum disorder: A novel test of multitasking. Brain and Cognition, 61(1), 14–24. doi: 10.1016/j.bandc.2005.12.006

Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., . . .

Denckla, M. B. (2002, December). Effects of IQ on Executive Function Measures in Children with ADHD. Child Neuropsychology (Neuropsychology, Development and Cognition: Section C), 8(1), 52–65. doi: 10.1076/chin.8.1.52.8719

Mahoney, J. L., & Stattin, H. (2000, April). Leisure activities and adolescent antisocial behavior: The role of structure and social context. Journal of Adolescence, 23(2), 113–127. doi: 10.1006/jado.2000.0302

Matute, E., Rosselli, M., Ardila, A., & Morales, G. (2004, October). Verbal and Nonverbal Fluency in Spanish-Speaking Children. Developmental Neuropsychology, 26(2), 647–660. doi: 10.1207/s15326942dn2602_7

McAuley, T., Christ, S. E., & White, D. A. (2011, July). Mapping the Development of Response Inhibition in Young Children Using a Modified Day-Night Task. Developmental Neuropsychology, 36(5), 539–551. doi: 10.1080/87565641.2010.549871

McDermott, J. M., Pérez-Edgar, K., & Fox, N. A. (2007, February). Variations of the flanker paradigm: Assessing selective attention in young children. Behavior Research Methods, 39(1), 62–70. doi: 10.3758/BF03192844

Meeks, C. B., & Mauldin, T. (1990). Children's time in structured and unstructured leisure activities. Lifestyles Family and Economic Issues, 11(3), 257–281. doi: 10.1007/BF00987003

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49(2), 270–291. doi: 10.1037/a0028228

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, January). Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781 [cs]. (arXiv: 1301.3781)

Miller, G. (2012, May). The Smartphone Psychology Manifesto. Perspectives on Psychological Science, 7(3), 221–237. doi: 10.1177/1745691612441215

Miller, M. R., Müller, U., Giesbrecht, G. F., Carpendale, J. I., & Kerns, K. A. (2013, October). The contribution of executive function and social understanding to preschoolers' letter and math skills. Cognitive Development, 28(4), 331–349. doi: 10.1016/j.cogdev.2012.10.005

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000, August). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. Cognitive Psychology, 41(1), 49–100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... Caspi, A. (2011, February). A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the National Academy of Sciences, 108(7), 2693–2698. doi: 10.1073/pnas.1010076108

Montez, P., Thompson, G., & Kello, C. T. (2015, November). The role of semantic clustering in optimal memory foraging. Cognitive Science, 39(8), 1925–1939. doi: http://0-dx.doi.org.libraries.colorado.edu/10.1111/cogs.12249

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012, April). Developing Cognitive Control Three Key Transitions. Current Directions in Psychological Science, 21(2), 71–77. doi: 10.1177/0963721412436807

Muthén, L. K., & Muthén, B. O. (1998). Mplus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nichols, R. C., & Bilbro, W. C. (1966). The diagnosis of twin zygosity. Acta Genetica et Statistica Medica, 16, 265–275.

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007, July). Socioeconomic gradients predict individual differences in neurocognitive abilities. Developmental Science, 10(4), 464–480. doi: 10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., & Farah, M. J. (2005, January). Neurocognitive correlates of socioeconomic status in kindergarten children. Developmental Science, 8(1), 74–87. doi: 10.1111/j.1467-7687.2005.00394.x

Nusbaum, E. C., & Silvia, P. J. (2011, January). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. Intelligence, 39(1), 36–45. doi: 10.1016/j.intell.2010.11.002

Osgood, D. W., Anderson, A. L., & Shaffer, J. N. (2005). Unstructured leisure in the after-school hours. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), Organized Activities as Contexts of Development: Extracurricular Activities, After-school and Community Programs (pp. 45–64). Mahwah, NJ: Erlbaum.

Owens, S. F., Rijsdijk, F., Picchioni, M. M., Stahl, D., Nenadic, I., Murray, R. M., & Toulopoulou, T. (2011, April). Genetic overlap between schizophrenia and selective components of executive function. Schizophrenia Research, 127(1–3), 181–187. doi: 10.1016/j.schres.2010.10.010

Paquet, L. (2001, June). Eliminating flanker effects and negative priming in the flankers task: Evidence for early selection. Psychonomic Bulletin & Review, 8(2), 301–306. doi: 10.3758/BF03196165

Raver, C. C., Blair, C., Willoughby, M., & Family Life Project Key Investigators. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. Developmental Psychology, 49(2), 292–304. doi: 10.1037/a0028343

Reitan, R. M., & Wolfson, D. (1994, September). A selective and critical review of neuropsychological deficits and the frontal lobes. Neuropsychology Review, 4(3), 161–198. doi: 10.1007/BF01874891

Rhea, S.-A., Gross, A. A., Haberstick, B. C., & Corley, R. P. (2006, December). Colorado Twin Registry. Twin Research and Human Genetics, 9(6), 941–949. doi: 10.1375/twin.9.6.941

Ridderinkhof, K. R., & van der Molen, M. W. (1995, August). A Psychophysiological Analysis of Developmental Differences in the Ability to Resist Interference. Child Development, 66(4), 1040. doi: 10.2307/1131797

Riva, D., Nichelli, F., & Devoti, M. (2000, February). Developmental Aspects of Verbal Fluency and Confrontation Naming in Children. Brain and Language, 71(2), 267–284. doi: 10.1006/brln.1999.2166

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. Journal of experimental psychology. General, 126(3), 211–227.

Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Roebers, C. M. (2012, July). Improving executive functions in 5- and 6-year-olds: Evaluation of a small group intervention in prekindergarten and kindergarten children: Improving executive functions in 5- and 6-year-olds. Infant and Child Development, 21(4), 411–429. doi: 10.1002/icd.752

Rueda, M., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004, January). Development of attentional networks in childhood. Neuropsychologia, 42(8), 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Checa, P., & Cómbita, L. M. (2012, February). Enhanced efficiency of the executive attention network after training in preschool children: Immediate changes and effects after two months. Developmental Cognitive Neuroscience, 2, S192–S204. doi: 10.1016/j.dcn.2011.09.004

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005, October). From The Cover: Training, maturation, and genetic influences on the development of executive attention. Proceedings of the National Academy of Sciences, 102(41), 14931–14936. doi: 10.1073/pnas.0506897102

Samuels, W. E., Tournaki, N., Blackman, S., & Zilinski, C. (2016, September). Executive functioning predicts academic achievement in middle school: A four-year longitudinal study. The Journal of Educational Research, 109(5), 478–490. doi: 10.1080/00220671.2014.979913

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004, April). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. Brain and Language, 89(1), 192–202. doi: 10.1016/S0093-934X(03)00367-5

Shallice, T., & Burgess, P. W. (1991, April). Deficits in Strategy Application Following Frontal Lobe Damage in Man. Brain, 114(2), 727–741. doi: 10.1093/brain/114.2.727

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013, September). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. Intelligence, 41(5), 328–340. doi: 10.1016/j.intell.2013.05.004

Snyder, H. R., & Munakata, Y. (2008, December). So many options, so little time: The roles of association and competition in underdetermined responding. Psychonomic Bulletin & Review, 15(6), 1083–1088. doi: 10.3758/PBR.15.6.1083

Snyder, H. R., & Munakata, Y. (2010, August). Becoming self-directed: Abstract representations support endogenous flexibility in children. Cognition, 116(2), 155–167. doi: 10.1016/j.cognition.2010.04.007

Snyder, H. R., & Munakata, Y. (2013, November). So many options, so little control: Abstract representations can reduce selection demands to increase children's self-directed flexibility. Journal of Experimental Child Psychology, 116(3), 659–673. doi: 10.1016/j.jecp.2013.07.010

St Clair-Thompson, H. L., & Gathercole, S. E. (2006, April). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. The Quarterly Journal of Experimental Psychology, 59(4), 745–759. doi: 10.1080/17470210500162854

Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. C. (2005). Sports Participation during Adolescence: A Shift from Environmental to Genetic Factors. Medicine & Science in Sports & Exercise, 37(4), 563–570.

Swan, G. E., & Carmelli, D. (2002, March). Evidence for Genetic Mediation of Executive Control A Study of Aging Male Twins. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57(2), P133–P143. doi: 10.1093/geronb/57.2.P133

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006, December). Hierarchical Dirichlet Processes. Journal of the American Statistical Association, 101(476), 1566–1581. doi: 10.1198/016214506000000302

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009, January). Training and transfer effects of executive functions in preschool children. Developmental Science, 12(1), 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Titz, C., & Karbach, J. (2014, November). Working memory and executive functions: effects of training on academic achievement. Psychological Research, 78(6), 852–868. doi: 10.1007/s00426-013-0537-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013, February). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? Journal of Child Psychology and Psychiatry, 54(2), 131–143. doi: 10.1111/jcpp.12001

Tranel, D., Hathaway-Nepple, J., & Anderson, S. W. (2007, March). Impaired behavior on real-world tasks following damage to the ventromedial prefrontal cortex. Journal of Clinical and Experimental Neuropsychology, 29(3), 319–332. doi: 10.1080/13803390600701376

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology, 11(1), 138–146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998, June). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. Neuropsychologia, 36(6), 499–504. doi: 10.1016/S0028-3932(97)00152-8

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, a. M. (1998, March). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. Journal of the International Neuropsychological Society, 4(02), 137–143.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. Psychological review, 114(1), 104–132.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011, February). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. The Quarterly Journal of Experimental Psychology, 64(3), 447–466. doi: 10.1080/17470218.2010.505292

van der Elst, W., van Boxtel, M. P. J., van Breukelen, G. J. P., & Jolles, J. (2006, January). Normative data for the Animal, Profession and Letter 'M' Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. Journal of the International Neuropsychological Society, 12(1), 80–89. doi: 10.1017/S1355617706060115

Vandewater, E. A., Rideout, V. J., Wartella, E. A., Huang, X., Lee, J. H., & Shim, M.-s. (2007, May). Digital Childhood: Electronic Media and Technology Use Among Infants, Toddlers, and Preschoolers. PEDIATRICS, 119(5), e1006–e1015. doi: 10.1542/peds.2006-1804

Wechsler, D. (1974). Manual for the Wechsler intelligence scale for children, revised. Psychological Corporation.

Welsh, M. C. (1991, January). Rule-guided behavior and self-monitoring on the tower of hanoi disk-transfer task. Cognitive Development, 6(1), 59–76. doi: 10.1016/0885-2014(91)90006-Y

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In Structural equation modeling: Concepts, issues, and applications (pp. 56–75). Thousand Oaks, CA, US: Sage Publications, Inc.

Wilson, B. A. (1993, September). Ecological validity of neuropsychological assessment: Do neuropsychological indexes predict performance in everyday activities? Applied and Preventive Psychology, 2(4), 209–215. doi: 10.1016/S0962-1849(05)80091-5

Zelazo, P. D., & Lyons, K. E. (2012, June). The Potential Benefits of Mindfulness Training in Early Childhood: A Developmental Social Cognitive Neuroscience Perspective. Child Development Perspectives, 6(2), 154–160. doi: 10.1111/j.1750-8606.2012.00241.x

Appendix A

Correlations across structured activity and verbal fluency measures.

Table A.1: Heterogeneous Correlation Matrix Relating Chapter 3

Study Variables (Pearson and Polychoric Estimates*)

	Variable	1	2	3	4	5	6	7	8	9	10
1	Y4 Gfactor										
2	Y7 Gfactor	.60									
3	Y16 Gfactor	.49	.66								
4	Y4 Vocabulary	.53	.37	.3							
5	Y7 Vocabulary	.44	.58	.34	.42						
6	Y16 Vocabulary	.46	.53	.58	.39	.56					
7	Parent Occupation	.20	.29	.26	.31	.24	.33				
8	Y4 CBCL Structured	.13	.08	.01	.02	.19	-.04	.09			
9	Y4 CBCL Club	.04	.11	.11	.03	.07	-.01	.01	.09		
10	Y4 CBCL Chores	.03	.04	.11	.09	-.02	.10	-.01	.13	.21	
11	Y3 FES Control	.00	-.15	-.02	-.18	-.2	-.2	-.14	.02	.12	.14
12	Y3 FES Organization	.00	-.11	-.01	-.13	-.01	-.05	-.04	.20	.10	.13
13	Y7 CBCL Chores	.06	.06	.11	.09	.03	.06	-.04	-.08	.09	.24
14	Y7 CBCL Structured	.07	.10	.12	.06	.06	.10	.15	.22	.11	.08
15	Y7 IAS Lessons	.11	.12	.15	.06	.15	.14	.20	.31	.16	.03
16	Y7 FES Control	-.13	-.12	-.05	-.06	-.16	-.10	-.17	-.08	.06	.13
17	Y7 FES Organization	.11	-.01	.03	-.01	.00	-.04	.05	.10	.17	.20
18	Y16 CBCL Club	.08	.23	.29	.16	.18	.16	.18	.10	.27	.19
19	Y16 CBCL Chores	-.01	.05	.00	.09	.05	.04	-.03	-.04	.07	.16
20	Y15 FES Control	.01	.02	.11	-.08	-.04	-.09	-.07	.06	.14	.21
21	Y15 FES Organization	.00	-.03	.10	-.21	-.08	-.04	-.02	.01	.17	-.01
22	Y4 VFsoft Total	.32	.16	.14	.27	.18	.16	.17	.07	.13	.06
23	Y4 VFsoft Switch	.19	.13	.10	.18	.20	.15	.08	.07	.08	.11
24	Y4 VFnoise Total	.35	.18	.11	.26	.20	.17	.06	.12	-.02	.02
25	Y4 VFnoise Switch	.32	.17	.18	.24	.19	.22	.03	.06	-.04	.07
26	Y4 VFround Total	.41	.31	.23	.25	.25	.30	.21	.05	-.03	.04
27	Y4 VFround Switch	.27	.21	.16	.13	.17	.16	.12	.09	-.06	-.10
28	Y7 VFround Total	.22	.45	.26	.24	.28	.20	.20	.11	.07	-.03
29	Y7 VFround Switch	.10	.29	.15	.14	.15	.18	.22	.09	.02	-.03
30	Y7 VFmetal Total	.08	.31	.14	.12	.25	.17	.12	-.05	.00	-.06
31	Y7 VFmetal Switch	.04	.21	.09	.10	.18	.12	.09	-.06	-.08	-.08
32	Y7 VFanimal Total	.30	.51	.35	.29	.38	.39	.19	.01	-.01	-.03
33	Y7 VFanimal Switch	.26	.43	.27	.23	.30	.32	.13	.01	.01	-.04
34	Y16 VFmetal Total	.15	.36	.52	.11	.25	.38	.16	.05	.04	.12
35	Y16 VFmetal Switch	.12	.33	.44	.11	.23	.32	.12	.02	.01	.13
36	Y16 VFround Total	.17	.32	.51	.15	.21	.28	.16	.02	.09	.12
37	Y16 VFround Switch	.05	.22	.37	.08	.11	.21	.10	-.03	.01	.08
38	Y4 VFsoft Cluster	.25	.20	.25	.24	.21	.27	.06	.09	.06	.04
39	Y4 VFnoise Cluster	.32	.15	.19	.23	.18	.19	.02	.06	-.03	.05
40	Y4 VFround Cluster	.38	.25	.24	.19	.29	.26	.12	.10	.02	-.01
41	Y7 VFround Cluster	.07	.15	.10	.10	.10	.13	.11	.06	.02	.02
42	Y7 VFmetal Cluster	-.04	.09	-.08	.00	.08	.06	.01	.01	-.03	-.05
43	Y7 VFanimal Cluster	.00	.05	-.09	.02	.06	.04	-.02	.08	-.01	-.04
44	Y16 VFmetal Cluster	-.01	.10	.24	.02	.08	.16	.00	-.05	.03	.10
45	Y16 VFround Cluster	-.04	.06	.09	.01	-.03	.03	.00	-.06	-.04	.03

	Variable	11	12	13	14	15	16	17	18	19	20
1	Y4 Gfactor										
2	Y7 Gfactor										
3	Y16 Gfactor										
4	Y4 Vocabulary										
5	Y7 Vocabulary										
6	Y16 Vocabulary										
7	Parent Occupation										
8	Y4 CBCL Structured										
9	Y4 CBCL Club										
10	Y4 CBCL Chores										
11	Y3 FES Control										
12	Y3 FES Organization	.45									
13	Y7 CBCL Chores	.14	.12								
14	Y7 CBCL Structured	-.04	-.08	.07							
15	Y7 IAS Lessons	-.06	-.02	.15	.36						
16	Y7 FES Control	.32	.20	.05	.05	-.04					
17	Y7 FES Organization	.35	.59	.12	-.02	.07	.26				
18	Y16 CBCL Club	.09	.05	.23	.25	.24	.05	.08			
19	Y16 CBCL Chores	.13	.21	.23	-.01	-.11	.15	.14	.15		
20	Y15 FES Control	.35	.39	.07	.00	-.01	.44	.41	.20	.04	
21	Y15 FES Organization	.18	.61	-.14	.07	.11	.19	.62	.09	-.02	.51
22	Y4 VFsoft Total	-.02	.01	-.01	.07	.08	.04	.16	.00	.03	.05
23	Y4 VFsoft Switch	.07	.00	.02	.01	.07	.04	.11	.02	.07	.04
24	Y4 VFnoise Total	-.12	-.10	-.05	.12	.07	-.02	-.09	-.01	.06	-.12
25	Y4 VFnoise Switch	-.08	-.11	.07	.12	.02	.02	-.01	.02	.03	-.09
26	Y4 VFround Total	.00	.08	-.05	-.02	.05	-.03	.12	.16	.14	.02
27	Y4 VFround Switch	-.04	-.02	-.10	-.02	.12	-.03	.06	.02	-.03	-.11
28	Y7 VFround Total	-.11	-.05	.09	.07	.08	.03	.02	.21	-.03	.10
29	Y7 VFround Switch	-.10	-.02	.09	-.06	.08	.01	.03	.03	.05	.00
30	Y7 VFmetal Total	-.04	.07	.03	-.05	-.03	-.02	-.05	.07	.01	-.04
31	Y7 VFmetal Switch	-.08	.01	.04	-.04	-.04	.00	-.03	.12	.02	-.04
32	Y7 VFanimal Total	-.15	-.11	.13	-.03	.06	.00	.05	.19	.11	-.06
33	Y7 VFanimal Switch	-.13	-.16	.10	-.04	.09	-.03	.01	.10	.02	-.09
34	Y16 VFmetal Total	-.01	-.09	.16	.10	.17	.00	-.05	.17	-.07	-.02
35	Y16 VFmetal Switch	.01	-.08	.16	.09	.12	.00	-.05	.13	-.07	-.06
36	Y16 VFround Total	.04	-.03	.06	.03	.11	.03	.04	.16	.02	-.07
37	Y16 VFround Switch	.03	.04	.01	.05	.04	.06	.03	.06	-.03	-.09
38	Y4 VFsoft Cluster	-.11	-.03	.05	.12	.22	-.07	.06	.13	.00	-.05
39	Y4 VFnoise Cluster	-.09	-.07	.07	.10	.04	-.03	.00	.06	.01	-.07
40	Y4 VFround Cluster	-.05	.07	.01	.05	.13	-.12	.13	.11	.11	.02
41	Y7 VFround Cluster	-.03	.02	.09	-.05	.08	.03	.09	.04	-.01	.03
42	Y7 VFmetal Cluster	-.02	.07	.06	-.05	-.07	.00	.01	.10	.11	-.05
43	Y7 VFanimal Cluster	-.14	-.04	.05	-.02	.09	-.03	-.06	-.09	-.16	-.10
44	Y16 VFmetal Cluster	.00	-.04	.10	.06	.03	-.03	-.05	.06	.00	-.08
45	Y16 VFround Cluster	.03	.03	-.05	.09	.01	.04	-.01	.06	-.07	-.12

	Variable	21	22	23	24	25	26	27	28	29	30
1	Y4 Gfactor										
2	Y7 Gfactor										
3	Y16 Gfactor										
4	Y4 Vocabulary										
5	Y7 Vocabulary										
6	Y16 Vocabulary										
7	Parent Occupation										
8	Y4 CBCL Structured										
9	Y4 CBCL Club										
10	Y4 CBCL Chores										
11	Y3 FES Control										
12	Y3 FES Organization										
13	Y7 CBCL Chores										
14	Y7 CBCL Structured										
15	Y7 IAS Lessons										
16	Y7 FES Control										
17	Y7 FES Organization										
18	Y16 CBCL Club										
19	Y16 CBCL Chores										
20	Y15 FES Control										
21	Y15 FES Organization										
22	Y4 VFsoft Total	.03									
23	Y4 VFsoft Switch	-.08	.73								
24	Y4 VFnoise Total	-.07	.34	.20							
25	Y4 VFnoise Switch	-.01	.20	.21	.72						
26	Y4 VFround Total	-.04	.29	.22	.28	.20					
27	Y4 VFround Switch	-.01	.17	.17	.06	.07	.51				
28	Y7 VFround Total	-.06	.09	-.04	.12	.07	.16	.07			
29	Y7 VFround Switch	.01	.04	-.05	.00	.03	.11	.03	.61		
30	Y7 VFmetal Total	.00	.03	-.05	.12	-.02	.07	.00	.37	.20	
31	Y7 VFmetal Switch	-.05	-.02	-.09	.05	-.04	.03	-.05	.27	.21	.7
32	Y7 VFanimal Total	-.22	.19	.16	.15	.17	.29	.18	.48	.34	.35
33	Y7 VFanimal Switch	-.18	.17	.16	.11	.17	.24	.16	.40	.29	.29
34	Y16 VFmetal Total	.00	-.01	.01	.18	.19	.06	-.01	.28	.16	.26
35	Y16 VFmetal Switch	-.05	.01	.05	.16	.19	.01	-.01	.27	.17	.24
36	Y16 VFround Total	-.08	.01	.01	.17	.13	.12	.01	.24	.13	.16
37	Y16 VFround Switch	-.07	-.10	-.05	.11	.05	.07	-.01	.22	.10	.12
38	Y4 VFsoft Cluster	.01	.55	.41	.26	.28	.16	.12	.08	.08	.05
39	Y4 VFnoise Cluster	-.04	.16	.20	.63	.78	.15	.04	.04	-.03	.05
40	Y4 VFround Cluster	.02	.19	.20	.20	.25	.65	.33	.16	.13	.07
41	Y7 VFround Cluster	.06	.02	-.02	.00	.01	.01	.00	.30	.62	.13
42	Y7 VFmetal Cluster	.05	-.04	-.05	.09	.04	.01	-.04	.14	.17	.46
43	Y7 VFanimal Cluster	.02	-.05	-.03	.03	.03	.08	.04	.02	.04	.09
44	Y16 VFmetal Cluster	-.04	-.03	.02	.08	.09	-.02	.01	.05	.04	.13
45	Y16 VFround Cluster	-.10	-.16	-.09	-.01	-.06	-.04	-.03	.09	.00	.03

	Variable	31	32	33	34	35	36	37	38	39	40
1	Y4 Gfactor										
2	Y7 Gfactor										
3	Y16 Gfactor										
4	Y4 Vocabulary										
5	Y7 Vocabulary										
6	Y16 Vocabulary										
7	Parent Occupation										
8	Y4 CBCL Structured										
9	Y4 CBCL Club										
10	Y4 CBCL Chores										
11	Y3 FES Control										
12	Y3 FES Organization										
13	Y7 CBCL Chores										
14	Y7 CBCL Structured										
15	Y7 IAS Lessons										
16	Y7 FES Control										
17	Y7 FES Organization										
18	Y16 CBCL Club										
19	Y16 CBCL Chores										
20	Y15 FES Control										
21	Y15 FES Organization										
22	Y4 VFsoft Total										
23	Y4 VFsoft Switch										
24	Y4 VFnoise Total										
25	Y4 VFnoise Switch										
26	Y4 VFround Total										
27	Y4 VFround Switch										
28	Y7 VFround Total										
29	Y7 VFround Switch										
30	Y7 VFmetal Total										
31	Y7 VFmetal Switch										
32	Y7 VFanimal Total	.29									
33	Y7 VFanimal Switch	.22	.88								
34	Y16 VFmetal Total	.20	.34	.29							
35	Y16 VFmetal Switch	.20	.31	.30	.87						
36	Y16 VFround Total	.13	.27	.19	.58	.49					
37	Y16 VFround Switch	.13	.15	.08	.42	.43	.79				
38	Y4 VFsoft Cluster	.10	.21	.20	.14	.12	.15	.03			
39	Y4 VFnoise Cluster	.01	.13	.15	.19	.16	.12	.04	.31		
40	Y4 VFround Cluster	.08	.25	.23	.13	.08	.17	.07	.29	.27	
41	Y7 VFround Cluster	.20	.10	.07	.10	.09	.06	.04	.06	-.05	-.03
42	Y7 VFmetal Cluster	.63	.18	.12	.06	.09	.09	.12	.05	.07	.07
43	Y7 VFanimal Cluster	.06	.17	.42	.03	.03	-.07	-.10	-.03	-.02	.09
44	Y16 VFmetal Cluster	.17	.13	.17	.34	.6	.17	.21	.10	.08	-.03
45	Y16 VFround Cluster	.06	.02	-.01	.11	.20	.31	.69	-.07	-.06	-.03

	Variable	41	42	43	44	45
1	Y4 Gfactor					
2	Y7 Gfactor					
3	Y16 Gfactor					
4	Y4 Vocabulary					
5	Y7 Vocabulary					
6	Y16 Vocabulary					
7	Parent Occupation					
8	Y4 CBCL Structured					
9	Y4 CBCL Club					
10	Y4 CBCL Chores					
11	Y3 FES Control					
12	Y3 FES Organization					
13	Y7 CBCL Chores					
14	Y7 CBCL Structured					
15	Y7 IAS Lessons					
16	Y7 FES Control					
17	Y7 FES Organization					
18	Y16 CBCL Club					
19	Y16 CBCL Chores					
20	Y15 FES Control					
21	Y15 FES Organization					
22	Y4 VFsoft Total					
23	Y4 VFsoft Switch					
24	Y4 VFnoise Total					
25	Y4 VFnoise Switch					
26	Y4 VFround Total					
27	Y4 VFround Switch					
28	Y7 VFround Total					
29	Y7 VFround Switch					
30	Y7 VFmetal Total					
31	Y7 VFmetal Switch					
32	Y7 VFanimal Total					
33	Y7 VFanimal Switch					
34	Y16 VFmetal Total					
35	Y16 VFmetal Switch					
36	Y16 VFround Total					
37	Y16 VFround Switch					
38	Y4 VFsoft Cluster					
39	Y4 VFnoise Cluster					
40	Y4 VFround Cluster					
41	Y7 VFround Cluster					
42	Y7 VFmetal Cluster	.09				
43	Y7 VFanimal Cluster	.02	.00			
44	Y16 VFmetal Cluster	-.03	.13	.02		
45	Y16 VFround Cluster	-.01	.09	-.09	.17	

*Note: to provide estimate unbiased by familial clustering, only Twin 1 correlations shown.

Appendix B

Moos Family Environment Subscales (MFES)

Table B.1: Each item was measured on a binary scale at age 3 (True/False), and on a 5-point Likert scale at ages 7 and 15 (anchored by Strongly disagree = 1; Strongly agree = 5). Reverse-coded items indicated with (R).

Control Subscale	Age 3	Age 7	Age 15
Family members are rarely ordered around. (R)		X	
There are very few rules to follow in our family. (R)	X	X	X
There is one family member who makes most of the decisions.		X	
There are set ways of doing things at home.	X	X	X
There is a strong emphasis on following rules in our family.	X	X	X
Everyone has an equal say in family decisions. (R)		X	
We can do whatever we want to in our family. (R)		X	
Rules are pretty inflexible in our household.	X	X	
*Rules are pretty flexible in our household. (R)			X
You cant get away with much in our family.	X	X	X

Organization Subscale	Age 3	Age 7	Age 15
Activities in our family are pretty carefully planned.		X	
We are generally very neat and orderly.	X	X	X
Its often hard to find things when you need them in our household. (R)		X	
Being on time is very important in our family.	X	X	X
People change their minds often in our family. (R)		X	
Family members make sure their rooms are neat.	X	X	X
Each persons duties are clearly defined in our family.	X	X	X
Money is not handled very carefully in our family. (R)		X	
Dishes are usually done immediately after eating.	X	X	X

Appendix C

Specific Cognitive Battery Subtests used to Derive g-Factor

- Verbal Tasks:
  - ∗ Vocabulary: WPPSI (Year 4); WISC-R (Year 7); Other (Year 16)
  - ∗ Semantic Verbal Fluency (Y4-16)
  - ∗ Word Beginnings and Endings (Y16)
- Spatial Tasks:
  - ∗ McCarthy Puzzle Solving (Y4)
  - ∗ WPPSI Block Design (Y4)
  - ∗ Card Rotation (Y16)
  - ∗ Ravens Progressive Matrices (Y16)
  - ∗ Hidden Patterns (Y16)
- Perceptual Speed Tasks (Y4-Y16):
- Colorado Perceptual Speed Task:
  - ∗ - Varied Dots
  - ∗ - Identical Pictures
- Memory Tasks (Y4-Y16):
  - ∗ Picture Memory Immediate Recognition
  - ∗ Picture Memory Delayed Recognition
- Other (Y16): Subtraction and Multiplication