

# Cognitive Outcomes in Early-Treated Adults With Phenylketonuria (PKU): A Comprehensive Picture Across Domains

Liana Palermo  
Aston University; and Magna Græcia University

Tarekegn Geberhiwot  
Queen Elizabeth Hospital, Birmingham, United Kingdom

Anita MacDonald  
Birmingham Children's Hospital, Birmingham, United Kingdom

Ellie Limback  
Aston University

S. Kate Hall  
UK Newborn Screening Laboratories Network, Manchester,  
United Kingdom

Cristina Romani  
Aston University

**Objective:** Phenylketonuria (PKU) is an inherited metabolic disease which affects cognitive functions due to an inability to metabolize phenylalanine which leads to the accumulation of toxic by-products (Phe) in the brain. PKU can be effectively treated with a low phenylalanine diet, but some cognitive deficits remain. Studies have reported impairments, especially for processing speed and executive functions, but there is a lack of comprehensive assessment across cognitive domains. Moreover, it is important to establish outcomes in early treated adults with PKU (AwPKU) who have better metabolic control than groups previously reported in the literature. **Method:** We tested 37 AwPKU with an unprecedented number of tasks ( $N = 28$ ) and measures ( $N = 44$ ) and compared results with 30 controls matched for age and education. **Results:** We found (a) group impairments, particularly in tasks tapping speed of processing and complex executive functions; (b) high variability across participants, with a sizable number of AwPKU with completely normal performance (about 38%); (c) but also a sizable number of participants who were clearly impaired (about 24%); and (d) good performance in tasks tapping verbal learning, verbal memory and orthographic processing, indicating no generalized learning impairment. **Conclusion:** Our results indicate good outcomes, but also that deficits are still present with current treatment policies.

## General Scientific Summary

This study has advanced our understanding of the cognitive outcome in adults with phenylketonuria (AwPKU), highlighting the success, but also some limitations, of the clinical management. Our results indicate extreme variability across participants and across cognitive domains. As a group, AwPKU showed good performance in learning tasks, but reduced speed of processing and complex executive function deficits. This calls for a better understanding of what causes these impairments.

**Keywords:** Adult PKU, metabolic diseases, cognitive impairments, executive functions and speed of processing, memory and learning

**Supplemental materials:** <http://dx.doi.org/10.1037/neu0000337.supp>

This article was published Online First January 12, 2017.

Liana Palermo, School of Life and Health Sciences, Aston University and Department of Medical and Surgical Sciences, Magna Græcia University; Tarekegn Geberhiwot, IMD Department, Queen Elizabeth Hospital, Birmingham, United Kingdom; Anita MacDonald, IMD Department, Birmingham Children's Hospital, Birmingham, United Kingdom; Ellie Limback, School of Life and Health Sciences, Aston University; S. Kate Hall, UK Newborn Screening Laboratories Network, Manchester, United Kingdom; Cristina Romani, School of Life and Health Sciences, Aston University.

This research was supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme granted to Liana Palermo under the supervision of Cristina Romani and by a grant of the University Hospital Birmingham Charity to Tarekegn Geberhiwot.

We thank Andrew Olson for helpful comments on a version of the manuscript.

This article has been published under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Copyright for this article is retained by the author(s). Author(s) grant(s) the American Psychological Association the exclusive right to publish the article and identify itself as the original publisher.

Correspondence concerning this article should be addressed to Cristina Romani, School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, England. E-mail: [C.Romani@Aston.ac.uk](mailto:C.Romani@Aston.ac.uk)

Phenylketonuria (PKU, ORPHA716) is a disorder caused by an inborn error in amino acid metabolism that, if untreated, causes severe mental retardation, microcephaly, seizures and behavioral problems (Blau, van Spronsen, & Levy, 2010). The metabolic defect underlying the classical form of PKU is a mutation in the gene coding for the enzyme phenylalanine hydroxylase (PAH), which is responsible for the conversion of the amino acid Phenylalanine (Phe) into *tyrosine*. The disrupted metabolism causes accumulation of Phe in the blood, toxic concentrations in the brain, and a reduced level of the amino acid tyrosine. The introduction of newborn screening programs and early treatment with a low-phenylalanine diet has prevented severe neurological damage and mental disability in patients with PKU, but number of studies have documented that even early treated patients suffer from a variety of mild cognitive impairments (see review later on). Still, there is a need of studies which document the performance of more recent cohorts of patients who have followed a stricter diet, and who are old enough to be assessed when brain development and cognitive development are completed. Linked to this need is the need of comprehensive assessments of cognitive functions across domains, because impairments may affect selective domains, and cognitive profile may differ at different developmental ages.

Most studies with early treated adult patients with PKU (from now on AwPKU) have reported impairments in speed of processing and executive functions, but only a very limited number of cognitive functions have been examined by each study (for a review see Christ, Huijbregts, de Sonneville, & White, 2010; Janzen & Nguyen, 2010; Moyle, Fox, Arthur, Bynevelt, & Burnett, 2007). Comparing outcomes across studies, however, is difficult. Functions are not always assessed in the same way and with the same tests and different cohorts of patients are bound to differ somehow for severity, age and socioeconomic status. This makes difficult to compare performance on the same functions across studies and to establish which functions, if any, are most affected by PKU. Moreover, although executive functions and visual attentional skills have been extensively studied, other functions have received less attention.

The lack of *comprehensive* neuropsychological assessments is one of the weaknesses in adult PKU research (but see Brumm, Azen, & Moats, et al., 2004 for an exception). Such assessments, instead, are crucial both for the clinical management of these patients and for a better understanding of the neurophysiological basis of the disease. To accurately evaluate the success of current treatment we need to identify how cognitive performance is affected *across* domains. Only in this way, can we provide patients with the correct educational and psychological support. In addition, we need to know if some functions are more affected than others to understand if the disease disrupts the functioning of specific brain areas and/or specific neurophysiological mechanisms (e.g., see the hypothesis of dopamine depletion following lack of precursor tyrosine vs. the hypothesis of a toxic effect on myelin; Anderson & Leuzzi, 2010; Anderson et al., 2007). Examining outcomes in adult patients is particularly important. Certain functions may only show a developmental delay, so that normal levels are reached even if at a later time. Alternatively, impairments may become more severe or even emerge only with time. Good baseline data is required to track possible disease progression with aging and to compare efficacy of possible treatments that

are alternatives to dietary control (pharmacological treatments, enzyme replacement therapy: see Biomarin, 2016).

The seminal article by Brumm et al. (2004) provided some preliminary evidence of impairments across a variety of cognitive functions for a group of AwPKU. However, the number of patients was relatively limited ( $N = 22$ ). More important, on average, the group had relatively poor metabolic control with an average blood Phe of 702  $\mu\text{mol/L}$  before 10 years of age and  $>1000 \mu\text{mol/L}$  later on. These values are well above current European guidelines (above 18 years: 120–600  $\mu\text{mol/L}$ ; MacDonald personal communication, forthcoming at <http://www.espku.org/who-we-are/european-guidelines>) and American guidelines (American College of Medical Genetics and Genomics: 120–360  $\mu\text{mol/L}$  throughout life; see Vockley et al., 2014). It is important, therefore, to establish if cognitive impairments are still present in AwPKU with better metabolic control reflecting more recent advice to maintain stricter blood Phe control.

In children with PKU, the deficits reported most often involve executive functions and a reduction in processing speed (Albrecht, Garbade, & Burgard, 2009; DeRoche & Welsh, 2008). As mentioned, studies of adult patients have reported similar deficits, but with some inconsistencies in their results, possibly, due to smaller number of participants in each study. We now briefly review these studies as well as studies which have assessed other cognitive functions, before outlining the plan of our study.

## Review of Literature

### Impairments in Executive Functions

Executive functions refer to a heterogeneous set of functions which are necessary to plan and complete tasks in spite of potentially distracting or irrelevant information (e.g., Alvarez & Emory, 2006; Diamond, 2013; Elliott, 2003; Miller & Cohen, 2001). We will review results according to *type* of executive function, subdivided into (a) complex functions including abstract reasoning and planning, (b) working memory, (c) sustained attention, and (d) inhibitory control. Although some researchers make fewer distinctions, we kept functions as separate as possible so that possible differences can be evaluated (see Diamond, 2013; Jurado & Rosselli, 2007; Latzman & Markon, 2010; Miyake et al., 2000, for similar subdivisions).

**Complex executive functions.** We have included in this category tasks requiring rule extraction, planning and flexibility/switching (e.g., Dubois, Slachevsky, Litvan, & Pillon, 2000; Stuss, & Benson, 1986). Typical tests used to assess these functions include the Wisconsin Card Sorting Test (WCST), which requires participants to infer, using feedback from the examiner, the right rule to sort cards into piles and, then, to change sorting criterion accordingly when feedback is changed and the Tower of Hanoi, a puzzle which requires forming a tower by moving circles from one peg to another, following rules. Impairments in the WCST have been reported by Brumm et al. (2004;  $N = 24$ ; current Phe 157–1,713  $\mu\text{mol/L}$ ), by Nardecchia et al. (2015;  $N = 14$  adults; current Phe not reported) and by Smith, Klim, Mallozzi, and Hanley (1996;  $N = 22$ ; current Phe level 200–1,879  $\mu\text{mol/L}$ ), but not by Ris, Williams, Hunt, Berry, and Leslie (1994;  $N = 25$ ; current average Phe 1,320  $\mu\text{mol/L}$ ). Channon, German, Cassina, and Lee (2004) also did not find any deficit ( $N = 20$ ; current Phe 333–1,432  $\mu\text{mol/L}$ ), using a similar test (the Brixton test).

Other tasks that are commonly used to assess executive functions include the *Trail Making Test B* and *Verbal Fluency*. The Trail Making A requires connecting circles containing numbers, scattered randomly on a piece of paper, in ascending number order. The Trail Making B is more complicated because it requires alternating between numbers and letters. The difference between A and B is considered a measure of executive functions (Sánchez-Cubillo et al., 2009). AwPKU have shown normal performance in the *Trail Making B* (Moyle, Fox, Bynevelt, Arthur, & Burnett, 2007;  $N = 12$ , current Phe not reported, and Brumm et al., 2004). Verbal fluency requires participants to come up with as many words as possible in a unit of time consistent with a given criterion (beginning with a given letter or belonging to a given semantic category). Verbal fluency is believed to be part of executive function because it requires *planning* a systematic search of the lexicon to avoid wasting time on exhausted areas (Alvarez & Emory, 2006; Baldo, Shimamura, Delis, Kramer, & Kaplan, 2001; Latzman & Markon, 2010). Verbal fluency was reported to be impaired in AwPKU by Channon et al. (2004) and Brumm et al. (2004), but not by Moyle, Fox, Bynevelt, et al. (2007).

**Working memory/short-term memory.** This is the capacity which allows us to keep in mind, manipulate and refresh the information necessary to complete a task (see Baddeley, 1986; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Typical tasks used to tap this function are forward and backward digit span tasks (where sequences of digits have to be repeated back in the order in which they were presented or in reverse order). The Corsi Block test is a typical visuospatial working memory task (it requires participants to touch a sequence of blocks in the same order in which they were touched by the examiner or in reverse order). Most studies have reported working memory to be impaired in AwPKU (Bik-Multanowski, Pietrzyk, & Mozrzymas, 2011; Brumm et al., 2004; Channon et al., 2004; Channon, Goodman, Zlotowitz, Mockler, & Lee, 2007; Channon, Mockler, & Lee, 2005; but see Moyle, Fox, Bynevelt, et al., 2007 for negative results).

**Sustained attention.** This is the capacity to keep in mind target information for a sustained amount of time in spite of interfering information being presented (Sarter, Givens, & Bruno, 2001; Wilkins, Shallice, & McCarthy, 1987). Typical tasks used to tap this function are the continuous performance task or the rapid visual information processing task, in which participants are presented with sequences of digits and must detect target sequences. Sustained attention has been reported to be impaired in AwPKU (Bik-Multanowski et al., 2011; Schmidt et al., 1994; Weglage et al., 2013).

**Inhibitory control.** This refers to the ability to flexibly modify answers and inhibit inappropriate responses depending on task demands (e.g., Latzman & Markon, 2010; Miyake et al., 2000). A prototypical task tapping this ability is the *Stroop Test* where participants have to name the ink color of words whose meaning refers to a different color (say “yellow” for the word “red” written with yellow ink), thus, suppressing the tendency to read the word. Impairments have been reported in children with PKU, but only sporadically in adults (for a review of children studies, see DeRoche & Welsh, 2008; AwPKU impaired in Sundermann et al., 2011,  $N = 17$ ; current Phe average 1140  $\mu\text{mol/L}$ ; not impaired in Brumm et al., 2004; Feldmann, Denecke, Grenzbach, & Weglage, 2005). No impairment has also been reported in other tasks tapping

inhibitory control such as the *Hayling Sentence Completion Test* which involves completing sentences as quickly as possible with nonsensical words (Channon et al., 2004) and the *Flanker Test* in which participants have to respond to the direction of a central arrow, ignoring flanker arrows that may point in opposite ways (Channon et al., 2005).

### Impairments in Speed of Processing

As evident from the review above, results in the assessment of executive functions have been mixed, in spite of groups being similar for Phe levels at the time of testing (although historical Phe are not always reported and could be responsible for variations). Negative findings are particularly evident for inhibitory control, which is considered a typical executive function. These mixed results have led some authors to suggest that the main residual deficit in AwPKU in one of speed of processing with deficits in executive tasks being largely resolved (e.g., Channon et al., 2005, 2007; Moyle, Fox, Bynevelt, et al., 2007; Feldmann et al., 2005). Channon et al. (2005, 2007), for example, compared the performance of 25 AwPKU (concurrent Phe level 221–1233  $\mu\text{mol/L}$ ) and 25 matched controls on two executive tasks tapping working memory (*n-back*; Braver et al., 1997) and inhibition (Flanker inhibitory task), and two nonexecutive tasks (one involving learning an object location, and one involving categorizing objects for shape or function). AwPKU showed only a speed deficit in the working memory task and performed normally in the other tasks, prompting the conclusion that the only remaining deficit is one of speed. This conclusion, however, may be premature. Although inhibitory control may normalize in adulthood, other executive functions may remain suboptimal.

### Impairments in Other Cognitive Functions

**Visuomotor coordination.** Impairments have been reported both in children and adults, but it is unclear whether deficits are primary or a consequence of a more general speed reduction, because performance is generally tested with timed tasks. In adults, deficits have been found using tasks like the *Grooved Pegboard Test*, which requests participants to insert pegs into small holes with one hand, as quickly as possible (see Griffiths, Paterson, & Harvie, 1995; Pietz et al., 1998, but see Brumm et al., 2004, for negative results), but not with the *Digit-Symbol Coding Test* of the WAIS (Brumm et al., 2004) where participants have to write numbers below symbols according to given letter-symbol pairings. Arguably, digit-symbol coding requires a variety of skills (keeping in mind the number-symbol pairings and switching back on forth between the pairings and the symbols under which the numbers have to be written; see also Davis & Pierson, 2012). However, it has a clear visuomotor component because visual information has to guide the motor response and, for this reason, has been considered by Brumm et al. (2004) to tap mainly visuomotor coordination.

**Language processing, and memory and learning.** Impairments in these domains have been reported less frequently in early treated PKU children (for a review see Janzen & Nguyen, 2010). In AwPKU, results are inconsistent. Typical tasks include learning a list of words across trials (e.g., Rey Word Test), learning the positions of shapes (e.g., paired associated visual learning) or

drawing a complex picture from memory (Rey–Osterrieth Complex Figure). No memory impairment was reported by Griffiths et al. (1995) in the verbal domain and by Channon et al. (2004) in the visuospatial domain. In contrast, memory impairments have been reported by Brumm et al. (2004) and Smith et al. (1996) across domains and by Griffiths et al. (1995) in the visuospatial domain.<sup>1</sup> In addition, Brumm et al. (2004) have reported impairments in expressive naming (Boston Naming Test), but not in receptive vocabulary (Peabody Picture Vocabulary Test).

### Plan of Study and Data Analyses

In sum, for most cognitive functions including complex executive functions, visuomotor coordination, language processing, and memory and learning, results have been mixed. Deficits have been reported more consistently for working memory and sustained attention and normal performance for inhibitory control. Some authors have suggested that only deficits of speed of processing characterize AwPKU.

Our study will establish the cognitive profile of a group of 37 AwPKU with better metabolic control than reported in previous studies (average 432  $\mu\text{mol/L}$ , in childhood; <850  $\mu\text{mol/L}$  later on). Our study is similar to Brumm et al. (2004) in assessing a large number of functions and using a variety of tests for each function to reduce error. Differently from Brumm et al., we have decomposed executive functions in a number of subtypes, assessed visuospatial attention as a separate domain, and included assessment of orthographic skills.

Besides performance in individual tasks and functions, we will consider two general measures of performance across domains: (a) an overall standard score, which averages across cognitive tests and measures (RTs and accuracy); and (b) the proportion of impaired scores. This second measure is also very important because good performance in a number of tasks may counterbalance and mask severe impairments in others and result in an overall score which is normal or close to normal. Therefore, we will also consider, for each PKU participant, the number of cognitive domains/measure where severe difficulties are encountered. Even control participants are anticipated to demonstrate some poor scores, but this number may be much higher in AwPKU. For these measures (overall standardized score and rate of impaired scores) we will report both group averages and proportions of individuals impaired. To be conservative, we will consider impaired scores which deviate 2 or more standard deviations from the mean of a control group that we tested for comparison and was matched to the AwPKU for age and education ( $z$  scores = >2). We will consider scores which are within 0.5 standard deviations from the control mean ( $z$  scores = <.05) to be clearly normal.

## Method

### Participants and Procedure

Thirty-seven early treated patients with classical PKU were recruited from a pool of patients currently followed by the Department of Inherited Metabolic Disorders at the Queen Elisabeth Hospital in Birmingham who had been continuously treated with a low-phenylalanine diet since birth. Diagnosis was through newborn screening conducted at 5–7 days after birth. At the time of

testing 7 patients were on an unrestricted diet and 30 on a low phenylalanine diet. We invited all early treated patients attending the clinic to participate, plus patients who were still contactable, but no longer in regular follow-up. All individuals who responded to the invitation were tested. No patient had been diagnosed with any psychiatric disorder. Data on historical Phe levels were obtained from the PKU database at The Clinical Chemistry Department at Birmingham Children's Hospital.

The PKU participants were compared with a group of 30 healthy control participants matched for age, gender, and educational status. They were recruited through the Aston University volunteering website.

Participants were tested in a quiet room in two separate testing sessions, each lasting approximately 3 hours. Blood Phe concentrations were measured prior to each testing session to determine current Phe levels. The research was approved by the NHS and Aston University Ethics committees. All participants gave voluntary informed consent to take part. All efforts were made to administer all tasks to all participants, but some data points are missing because not everybody returned for the second session (see Tables 1–3 for exact numbers). Whenever possible we report performance in terms of error rates so that like for speed measures higher scores indicate lower performance.

### Tasks

IQ was measured using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), which includes the following subtests: Vocabulary, Block Design, Similarities, and Matrix Reasoning. These subtests are similar in format to their WISC–III and WAIS–III counterparts. They are those with the highest loadings on general intellectual functioning (Wechsler, 1991, 1997) and allow estimates of both Verbal and Performance IQ. In addition, participants were given an extensive neuropsychological battery. Tasks are briefly described below. A more extensive description is provided in the supplementary materials.

**Visuospatial attention.** This included measures from 6 tasks: (a) *Simple Detection*: Press a response button as soon as a ladybird appeared on the screen; (b) *Detection with Distractors*: Press a button whenever a ladybird appeared on the screen alone or with a green bug; in the second part of the task the instruction was changed to press a button whenever a green bug appeared on the screen alone or with a ladybird; (c) *Choice Reaction Time*: Press either a left or right response key consistent with the direction of an arrow centrally presented; (d) *Feature Search*: Detect a target among distractors not sharing features by pressing a 'yes' or 'no' button (e.g., a red ladybird among green bugs); and (e) *Conjoined Search*: Detect a target among distractors sharing features (e.g., red ladybird among red bugs and green ladybirds). Both RTs (reaction time (RT) from now on) and accuracy measures (error rates) were taken.

**Visuomotor coordination.** This included measures from 2 tasks: (a) *Grooved Pegboard Test*: Put pegs into the holes of a

<sup>1</sup> Note that a limit of the Complex Figure Test (Rey, 1941) when used to test visuospatial memory is that it also taps executive function skills because it requires planning and ongoing monitoring of performance (Watanabe et al., 2005). Thus, it is not clear whether an impairment is due to a visuospatial deficit or an executive deficit.

Table 1  
Demographic Information, Metabolic Control, and General Cognitive Performance for 30 Control and 37 PKU Participants

Variable	Controls		PKU participants		Controls vs. PKU
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age	<b>27.6</b>	7.4	<b>27.5</b>	7.3	$t_{(1, 65)} = 0.08; p = .94$
Education years	<b>15.2</b>	1.7	<b>14.4</b>	1.9	$t_{(1, 65)} = 1.6; p = .12$
Gender (M/F)	<b>10/20</b>		<b>13/24</b>		$\chi^2_{(1)} = .02; p = .88$
Childhood (mean <i>N</i> obs. = 197; <i>SD</i> = 165)					
Phe average			<b>432</b>	243	
Phe fluctuation ( <i>SD</i> )			<b>205</b>	63	
Adolescence (mean <i>N</i> obs. = 77; <i>SD</i> = 70)					
Phe average			<b>721</b>	340	
Phe fluctuation ( <i>SD</i> )			<b>157</b>	58	
Adulthood (mean <i>N</i> obs. = 65; <i>SD</i> = 74)					
Phe average			<b>802</b>	324	
Phe fluctuation ( <i>SD</i> )			<b>137</b>	68	
Lifetime (mean <i>N</i> obs. = 340; <i>SD</i> = 241)					
Phe average			<b>634</b>	291	
Phe fluctuation ( <i>SD</i> )			<b>164</b>	58	
Current Phe			<b>720</b>	343	
Cognitive performance					
FIQ	<b>113.8</b>	10.9	<b>103.9</b>	14.3	$t_{(1, 65)} = 3.1; p = .003$
VIQ	<b>112.2</b>	10.2	<b>102.3</b>	12.9	$t_{(1, 65)} = 3.1; p = .001$
PIQ	<b>112.2</b>	11.7	<b>104.5</b>	15.1	$t_{(1, 65)} = 2.2; p = .03$
Overall <i>z</i> score	<b>.0</b>	.3	<b>.5</b>	.7	$t_{(1, 65)} = -3.8; p < .001$
Impaired participants					
<i>N</i>	<b>0</b>		2		
%	<b>0</b>		5.4		
Impaired measures (%)	<b>3.8</b>	3.7	<b>13.2</b>	15.1	$t_{(1, 65)} = -3.4; p = .001$
Impaired participants					
<i>N</i>	<b>1</b>		<b>14</b>		
%	<b>3.3</b>		<b>37.8</b>		$\chi^2_{(1)} = 11.35; p < .001$

Note. Blood Phe measured in  $\mu\text{mol/L}$ . Impaired =  $> 2$  *z* scores from control means. PKU = phenylketonuria; Phe = Phenylalanine.

board using only one hand as quickly as possible and (b) *Digit Symbol Task*: Fill as many boxes as possible with symbols corresponding with numbers in 90 s.

**Complex executive functions.** This included measures from 4 tasks which, to be completed involve planning, flexibility and abstract thinking. First, *The Wisconsin Card Sorting Test*, (64 Cards Version) in which participants have to discover the rules to match cards from a deck with four reference cards according to shape, number and color using feedback. Flexibility is required when the sorting rule is changed unknown to the participant and the new rule has to be discovered. A number of detailed measures are reported in Table 3 below. To compute an aggregate score for 'Complex executive functions' we used total number of errors. Second was difference in speed between *Trail Making Test B-A* (A involves connecting circles containing numbers in ascending order as quickly as possible; B also involves connecting circles in ascending order but alternating between number and letters). Only speed measures are taken because errors are rare. Third, for *Fluency*. For letter fluency, participants generate as many words as possible starting with a given letter in one minute of time (letters: C, F and L). For semantic fluency, they generate name of animals. We only considered semantic fluency in the aggregate score for 'Complex Executive functions' as better reflecting the ability to carry-out an efficient lexical search. And finally, for *The Tower of Hanoi* puzzle, participants move a number of rings of different sizes across three pegs to form a tower on the last peg following specific constraints. Our score is based on the percentage of solved trials of different complexity (3, 4, 5 rings).

**Inhibitory control.** This included measures from 2 tasks: (a) *Stroop* interference: Difference in time and errors between reporting the ink color of words where the color of the ink was incongruent with the meaning of the word; "red" written with yellow ink, or congruent, "red" written with red ink; (b) *Semantic Interference*: Differences in naming between the first and last exemplar of a series of semantically related nouns in terms of RT and errors.

**Short-term Memory/Working Memory.** This included measures from 3 tasks: (a) *Digit Span*: Repeat a sequence of digits spoken by the examiner, soon after presentation; (b) *Nonword Repetition*: Repeat a sequence of nonwords spoken by the examiner, soon after presentation; and (c) the *Corsi Block Tapping Test*: The examiner taps a sequence of blocks and the participant has to reproduce the sequence in the same order.

**Sustained Attention.** Percentage correct from the Rapid Visual Information Processing task: detect three target sequences of 3 digits by pressing the response key when the last number of the sequence appears on the screen.

**Orthographic Language.** This involved measures from 6 tasks: (a) and (b) *Word and Nonword Reading*: Read as fast as possible an English or a made up word and both RT and accuracy measures (error rates) are taken; (c) and (d) *Word and Nonword Spelling*: Spell words/nonwords to dictation; (e) *Phoneme Deletion*: Delete a sound from a word (e.g., powder; /d/ > power); and (f) *Spoonerisms*: Exchange the initial sounds of two words to produce two different words (e.g., bad-sin > sad-bin). These last two tasks were included because performance shows strong cor-

Table 2  
*Speed and Accuracy in Visuospatial Attention, Visuomotor, and Visuospatial Memory Tasks in Control and PKU Participants*

Domain/task	Controls		PKU participants				Controls vs. PKU	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>N</i>	% <i>z</i> = > 2	% <i>z</i> = > 1	<i>t</i> and <i>p</i>
<b>Visuospatial attention</b>								
Simple detection (RT – ms)	<b>316</b>	56.9	332	52.9	31	6.5	16.5	$t_{(1, 59)} = -1.1; p = .27$
Detention with distractors								
RT – ms	<b>407</b>	72.2	<b>438</b>	70.6	31	3.2	22.6	$t_{(1, 59)} = -1.7; p = .09$
% errors	<b>.6</b>	.6	<b>.8</b>	1.1	31	6.5	19.4	$t_{(1, 59)} = -.8; p = .45$
Choice RT								
RT – ms <sup>a</sup>	<b>281</b>	31.3	<b>307</b>	42	37	16.2	32.4	$t_{(1, 65)} = -2.8; p < .01$
% errors	<b>.4</b>	.8	<b>.5</b>	.8	37	8.1	8.1	$t_{(1, 65)} = -.4; p = .68$
Feature search								
RT – ms <sup>a</sup>	<b>498</b>	71.7	<b>606</b>	167.6	31	32.3	41.9	$t_{(1, 59)} = -3.2; p < .01$
% errors	<b>2</b>	2.3	<b>1.6</b>	2.6	31	6.5	16.1	$t_{(1, 59)} = .7; p = .50$
Conjoined search								
RT – ms <sup>a</sup>	<b>841</b>	126.3	<b>1,008</b>	239.7	31	25.8	38.7	$t_{(1, 59)} = -3.4; p < .01$
% errors	<b>3.1</b>	4.4	<b>2.4</b>	2.7	31	0	3.2	$t_{(1, 59)} = .8; p = .43$
<b>Visuomotor coordination</b>								
Pegboard (time – seconds) <sup>a</sup>	<b>59.2</b>	5.1	<b>65.1</b>	11.7	37	18.9	43.2	$t_{(1, 65)} = -2.6; p = .01$
Digit symbol (% errors in 90 s) <sup>a</sup>	<b>27.1</b>	10.2	<b>37.3</b>	11.3	31	19.4	41.9	$t_{(1, 59)} = -3.7; p < .01$
<b>Visuospatial memory and learning</b>								
Paired Associate Verbal Learning (% errors)	<b>1.9</b>	1.9	<b>3.1</b>	3.82	31	12.9	16.1	$t_{(1, 59)} = -1.5; p = .15$
Delayed matching to sample (% errors)	<b>10.5</b>	8.3	<b>14.1</b>	8.5	37	10.8	29.7	$t_{(1, 65)} = -1.7; p = .09$

Note. Percentage of Phenylketonuria (PKU) participants with *z* score = >2 (expected % 2.3) is also reported. RT = reaction time.

<sup>a</sup>Tasks where there is a significant difference with controls.

relations with orthographic skills (Landerl, Wimmer, & Frith, 1997; Romani, Tsouknida, & Olson, 2015).

**Spoken Language.** This included measures from four tasks: (a) *Picture Naming*: Name a picture as fast as possible; (b) *Color naming*: Name as fast as possible the ink color of three Xs or colored words; only the congruent condition is considered here, where the color of the ink matched the meaning of the word, for example, “red” written with red ink. Both RT and accuracy measures (error rates) were taken; (c) *Similarities* from the WASI: Describe how similar in meaning two words are; and (d) *Vocabulary* from the WASI: Define a word.

**Verbal Memory and Learning.** This included measures from 2 tasks: *The Rey Auditory Verbal Learning Test* (learning, immediate recall and delayed recall of a list of 15 words) and *Paired Associates Verbal Learning* (learning the association between a made-up word and the picture of an object or animal).

**Visual Memory and Learning.** This included measures from 2 tasks: (a) *Delayed Matching to Sample*: Recognize a previously seen pattern among distractors and (b) *Paired Associates Visual Learning*: Learn to associate shapes with locations.

## Results

Demographic and Phe data, as well as general cognitive performance, are reported in Table 1. Metabolic values are reported in three age bands: childhood (0–10 years old), adolescence (11–16 years old), and adulthood (17 years to present), as well as at testing time and across the lifetime. We report metabolic control both in terms of mean Phe level and Phe fluctuation. The Phe level in each band was calculated by averaging the Phe medians for each year in the band; the Phe fluctuation was calculated by averaging standard

deviations from median values (*SD*) for each year in the band. For each AwPKU, the current Phe was calculated averaging Phe level at Session 1 and Phe level at Session 2. Across the group, Phe levels were better controlled in childhood; diet was progressively relaxed after early childhood with increasing blood Phe levels.

## General Cognitive Performance and Variability Between Participants

Overall, AwPKU had a full scale IQ in the average range (only one impaired patient), but significantly lower than matched controls (see also DeRoche & Welsh, 2008; and Moyle, Fox, Arthur, et al., 2007, for a meta-analysis). Considering the whole neuropsychological assessment together (44 measures), the mean AwPKU *z* score was only slightly below the control average (*z* score = 0.5). However, a different picture emerged when performance with individual tasks was considered. Specifically, for each participant we calculated the percentage of impaired measures (a measure was considered impaired if it was >2 *z* score from the mean of our control group) on the overall number of available measures. On average, 4% ( $\pm 3.7\%$ ) of measures were impaired in controls versus 13% ( $\pm 15.1\%$ ) AwPKU. In 17/37 (46%) of PKU participants, the rate of impaired measures was significantly above control average. Moreover, 9 AwPKU (24% of the PKU sample) showed both a pathological rate of impaired measures (>2 *z* scores) and an overall *z* score below average ( $= >1$  *z* score), thus showing overall a clear impairment. In contrast, 38% of PKU participants (14/37) showed a completely normal profile with *z* scores = < 0.5, both for rate of impaired measures and overall average performance.

Table 3  
Executive Function Tasks in Control and PKU Participants

Domain/task	Controls		PKU participants				Controls vs. PKU	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>N</i>	% <i>z</i> = > 2	% <i>z</i> = > 1	<i>t</i> and <i>p</i>
Complex executive functions								
WCST								
Total errors <sup>a</sup>	<b>10.8</b>	4.9	<b>14</b>	8.2	37	24.3	27	$t_{(1, 65)} = -2.0; p = .05$
Perseverative responses	<b>6.8</b>	4.4	<b>8.1</b>	5.1	37	10.8	27	$t_{(1, 65)} = -1.1; p = .27$
Perseverative errors	<b>6.1</b>	3.3	<b>7.4</b>	4.3	37	10.8	24.3	$t_{(1, 65)} = -1.4; p = .16$
<i>N</i> of completed categories <sup>a</sup>	<b>4.4</b>	.9	<b>3.9</b>	1.2	37	13.5	32.4	$t_{(1, 65)} = 2.1; p = .04$
Trail making: B–A (Time – seconds)	<b>20.7</b>	13.3	<b>19</b>	10.8	37	5.4	10.8	$t_{(1, 65)} = -.6; p = .57$
Tower of Hanoi (% of not solved trials) <sup>a</sup>	<b>2.9</b>	7.9	<b>20.6</b>	28.2	27	33.3	48.1	$t_{(1, 52)} = -3.1; p < .01$
Verbal fluency								
Letter (correct answers)	<b>41.8</b>	12.9	<b>35.9</b>	11.4	37	2.7	27	$t_{(1, 65)} = 1.9; p = .06$
Semantic (correct answers) <sup>a</sup>	<b>25</b>	5	<b>21.4</b>	6.2	37	10.8	37.8	$t_{(1, 65)} = 2.6; p = .01$
Inhibitory control								
Stroop Test: Interference								
RT – ms	<b>94.3</b>	51.9	<b>111.3</b>	86.7	30	16.7	23.3	$t_{(1, 58)} = -1.4; p = .18$
% errors	<b>.8</b>	1	<b>.5</b>	1.1	30	3.3	13.3	$t_{(1, 58)} = .5; p = .63$
Naming: Semantic interference								
Position 5–1 (ms.)	<b>145</b>	62.9	<b>173</b>	94.6	30	16.7	33.3	$t_{(1, 58)} = -1.4; p = .18$
Position 5–1 (% errors)	<b>1.1</b>	1.6	<b>1.3</b>	2	31	9.7	16.1	$t_{(1, 59)} = .3; p = .74$
Short-term memory								
Digit span <sup>a</sup>	<b>6.5</b>	.9	<b>6.1</b>	1	37	10.8	35.1	$t_{(1, 65)} = 2.0; p = .05$
Nonword repetition (% errors) <sup>a</sup>	<b>39.3</b>	10.4	<b>48.2</b>	12.5	31	19.4	45.2	$t_{(1, 59)} = -3.0; p < .01$
Corsi block-tapping test span	<b>5.7</b>	<b>.9</b>	<b>5.3</b>	.9	37	2.7	29.7	$t_{(1, 65)} = 1.7; p = .10$
Sustained attention								
RVP (% of errors) <sup>a</sup>	<b>13.3</b>	8.8	<b>18.9</b>	11.3	37	13.5	35.1	$t_{(1, 65)} = -2.2; p = .03$

Note. Percentage of phenylketonuria (PKU) participants with *z* score = >2 (expected % 2.3) is also reported. WCST = Wisconsin Card Sorting Test; RVP = Rapid Visual Information Processing.

<sup>a</sup>Tasks where there is a significant difference with controls.

AwPKU with impaired ( $N = 9$ ) versus normal ( $N = 14$ ) performance did not have significantly different Phe levels during childhood ( $\mu\text{mol/L}$  469 vs. 485;  $t(1,18) = -1.21; p = .90$ ), but differed significantly for metabolic control in adolescence ( $\mu\text{mol/L}$  659 vs. 944;  $t(1,19) = -2.2; p = .04$ ) and adulthood ( $\mu\text{mol/L}$  782 vs. 1128;  $t(1,21) = -3.3; p = .003$ ) and for current Phe ( $\mu\text{mol/L}$  713 vs. 1,059;  $t(1,20) = -2.7; p = .01$ ). They also differed in terms of Phe variability (in terms of average *SD* per year) in adolescence ( $\mu\text{mol/L}$  *SD* 148 vs. 214;  $t(1,19) = -3.5; p = .002$ ) and through the life span ( $\mu\text{mol/L}$  *SD* 162 vs. 212;  $t(1,21) = -2.5; p = .02$ ), but variability was not different at other times.

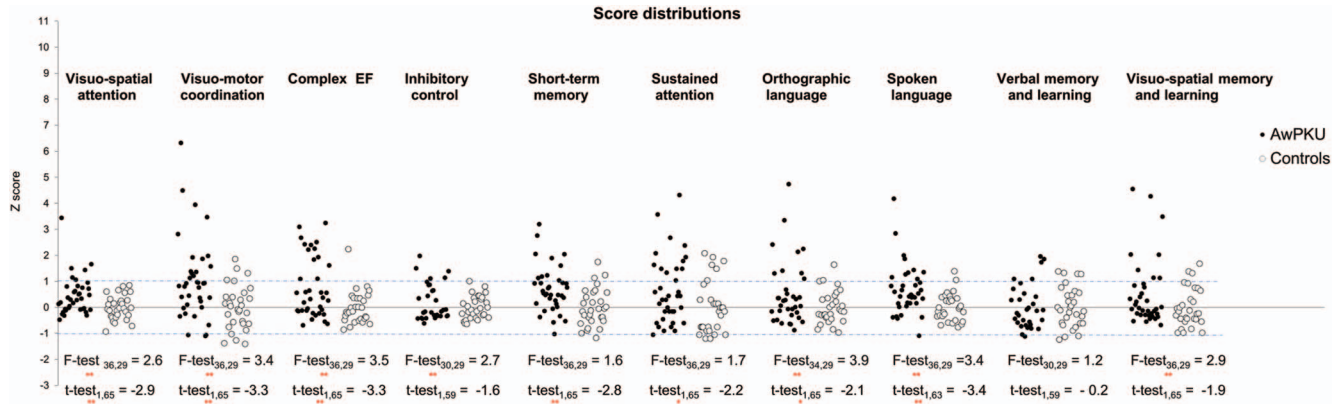
Figure 1 shows that although, in each domain, most scores of AwPKU fall within 1 *SD* of the control mean, there are a substantial number of AwPKU performing below average. The only domain where this does not occur is verbal memory and learning where the distribution of scores appears similar in the PKU and control group.

### Variability Across Domains

Performance of the PKU and control groups on individual tasks is shown in Tables 2–4. Table 2 shows results for visuospatial attention, visuomotor coordination and visuospatial memory. AwPKU are systematically impaired in tasks requiring visuospa-

tial attention except for the easiest tasks (Simple Detection and Detection with Distractors). No significant impairments are seen in accuracy measures, but this could be due to lack of sensitivity of these measures (performance being close to 100% correct across groups). Performance on the tasks tapping visuospatial memory and learning failed to reach significance when the tasks were considered individually, but was marginally lower when scores are aggregated, as shown in Figure 1. Performance on tasks tapping visuomotor coordination is impaired, consistent with previous results (see Griffiths et al., 1995; Pietz et al., 1998).

Table 3 shows results for executive functions. There is a contrast between tasks requiring higher order executive functions where differences are highly significant (WCST, Tower of Hanoi, Semantic Fluency) and tasks which arguably require less planning and monitoring, where no significant difference is present (the Corsi Block, the Letter Fluency and the Trial making B-A). There is also no impairment in tasks tapping inhibitory control, consistent with previous studies (Brumm et al., 2004; Channon et al., 2004, 2005). Finally, there is a significant impairment of sustained attention (see also Bik-Multanowski et al., 2011; Schmidt et al., 1994; Weglage et al., 2013). It is important to note that impairments are seen in executive functions even if tasks are untimed or have a minimal timing component. The WCST is untimed. The Tower of Hanoi is timed (maximum of 6 min for each trial), but



*Figure 1.* Distribution of scores for the Control and PKU groups across speed tasks. Each point is the performance of an individual participant;  $t$  tests assess differences in average performance between groups; Variance ratio  $F$  test assess differences in variability. Tasks were aggregated across domains by averaging  $z$  scores across tasks/measures as follows: (a) Visuospatial attention: RTs in *Simple Detection*, and RT and accuracy in *Detection with Distractors*, *Choice Reaction Time*, *Feature Search*, *Conjoined Search*; (b) visuo-motor coordination: *Grooved Pegboard Test*, *Digit symbol*; (c) complex executive functions: *WCST score*, *Tower of Hanoi scores*, *Semantic fluency*, *Trial Making B-A*; (d) inhibitory controls: *Stroop interference effect*; *Semantic interference effects* (both RTs and accuracy); (e) STM (short-term memory): *Digit Span*, *Nonword repetition*, *Corsi Span*; (f) Sustained attention: *Rapid Visual Information Processing -RVP*; (g) Orthographic Language: RT and accuracy in *words and nonword reading*, accuracy in *word and nonword spelling*, *spoonerisms*, and *phoneme deletions*; (h) Spoken Language: RT and accuracy in *picture naming* and *color naming*, *Vocabulary WASI*, *Similarities WASI*; (i) verbal memory and learning: *The Rey Auditory Verbal Learning Test* and *Paired Associates Verbal Learning*; (j) Visuospatial memory and learning: *Delayed Matching to Sample* and *Paired Associates Visual Learning*. \*  $p < 0.05$ . \*\*  $p < 0.01$ . Please see the online article for the color version of this figure.

participants rarely fail due to time constraints. Verbal fluency has a speed element because scores are based on the number of words produced in a fixed unit of time. However, the quality of the lexicon (its size and structure) and the strategies adopted to search it are more important than speed of access per se. Similarly, the task tapping sustained attention involved detection of number sequences within time constraints. However, the detection window included two digits following the last digit of the target sequence (totaling 1,800 ms) which should minimize difficulties with speed. The fact that, as a group, AwPKU show impairments in all of these tasks argues against the hypothesis that, in adulthood, deficits *only* involve a reduction of speed of processing.

Table 4 shows results for language tasks. Performance on tasks probing orthographic knowledge is good. Speed in word and nonword reading is impaired, consistent with a generalized reduction in speed of processing, but accuracy is good, as is performance in spelling, spoonerisms and the phoneme deletion task. Word spelling is numerically better than in the controls. Note also that word reading speed was relatively less impaired than nonword reading speed, group by task interaction:  $F_{(1, 59)} = 10.5$ ,  $p < .01$ ; partial eta-squared = .2, consistent with the deficit being more in components involved in perceptual analysis than in lexical access. Accuracy in picture naming was good. More taxing tasks probing vocabulary (vocabulary and similarity subtests of the WASI) were impaired, but these tasks have a strong reasoning component which could be responsible for the deficit. Finally, performance was excellent in memory and learning tasks. Combined, these results indicate excellent ability to learn in our AwPKU group.

To better compare the relative severity of impairments with speed and accuracy measures Figure 2 reports severity of PKU impairment in terms of Cohen's  $d$  for different domains. The Cohen's  $d$  is the difference between the mean performance of PKU and control groups divided by the pooled standard deviation (see Rosenthal, 1994). In our case, the higher the score the bigger the difference between the two groups.

It is clear that speed of processing is systematically impaired across domains with significant differences in most tasks and, on average, a medium Cohen's  $d$  (.58). Accuracy measures return a lower average Cohen's  $d$  (.26), but there is high variability across domains, making this figure poorly representative. The Cohen's  $d$  remains high for tasks which are taxing and require complex skills, including *planning* and *switching* (WCST, Tower of Hanoi, and Fluency), *monitoring* (short-term memory [STM], Sustained attention) and *verbal reasoning* (Similarities, Vocabulary of WASI).

## General Discussion

We have carried out a comprehensive neuropsychological investigation in a group of 37 AwPKU with relatively good metabolic control since childhood and 30 matched controls. This has advanced our understanding of the effect of PKU on cognition by identifying which cognitive functions are generally spared and which continue to be affected and by providing a more complete picture and the severity and variability of impairments in this population.



Table 4  
*Speed and Accuracy in Orthographic and Spoken Language Tasks and Verbal Memory and Learning Tasks in Control and PKU Participants*

Domain/task	Controls		PKU participants					Controls vs. PKU <i>t</i> and <i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>N</i>	% <i>z</i> = > 2	% <i>z</i> = > 1	
Orthographic language								
Word reading								
RT – ms <sup>a</sup>	<b>507</b>	93.3	<b>574</b>	120	30	6.7	30	$t_{(1, 58)} = -2.4; p = .02$
% errors	<b>.6</b>	.8	<b>.7</b>	1	30	10	16.7	$t_{(1, 58)} = -.2; p = .84$
Nonword reading								
RT – ms <sup>a</sup>	<b>604</b>	103.6	<b>795</b>	296.7	31	38.7	51.6	$t_{(1, 59)} = -3.3; p < .00$
% errors	<b>5.8</b>	5.7	<b>7.9</b>	10.8	31	12.9	22.6	$t_{(1, 59)} = -1.0; p = .33$
Word spelling (% errors)	<b>4.8</b>	5.7	<b>3.9</b>	4.7	29	3.4	13.8	$t_{(1, 57)} = .7; p = .50$
Nonword spelling (% errors)	<b>11.9</b>	7.6	<b>13.3</b>	8.2	35	8.6	20	$t_{(1, 63)} = -.7; p = .49$
Spoonerisms (% errors)	<b>6.8</b>	6.6	<b>10.9</b>	14.7	30	13.3	23.3	$t_{(1, 58)} = -1.4; p = .17$
Phoneme deletion (% errors)	<b>11.2</b>	8.7	<b>14.5</b>	13.1	35	14.3	25.7	$t_{(1, 63)} = -1.2; p = .24$
Spoken language								
Pictures naming								
RT – ms	<b>835</b>	138.1	<b>897</b>	138.9	29	3.4	27.6	$t_{(1, 57)} = -1.7; p = .09$
% errors	<b>7.1</b>	5.1	<b>8.5</b>	11	30	3.3	6.7	$t_{(1, 58)} = -.6; p = .53$
Color naming								
RT – ms <sup>a</sup>	<b>617</b>	98.2	<b>731</b>	141.8	29	17.2	44.8	$t_{(1, 57)} = -3.6; p < .01$
% errors	<b>.1</b>	.2	<b>.1</b>	.2	29	3.4	6.7	$t_{(1, 57)} = -.1; p = .89$
Vocabulary WASI <sup>a</sup> (correct answer raw score)	<b>64.2</b>	7.2	<b>58.4</b>	8.7	37	10.8	40.5	$t_{(1, 65)} = -2.9; p < .01$
Similarities WASI <sup>a</sup> (correct answer raw score)	<b>39.2</b>	3.6	<b>36</b>	5.8	37	29.7	56.8	$t_{(1, 65)} = -2.6; p = .01$
Verbal memory and learning								
Rey Auditory Verbal Learning Test								
Trial A1–A5 (% errors)	<b>20.5</b>	8.4	<b>24.1</b>	11.2	31	16.1	32.3	$t_{(1, 59)} = -1.4; p = .17$
Retention (% errors A6)	<b>15.8</b>	16.1	<b>18.3</b>	15.6	31	3.2	22.6	$t_{(1, 59)} = -.6; p = .54$
Delayed recall (% errors)	<b>15.1</b>	15.1	<b>16.3</b>	15.9	31	6.5	19.4	$t_{(1, 59)} = -.3; p = .76$
Paired Associate Verbal Learning								
Trial 1–5 (% errors)	<b>47</b>	24.9	<b>43.5</b>	22.9	31	0	12.9	$t_{(1, 59)} = .6; p = .58$
Delayed recall (% errors)	<b>29.3</b>	24.5	<b>20.4</b>	23.7	31	3.2	9.7	$t_{(1, 59)} = 1.4; p = .16$

Note. Percentage of phenylketonuria (PKU) participants with *z* score = > 2 (expected % 2.3) is reported. WASI = Wechsler Abbreviated Scale of Intelligence.

<sup>a</sup>Tasks where there is a significant difference with controls.

## Cognitive Impairments

We have shown significant and widespread cognitive impairments, even for a group of AwPKU who have had better metabolic control than reported in previous studies. Europe recommendation for target Phe levels have been very variables across countries (Blau et al., 2010). New forthcoming European guidelines from ESPKU suggest a target of <360 μmol/L in childhood and <600 μmol/L afterward. Our group does not completely meet these guidelines but comes closer than previously reported groups (e.g., concurrent Phe level in our group = 720 μmol/L compared with most groups were it is ~1,000 μmol/L; see Introduction). Testing groups with better control will continue to be important in the future to provide firm evidence of the success and limitations of current treatment.

In our PKU group, impairments were present across a variety of domains showing that, although reduction in speed of processing is a factor underlying cognitive difficulties, impairments are also present in tasks with no or a minimal speed component. Impairments were present especially in complex tasks involving reasoning, planning and monitoring, and, more generally, in tasks which required the orchestration of a variety of skills. In contrast, our

PKU group did not show learning difficulties and performance in tasks of verbal and visual learning and memory was very good. Performance in spelling, another task which relies heavily on learning and on the capacity to store information, was also excellent, as was performance on tasks which relied heavily on orthographic knowledge (phoneme deletion, spoonerisms). Good performance here is not due to lack of test sensitivity. On the contrary, these learning and orthographic tasks are challenging even for control participants and among the most sensitive tasks to detect impairments in other clinical populations such as adults with developmental dyslexia (e.g., see Romani, Di Betta, Tsounkida, & Olson, 2008). The good performance on these tasks may also reflect the good support received by our participants with PKU and/or acquired coping strategies. AwPKU may have learned to rely more on stored knowledge to compensate for their weaknesses in speed of processing.

**Speed of processing.** In PKU studies, speed of processing has, generally, been measured with visuospatial tasks where participants are asked to detect targets in different types of displays. In our study, we have also shown impairments in tasks involving picture naming, and word and nonword reading. Instead, we have

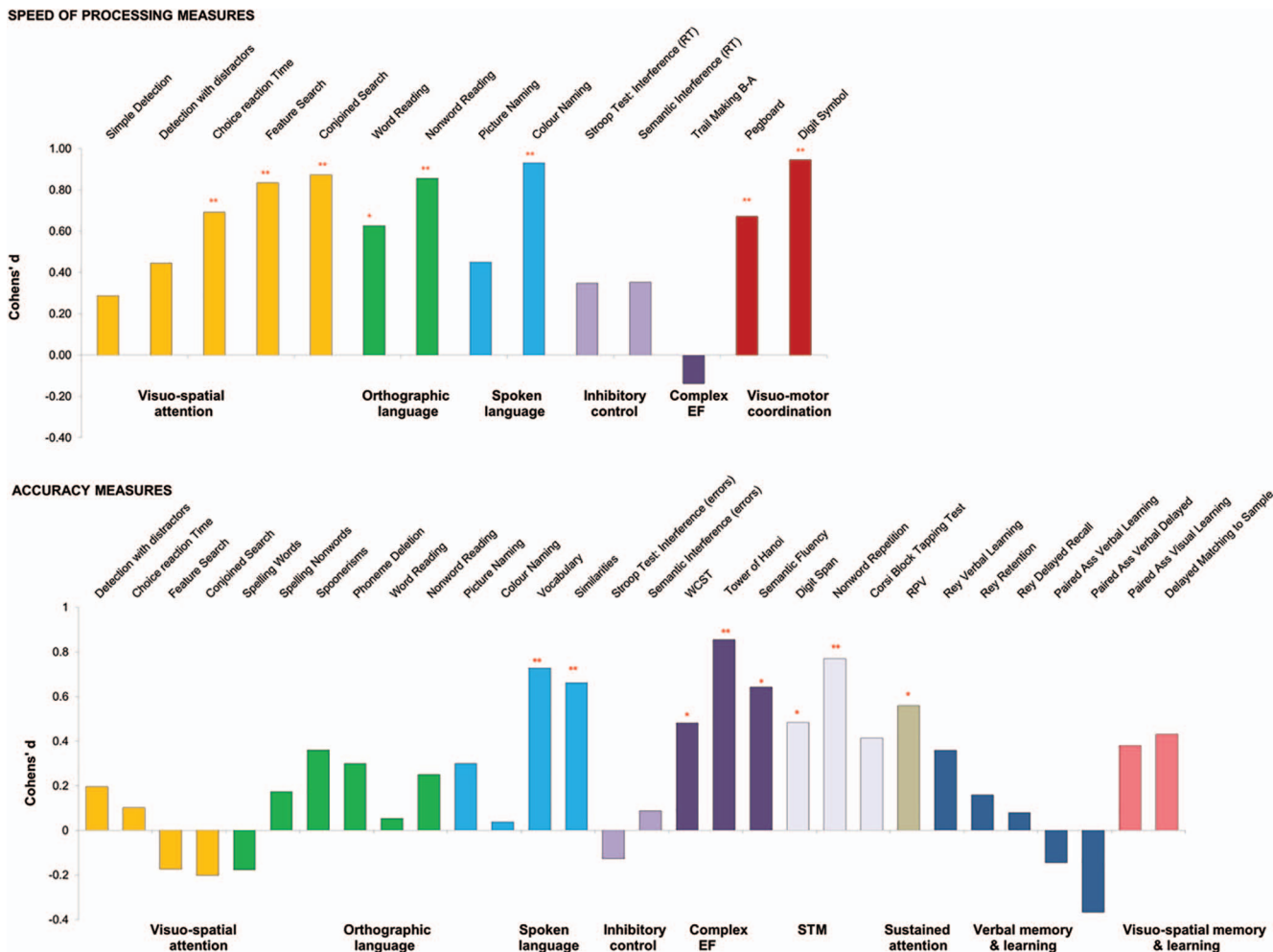


Figure 2. Differences in performance of the PKU group from the control group in terms of Cohen's *d* across functions; Cohen's *d* = difference between the two groups averages divided by the pooled *SD*. Asterisks indicate a significant difference between AwPKU and Control. \*  $p < 0.05$ . \*\*  $p < 0.01$ . See the online article for the color version of this figure.

shown no impairment in tasks which require speed but minimize cognitive processing by only requiring a response when a given stimulus appears on the screen (Simple Detection and Detection with Distractors). This result is important because it shows that AwPKU experience *cognitive* slowing, but not a *peripheral* reduction in motor speed. If PKU caused a generalized speed deficit, impacting equally on the central and the peripheral nervous system, then a task requiring simple target detection (particularly in the absence of distractors) should be especially affected. Here, the main factor affecting performance should be the speed with which action potentials travel along the long and heavily myelinated axons of the motor neurons. If myelination of these axons were degraded, speed would be impaired. Our results, instead, indicate that slowing occurs at the cognitive level, consistent with the toxic effects of high Phe levels on the oligodendroglia in the central nervous system (see Möller et al., 2003; Moyle, Fox, Arthur, et al., 2007), but not on the Schwann cells in the peripheral nervous system (see Joseph & Dyer, 2003). Consistent with our results, a

recent meta-analysis (Albrecht et al., 2009) reported choice RT to be the speed test most sensitive to phenylalanine concentrations.

**Impairment of executive functions.** We have found significant impairments in tasks tapping complex executive functions, in particular functions involving planning/cognitive flexibility (Tower of Hanoi, WCST), monitoring (STM, sustained attention) and verbal reasoning (vocabulary and similarities from the WASI). Other tasks which are impaired, like the digit symbol task, also require a combination of complex skills including switching and monitoring (for similar impairments in higher order executive functions see Brumm et al., 2004; Smith et al., 1996; Nardecchia et al., 2015). Instead, we have found no impairments in tasks tapping inhibitory control consistent with previous results (Brumm et al., 2004; Channon et al., 2004, 2005). This pattern is the opposite of what is seen in children and adolescents where inhibitory control seems to be particularly affected, whereas working memory and planning are less affected (see DeRoche & Welsh, 2008). This may be explained by the different developmental

trajectory of these functions. Inhibitory control develops in childhood and peaks in early adolescence whereas planning and monitoring functions may continue to develop through adulthood (see Anderson, 2002; Best & Miller, 2010; Romine & Reynolds, 2005). It is possible that some functions, like inhibitory control, only show a developmental delay, but eventually catch up. Instead, other executive functions involving different skills (monitoring) or more complex skills (planning, reasoning) either plateau later on or, more likely, will continue to show deficits because they are more sensitive to impairment.

### Variability

We have shown *extreme* variability across participants. Only 2/37 = 5.4% of our PKU participants showed a *severe* impairment in terms of average performance (with overall performance being below 2 *SD* from the mean), but a much higher number showed an abnormal cognitive profile in terms of proportion of impaired measures (46%). Overall about 1/4 of our sample (24%) displayed a clear cognitive impairment (in terms of both average performance and cognitive profile), whereas 38% performed as well as the controls with *z* scores always = <.05. It is important to note that AwPKU with a normal versus impaired cognition significantly differed for metabolic control (Phe average in adolescence and adulthood and Phe variability across the lifespan), underscoring the importance of maintaining control at all ages.

### Conclusion

We have carried out a cognitive assessment of a sample of adults with PKU using an unprecedented number of tasks across domains. Our results show the importance of carrying out such a comprehensive assessment. This has allowed us to gather important information about the type, severity and distribution of cognitive impairments in AwPKU. Regarding type of impairments, we have found that even in a group of relatively well controlled AwPKU there are still impairments in complex executive functions and responses are slower across domains. Learning abilities, however, were well preserved and a reduction in speed was limited to cognitive tasks with no effects on peripheral motor speed. This indicates that not all cognitive functions are equally susceptible to the toxic effect of Phe. In terms of severity and distribution of impairments, our results show extreme variability with about 1/4 of AwPKU obtaining a completely normal cognitive profile, but about half showing significant impairments in at least some functions. These results highlight the success, but also the limitations of the clinical management received by these patients. A large number of AwPKU reach performance which is indistinguishable from their peers. This should be reassuring for them, their families and the professionals involved in their care. However, there are still impairments in a large number of AwPKU which may reduce their career potential and economic prospects. This variability calls for a better understanding of what causes these impairments and for better treatment. A possible cause of variability is, of course, dietary adherence and current and historical levels of Phe. We have seen that AwPKU with normal versus impaired cognitive performance differ significantly in level of metabolic control. Although our AwPKU showed on average a better metabolic control than that reported in previous studies, forthcoming European guidelines

or current American guidelines are even more stringent and could result in still better outcomes. Further studies should investigate how metabolic control at different ages may impact cognition and explore possible differences across cognitive domains (see Romani, Palermo, MacDonald, Limback, Hall, & Geberhiwot, 2017).

### References

- Albrecht, J., Garbade, S. F., & Burgard, P. (2009). Neuropsychological speed tests and blood phenylalanine levels in patients with phenylketonuria: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *33*, 414–421. <http://dx.doi.org/10.1016/j.neubiorev.2008.11.001>
- Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*, 17–42. <http://dx.doi.org/10.1007/s11065-006-9002-x>
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*, 71–82. <http://dx.doi.org/10.1076/chin.8.2.71.8724>
- Anderson, P. J., & Leuzzi, V. (2010). White matter pathology in phenylketonuria. *Molecular Genetics and Metabolism*, *99*, 3–9. <http://dx.doi.org/10.1016/j.ymgme.2009.10.005>
- Anderson, P. J., Wood, S. J., Francis, D. E., Coleman, L., Anderson, V., & Boneh, A. (2007). Are neuropsychological impairments in children with early-treated phenylketonuria (PKU) related to white matter abnormalities or elevated phenylalanine levels? *Developmental Neuropsychology*, *32*, 645–668. <http://dx.doi.org/10.1080/87565640701375963>
- Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationship between cognitive and behavioral measure of executive function in children with brain disease. *Child Neuropsychology*, *8*, 231–240.
- Baddeley, A. D. (1986). *Working memory*. New York, NY: Oxford University Press.
- Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and design fluency in patients with frontal lobe lesions. *Journal of the International Neuropsychological Society*, *7*, 586–596. <http://dx.doi.org/10.1017/S1355617701755063>
- Benton, A. L., Hamsher, K. de, S., & Sivan, A. B. (1994). *Multilingual aphasia examination*. Iowa City, IA: AJA Associates.
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641–1660. <http://dx.doi.org/10.1111/j.1467-8624.2010.01499.x>
- Bik-Multanowski, M., Pietrzyk, J. J., & Mozrzymas, R. (2011). Routine use of CANTAB system for detection of neuropsychological deficits in patients with PKU. *Molecular Genetics and Metabolism*, *102*, 210–213. <http://dx.doi.org/10.1016/j.ymgme.2010.10.003>
- Biomarin. (2016). *BioMarin Phase 3 Study of pegvaliase for phenylketonuria (PKU) meets primary endpoint of blood phenylalanine (Phe) reduction (p<0.0001)*. (Press release). Retrieved from <http://investors.bmrn.com/releasedetail.cfm?ReleaseID=961419>
- Blau, N., van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *Lancet*, *376*, 1417–1427. [http://dx.doi.org/10.1016/S0140-6736\(10\)60961-0](http://dx.doi.org/10.1016/S0140-6736(10)60961-0)
- Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*, *5*, 49–62. <http://dx.doi.org/10.1006/nimg.1996.0247>
- Brumm, V. L., Azen, C., Moats, R. A., Stern, A. M., Broomand, C., Nelson, M. D., & Koch, R. (2004). Neuropsychological outcome of subjects participating in the PKU adult collaborative study: A preliminary review. *Journal of Inherited Metabolic Disease*, *27*, 549–566. <http://dx.doi.org/10.1023/B:BOLI.0000042985.02049.ff>
- Channon, S., German, E., Cassina, C., & Lee, P. (2004). Executive functioning, memory, and learning in phenylketonuria. *Neuropsychology*, *18*, 613–620. <http://dx.doi.org/10.1037/0894-4105.18.4.613>

- Channon, S., Goodman, G., Zlotowitz, S., Mockler, C., & Lee, P. J. (2007). Effects of dietary management of phenylketonuria on long-term cognitive outcome. *Archives of Disease in Childhood*, *92*, 213–218. <http://dx.doi.org/10.1136/adc.2006.104786>
- Channon, S., Mockler, C., & Lee, P. (2005). Executive functioning and speed of processing in phenylketonuria. *Neuropsychology*, *19*, 679–686. <http://dx.doi.org/10.1037/0894-4105.19.5.679>
- Christ, S. E., Huijbregts, S. C. J., de Sonnevile, L. M. J., & White, D. A. (2010). Executive function in early-treated phenylketonuria: Profile and underlying mechanisms. *Molecular Genetics and Metabolism*, *99*, 22–32. <http://dx.doi.org/10.1016/j.ymgme.2009.10.007>
- Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. Unpublished doctoral dissertation, McGill University, Montreal, Quebec, Canada.
- Davis, A. S., & Pierson, E. E. (2012). The relationship between the WAIS-III digit symbol Coding and executive functioning. *Applied Neuropsychology: Adult*, *19*, 192–197. <http://dx.doi.org/10.1080/09084282.2011.643958>
- DeRoche, K., & Welsh, M. (2008). Twenty-five years of research on neurocognitive outcomes in early-treated phenylketonuria: Intelligence and executive function. *Developmental Neuropsychology*, *33*, 474–504. <http://dx.doi.org/10.1080/87565640802101482>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. <http://dx.doi.org/10.1146/annurev-psych-113011-143750>
- Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: A Frontal Assessment Battery at bedside. *Neurology*, *55*, 1621–1626. <http://dx.doi.org/10.1212/WNL.55.11.1621>
- Elliott, R. (2003). Executive functions and their disorders. *British Medical Bulletin*, *65*, 49–59. <http://dx.doi.org/10.1093/bmb/65.1.49>
- Feldmann, R., Denecke, J., Grenzschach, M., & Weglage, J. (2005). Frontal lobe-dependent functions in treated phenylketonuria: Blood phenylalanine concentrations and long-term deficits in adolescents and young adults. *Journal of Inherited Metabolic Disease*, *28*, 445–455. <http://dx.doi.org/10.1007/s10545-005-0445-7>
- Griffiths, P., Paterson, L., & Harvie, A. (1995). Neuropsychological effects of subsequent exposure to phenylalanine in adolescents and young adults with early-treated phenylketonuria. *Journal of Intellectual Disability Research*, *39*, 365–372. <http://dx.doi.org/10.1111/j.1365-2788.1995.tb00540.x>
- Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*, 464–482. <http://dx.doi.org/10.1016/j.cognition.2005.02.006>
- Janzen, D., & Nguyen, M. (2010). Beyond executive function: Non-executive cognitive abilities in individuals with PKU. *Molecular Genetics and Metabolism*, *99*(Suppl. 1), S47–S51. <http://dx.doi.org/10.1016/j.ymgme.2009.10.009>
- Joseph, B., & Dyer, C. A. (2003). Relationship between myelin production and dopamine synthesis in the PKU mouse brain. *Journal of Neurochemistry*, *86*, 615–626. <http://dx.doi.org/10.1046/j.1471-4159.2003.01887.x>
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*, 213–233. <http://dx.doi.org/10.1007/s11065-007-9040-z>
- Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (2000). *WCST-64: Wisconsin Card Sorting Test-64 card version, professional manual*. Odessa, FL: Psychological Assessment Resources.
- Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German–English comparison. *Cognition*, *63*, 315–334. [http://dx.doi.org/10.1016/S0010-0277\(97\)00005-X](http://dx.doi.org/10.1016/S0010-0277(97)00005-X)
- Latzman, R. D., & Markon, K. E. (2010). The factor structure and age-related factorial invariance of the Delis–Kaplan Executive Function System (D-KEFS). *Assessment*, *17*, 172–184. <http://dx.doi.org/10.1177/1073191109356254>
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*, 222–243. <http://dx.doi.org/10.1037/a0017619>
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. <http://dx.doi.org/10.1146/annurev.neuro.24.1.167>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. <http://dx.doi.org/10.1006/cogp.1999.0734>
- Möller, H. E., Weglage, J., Bick, U., Wiedermann, D., Feldmann, R., & Ullrich, K. (2003). Brain imaging and proton magnetic resonance spectroscopy in patients with phenylketonuria. *Pediatrics*, *112*, 1580–1583.
- Moyle, J. J., Fox, A. M., Arthur, M., Bynevelt, M., & Burnett, J. R. (2007). Meta-analysis of neuropsychological symptoms of adolescents and adults with PKU. *Neuropsychology Review*, *17*, 91–101. <http://dx.doi.org/10.1007/s11065-007-9021-2>
- Moyle, J. J., Fox, A. M., Bynevelt, M., Arthur, M., & Burnett, J. R. (2007). A neuropsychological profile of off-diet adults with phenylketonuria. *Journal of Clinical and Experimental Neuropsychology*, *29*, 436–441. <http://dx.doi.org/10.1080/13803390600745829>
- Nardecchia, F., Manti, F., Chiarotti, F., Carducci, C., Carducci, C., & Leuzzi, V. (2015). Neurocognitive and neuroimaging outcome of early treated young adult PKU patients: A longitudinal study. *Molecular Genetics and Metabolism*, *115*, 84–90. <http://dx.doi.org/10.1016/j.ymgme.2015.04.003>
- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain and Language*, *103*, 175–176. <http://dx.doi.org/10.1016/j.bandl.2007.07.102>
- Pietz, J., Dunkelmann, R., Rupp, A., Rating, D., Meinck, H. M., Schmidt, H., & Bremer, H. J. (1998). Neurological outcome in adult patients with early-treated phenylketonuria. *European Journal of Pediatrics*, *157*, 824–830. <http://dx.doi.org/10.1007/s004310050945>
- Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique [Psychological examination of traumatic brain injury]. *Archives de Psychologie*, *28*, 286–340.
- Rey, A. (1964). *L'examen clinique en psychologie* [Clinical examination in psychology]. Paris, France: Presses Universitaires de France.
- Ris, M. D., Williams, S. E., Hunt, M. M., Berry, H. K., & Leslie, N. (1994). Early-treated phenylketonuria: Adult neuropsychologic outcome. *The Journal of Pediatrics*, *124*, 388–392. [http://dx.doi.org/10.1016/S0022-3476\(94\)70360-4](http://dx.doi.org/10.1016/S0022-3476(94)70360-4)
- Romani, C., Di Betta, A. M., Tsouknida, E., & Olson, A. (2008). Lexical and nonlexical processing in developmental dyslexia: A case for different resources and different impairments. *Cognitive Neuropsychology*, *25*, 798–830. <http://dx.doi.org/10.1080/02643290802347183>
- Romani, C., Palermo, L., MacDonald, A., Limback, E., Hall, S. K., & Geberhiwot, T. (2017). The impact of phenylalanine levels on cognitive outcomes in adult patients with Phenylketonuria: Effects across tasks and developmental stage. *Neuropsychology*, *31*, 242–254. <http://dx.doi.org/10.1037/neu0000336>
- Romani, C., Tsouknida, E., & Olson, A. (2015). Encoding order and developmental dyslexia: A family of skills predicting different orthographic components. *Quarterly Journal of Experimental Psychology*, *68*, 99–128. <http://dx.doi.org/10.1080/17470218.2014.938666>
- Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology*, *12*, 190–201. [http://dx.doi.org/10.1207/s15324826an1204\\_2](http://dx.doi.org/10.1207/s15324826an1204_2)

- Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology*, 2, 135–146. <http://dx.doi.org/10.1080/01688638008403788>
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York, NY: Sage.
- Sahakian, B., Jones, G., Levy, R., Gray, J., & Warburton, D. (1989). The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *The British Journal of Psychiatry*, 154, 797–800. <http://dx.doi.org/10.1192/bjp.154.6.797>
- Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain: A Journal of Neurology*, 111, 695–718. <http://dx.doi.org/10.1093/brain/111.3.695>
- Sánchez-Cubillo, I., Periañez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, F. (2009). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuospatial abilities. *Journal of the International Neuropsychological Society*, 15, 438–450. <http://dx.doi.org/10.1017/S1355617709090626>
- Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews*, 35, 146–160. [http://dx.doi.org/10.1016/S0165-0173\(01\)00044-3](http://dx.doi.org/10.1016/S0165-0173(01)00044-3)
- Schmidt, E., Rupp, A., Burgard, P., Pietz, J., Weglage, J., & de Sonneville, L. (1994). Sustained attention in adult phenylketonuria: The influence of the concurrent phenylalanine-blood-level. *Journal of Clinical and Experimental Neuropsychology*, 16, 681–688. <http://dx.doi.org/10.1080/01688639408402681>
- Schmidt, M. (1996). *Rey Auditory Verbal Learning Test: A handbook*. Los Angeles, CA: Western Psychological Services.
- Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London*, 298, 199–209. <http://dx.doi.org/10.1098/rstb.1982.0082>
- Smith, M. L., Klim, P., Mallozzi, E., & Hanley, W. B. (1996). A test of frontal-specificity hypothesis in the cognitive performance of adults with phenylketonuria. *Developmental Neuropsychology*, 12, 327–341. <http://dx.doi.org/10.1080/87565649609540656>
- Stemerdink, N., Van Der Molen, M., Kalverboer, A., Van Der Meere, J., Huisman, J., De Jong, L., . . . Van Spronsen, F. (1999). Prefrontal dysfunction in early and continuously treated phenylketonuria. *Developmental Neuropsychology*, 16, 29–57. <http://dx.doi.org/10.1207/S15326942DN160103>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. <http://dx.doi.org/10.1037/h0054651>
- Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York, NY: Raven Press.
- Sundermann, B., Pfeleiderer, B., Möller, H. E., Schwindt, W., Weglage, J., Lepsien, J., & Feldmann, R. (2011). Tackling frontal lobe-related functions in PKU through functional brain imaging: A Stroop task in adult patients. *Journal of Inherited Metabolic Disease*, 34, 711–721. <http://dx.doi.org/10.1007/s10545-011-9318-4>
- Trites, R. (1977). *Grooved Pegboard Test*. Lafayette, IN: Lafayette Instrument.
- VanZutphen, K., Packman, W., Sporri, L., Needham, M., Morgan, C., Weisiger, K., & Packman, S. (2007). Executive functioning in children and adolescents with phenylketonuria. *Clinical Genetics*, 72, 13–18. <http://dx.doi.org/10.1111/j.1399-0004.2007.00816.x>
- Vockley, J., Andersson, H. C., Antshel, K. M., Braverman, N. E., Burton, B. K., Frazier, D. M., . . . Berry, S. A. (2014). Phenylalanine hydroxylase deficiency: Diagnosis and management guideline. *Genetics in Medicine*, 16, 188–200. <http://dx.doi.org/10.1038/gim.2013.157>
- War Department, Adjutant General's Office. (1944). *Army Individual Test Battery. Manual of directions and scoring*. Washington, DC: Author.
- Watanabe, K., Ogino, T., Nakano, K., Hattori, J., Kado, Y., Sanada, S., & Ohtsuka, Y. (2005). The Rey–Osterrieth Complex Figure as a measure of executive function in childhood. *Brain & Development*, 27, 564–569. <http://dx.doi.org/10.1016/j.braindev.2005.02.007>
- Wechsler, D. (1991). *The Wechsler Intelligence Scale for Children—Third ed. (WISC—III)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (3rd ed.)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.
- Weglage, J., Fromm, J., van Teeffelen-Heithoff, A., Möller, H. E., Koletzko, B., Marquardt, T., . . . Feldmann, R. (2013). Neurocognitive functioning in adults with phenylketonuria: Results of a long-term study. *Molecular Genetics and Metabolism*, 110(Suppl), S44–S48. <http://dx.doi.org/10.1016/j.ymgme.2013.08.013>
- Welsh, M. C., Pennington, B. F., Ozonoff, S., Rouse, B., & McCabe, E. R. (1990). Neuropsychology of early-treated phenylketonuria: Specific executive function deficits. *Child Development*, 61, 1697–1713. <http://dx.doi.org/10.2307/1130832>
- Wilkins, A. J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia*, 25, 359–365. [http://dx.doi.org/10.1016/0028-3932\(87\)90024-8](http://dx.doi.org/10.1016/0028-3932(87)90024-8)

Received June 20, 2016

Revision received September 29, 2016

Accepted October 30, 2016 ■