NEUROPSYCHOLOGICAL ASSESSMENT IN INHERITED METABOLIC DISEASES: TRACKING OUTCOMES IN PHENYLKETONURIA (PKU)

LUCINDA SANGEETA THOMAS

Doctor of Philosophy

ASTON UNIVERSITY

September 2022

Aston University

Neuropsychological Assessment in Inherited Metabolic Diseases: Tracking Outcomes in Phenylketonuria (PKU)

Lucinda Sangeeta Thomas

Doctor of Philosophy

September 2022

Whilst treating PKU early dramatically reduces impairments caused by excess Phe, individuals with early-treated PKU demonstrate mild impairments in cognition, well-being, neuropathology, and neurophysiology. Information in the current literature about how metabolic control can impact these impairments in adulthood, however, remains limited. Furthermore, there remains a lack of understanding about how these impairments may interact with age-related neuropathological and neurophysiological alterations as early-treated adults with PKU begin to reach older age. This thesis includes two empirical investigations and one systematic review. The first empirical investigation compared impairments exhibited by young adults with PKU and healthy older adults to extrapolate how impairments may interact as adults with PKU reach older age. This investigation identified key overlaps in the profiles of cognitive impairment demonstrated by each population, in particular in the domains of speed of processing and executive function. The second empirical investigation explored whether middle-aged adults with PKU demonstrate accelerated effects of ageing on neuropsychological function by comparing outcomes in this population with both age-matched controls and younger adults with PKU. Overlaps in impaired cognitive domains between young and middle-aged AwPKU were identified, as well as poorer well-being in the middle-aged cohort compared to age-matched controls. The severity of impairments compared to healthy controls of the same age, however, was less significant in middle-aged cohorts. In addition, no significant impact of adult metabolic control on outcomes was identified in the middle-aged AwPKU included in this study. Finally, a systematic review considered within-participant studies investigating the effects of manipulating metabolic control on cognitive, well-being, and neurological outcomes in children and adults with PKU. The review indicated that metabolic control has a significant impact on all outcomes in both children and adults with PKU, as well as demonstrating that impairments in these domains may be reversible through decreasing current Phe levels at any age.

Keywords: PKU, Phenylketonuria, cognition, well-being, neurophysiology, ageing, adulthood, treatment, metabolic control

For Nanny,
Thank you for being my biggest cheerleader,
and for never letting me stop believing in myself.

Acknowledgements

Thank you to my supervisors, Dr Cristina Romani, Dr Gemma Mansell, and Dr Jane Waite, for your guidance and feedback throughout this process. Cristina, thank you for believing in me, and for always pushing me to be the best researcher I can be. Gemma and Jane, you have both been pillars of support and advice during the final stages of this PhD, and I am incredibly grateful to have had the opportunity to work with you. I would also like to thank Dr Andrew Olson, whose statistical knowledge and input has been invaluable to this thesis. Thanks also to the clinicians, research nurses, and dieticians at University College Hospital London and Salford Royal Hospital who worked tirelessly to help with recruitment and patient engagement, despite the ongoing pandemic and its profound impact on all aspects of the NHS. A special thank you, as well, to all my participants, without whom this thesis would not have been possible. I would also like to acknowledge the Fund for Women Graduates, whose generous living grant allowed me to see my PhD through to completion.

My time at Aston would not have been the same without the support and friendship of my amazing office-mates, past and present. Thank you to everyone in SW506 for all the conversations, advice, and support you've offered over the last four years. I especially want to thank Rachel Green, Emma Alving-Jessep, Foyzul Rahman, Ulrike Senftleben, and Lily Hawkins for keeping me sane, lifting me up, and always being available for a coffee when I needed one. Thanks also to Niteen Mulji and Jake Diggins for always being available, whether to lend a helping hand when experiments misbehaved, or to just chat about life over a cup of tea.

To my family, I can't begin to thank you enough for everything that you have done, and continue to do, for me. Mum, your strength and tenacity are second to none. Thank you for showing me that we are the masters of our own fate, and for always being there to listen and support me, even when you were facing your own challenges. Dad, you are patience personified. Thank you for your endless encouragement, advice, and hugs. Alex, thank you for being my rock, and for always being ready to help me de-stress with a whiskey and a board game (whether virtually or in person). Sarah, you are a beacon of positivity in my life. Thank you for always being there to make me smile, and to brighten up my day. Finally, to Ross. Since I started my journey back into academia six years ago, you have been my friend, my partner, my team-mate, and my love. Thank you for the hours of video games and trashy tv, the amazing home-cooked meals, and for always being on my side. I couldn't have done this without you.

Table of Contents

List of Abbreviations

ABC	Aberrant Behaviour Checklist
ADHD(-RS)	Attention Deficit Hyperactivity Disorder (- Rating Scale)
AFQ	Automated Fibre-Tract Quantification
ASRS	Adult ADHD Self Report Scale
AwPKU	Adults with PKU
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
BH4	Tetrahydrobiopterin (6R-BH4)
BNT	Boston Naming Task
BRIEF	Behaviour Rating Inventory of Executive Function
CATS	Comprehensive Affect Test System
CNV	Contingent Negative Variation
CwPKU	Children with PKU
DTI	Diffusion Tensor Imaging
ES	Effect Size
FDOPA	Fluorodopa
HADS	Hospital Anxiety and Depression Scale
HPA	Hyperphenylalaninemia
LTM	Long Term Memory
NEMI	New Metric Scale of Intelligence
NMDA	N-methyl-D-aspartate
OC	Older Controls
PAH	Phenylalanine Hydroxylase
PAL	Paired Associate Learning
PET	Positron Emission Topography
Phe	Phenylalanine
PKU	Phenylketonuria
Ppt	Participant
QoL	Quality of Life
ROI	Region Of Interest
RT	Response/Reaction time
RVP	Rapid Visual Processing
STM	Short Term Memory
Sz	Schizophrenia
VABS	Vine Adaptive Behaviour Scale
WAIS(-R)	Weschler Adult Intelligence Scale (- Revised)
WCST	Wisconsin Card Sorting Task
WM	White Matter
WML	White Matter Lesion
WMSS	White Matter Severity Score
WRAT	Wide range Achievement Test
WS	West Syndrome
YC	Younger Controls

List of Tables

List of Figures

Chapter 1: General Introduction and Literature Review

Introduction

Phenylketonuria (PKU) is a metabolic disorder affecting between 1 in 10,000 to 1 in 12,000 births (Christ, 2003; Scriver & Kaufman, 2001). Individuals with PKU are lacking in the genetic component required to code for the enzyme phenylalanine hydroxylase (PAH) which is essential for the hydroxylation of phenylalanine (Phe) from the diet into tyrosine in the liver (Choo et al., 1979; Güttler & Lou, 1986). This partial to complete enzyme inactivity therefore results in a lack of Phe hydroxylation, leading to a lack of tyrosine in PKU. Tyrosine is a known precursor for dopamine; therefore the lack of hydroxylation has a significant effect upon frontal lobe dopamine levels (Albrecht et al., 2009). The interrupted hydroxylation process also causes an accumulation of Phe in the blood, known as hyperphenylalaninemia, as well as in the tissues and other body fluids. The resultant increased levels of Phe passing through the blood-brain barrier has then been found to have an impact upon white matter structures, leading to lesions and demyelination of multiple neurological structures (Blau, van Spronsen, et al., 2010; Surtees & Blau, 2000). Left untreated, PKU can cause significant neurological impairment including intellectual disability, microcephaly, seizures, and behavioural difficulties (Albrecht et al., 2009; Blau, van Spronsen, et al., 2010).

In the 1950s, a low-Phe diet was introduced which has been proven to reduce blood-Phe concentration in people with PKU (Scriver & Kaufman, 2001), and consequently to dramatically reduce the cognitive, neurological, and behavioural impairments associated with the disorder. Following a low-Phe diet is now universally prescribed to infants diagnosed with PKU, with Manta-Vogli et al. (2020) stating that “dietary management with restriction of Phe intake remains the mainstay therapy of these patients” (p.628). Whilst maintaining low Phe levels during childhood is universally agreed to be of utmost importance to allow neurodevelopment to occur as unimpaired as possible, once the critical stage for neurodevelopment has passed, many individuals with PKU choose to abandon their low-Phe diet (Bekhof et al., 2003; Bowersox & Panel, 2001; Crone et al., 2005; Olsson et al., 2007; Rocha & MacDonald, 2018), with the proportion of patients attending clinics found to decrease once patients reach 13 years of age, and Phe levels found to significantly increase in adolescents with PKU from the age of 15 (Berry et al., 2013; Walter et al., 2002; Walter & White, 2004). Studies investigating why so many adolescents abandon their diet have suggested that children following the diet often have to eat separately from friends and family, leading to a loss of ‘family-time’ communication and rapport, or bullying from other

children when eating in a school or social environment (MacDonald, 2000). Studies surveying patients with PKU, and their caregivers, have also identified time constraints, stress associated with food preparation and record-keeping, and the restrictions that following a diet impose onto patients' social lives as the key obstacles to following a low-Phe diet as children move into adolescence (Bilginsoy et al., 2005). Furthermore, decreased motivation may result from an absence of immediate effects of abandoning a low-Phe diet, with family members often noticing the negative impacts on a patient's mood, attention span, and health before the patient themself (Macleod & Ney, 2010). The potential repercussions of abandoning a low-Phe diet in later life remain a subject of much debate.

This thesis aims to better understand the impact of early-treated PKU on cognitive, behavioural, and neurological health in adulthood, and the effect that metabolic control can have on these domains. The relationship between ageing and PKU-related impairment is investigated, as well as the potential reversibility of neurological damage caused by high Phe levels. The results of our investigations increase our understanding of the potential interactions between older age and PKU, as well as providing evidence of the impact of resuming or abandoning a low-Phe diet on cognitive, neurological, and well-being outcomes throughout the lifespan. The relationship between early-treated PKU and middle-age, however, remains unclear, with further investigations required with this age group before conclusions can be drawn about the impact that moving from young adulthood to middle-age has on adults with PKU (AwPKU) with varying levels of metabolic control. This research is vital to help clinicians better understand the impact of early-treated PKU on neurological health in later life and to allow them to provide their patients with the most informed advice possible regarding the impact that lifestyle and metabolic control may have on cognitive and behavioural well-being in middle- and older-age.

PKU treatment diet

The prescription of a modern PKU treatment diet is described by the PKU Dietary Handbook (MacDonald et al., 2020) as having three main aims: 1) to ensure strict control of patients' natural protein and/or Phe intake, to prevent excessive accumulation of Phe in the blood, and therefore in the brain; 2) to replace natural proteins with safe, Phe-free substitutes (both synthetic and naturally-occurring), and amino-acid supplements; 3) to the ensure that the treatment diet contains a balanced intake of necessary nutrients and energy, using vitamin and mineral supplements.

A standard low-Phe diet restricts patients' protein/Phe intake by at least 75% when compared to a regular diet. Phe comprises approximately 4-6% of all dietary protein-containing foods (Manta-Vogli et al., 2020), therefore such a significant reduction of Phe intake requires patients to avoid or limit all high protein foods, including but not limited to: meat, fish, animal milk and derivatives, cereals, Quorn, soya, pulses, gelatine, plant algae, aspartame etc. Regular foods which are very low in protein (<=25mg/100g Phe), such as vegan cheese, butter, sugar, and jams, can be eaten without measurement. Most fruits and vegetables (excluding potatoes), however, are slightly higher in Phe levels (<=75mg/100g) and therefore may still contribute to patients' daily Phe but shouldn't affect their blood Phe control (MacDonald et al., 2020). Other low-protein foods, such as potatoes and cereals, can also be consumed, but only in very restricted amounts that are limited to <50mg/100g of Phe (Manta-Vogli et al., 2020). Infants with PKU are treated with a low-Phe diet by restricting their protein intake through natural breast milk and regular formula, supplementing this instead with a specifically designed Phe-free infant formula.

As the regular foods available to patients following a low-Phe diet are significantly limited, a number of supplements and substitutes are required to encourage natural growth (Rocha et al., 2013), prevent protein and tyrosine deficiencies in children with PKU (van Spronsen et al., 2017), optimise metabolic control (MacDonald et al., 2003), block Phe transport across the blood-brain barrier (Pardridge, 1998), and help prevent neurological disability (Blau et al., 2011). Patients can also supplement their diet with synthetic low-protein foods, such as breads and pasta, which can provide up to 50% of their required energy intake (Pena et al., 2015). These synthetically manufactured foods, however, are often quite expensive, and may not be an option for patients whose access to them is not covered by government or insurance funding (MacDonald et al., 2020).

A low-Phe diet can be followed in one of two ways: 1) by prescribing a patient's total daily allowance of Phe and calculating all Phe-containing food into this allowance; or 2) by using a Phe exchange system wherein patients are given weighted portions of different foods, all of which contain the same amount of Phe (Ahring et al., 2009). Individuals with PKU can then use this low-Phe diet to reduce their blood-Phe concentrations to target levels set by their clinicians. Target blood-Phe levels vary across different countries. The current European PKU guidelines recommend target blood-Phe levels of 120-360µmol/L in children under 12 years and 120-600µmol/L from 12 years onwards (Van Wegberg et al., 2017). To reach these target levels, most people with PKU are able to tolerate less than 500mg of dietary Phe per day (MacDonald & White, 2015). The Phe tolerance of people with PKU, however, varies both between and within individuals, and can be impacted by a number of

factors, including the severity of their PKU (e.g. classical vs. mild PKU), their dosage, adherence, daily distribution of protein substitutes, their use of accompanying pharmaceutical treatments, growth rates, pregnancy, and catabolism (i.e. protein breakdown) during periods of illness. The severity of a low-Phe diet, therefore, can vary from person to person, based upon each individual's Phe tolerance. Dietary Phe intakes are often initially prescribed based upon patients' diagnosis, with patients with classical PKU (pre-treatment blood-Phe levels of >1200µmol/L) prescribed a Phe intake of <20mg/day, patients with moderate PKU (pre-treatment blood-Phe levels of 200-1200µmol/L) prescribed an intake of 20-25mg/day, patients with mild PKU (pre-treatment blood-Phe levels of 600-900µmol/L) prescribed an intake of 25-50mg/day, and patients with mild hyperphenylalaninemia (HPA; pre-treatment blood-Phe levels of <600µmol/L) not usually requiring any dietary Phe restriction (Manta-Vogli et al., 2020). Clinicians will often try to maximise patients' Phe intake based upon their observed Phe tolerance, in an attempt to make the diet more acceptable to patients by reducing the practical and social burdens associated with it, and increasing the nutritional benefits gained through consumption of natural, vs. synthetic, protein intake (MacDonald et al., 2020).

As touched upon earlier, as well as dietary intervention, pharmaceutical interventions are available for use in tandem with a low-Phe diet to decrease blood-Phe levels in people with PKU. One such intervention currently available in the USA and Europe, and to children with PKU in the UK, is sapropterin dihydrochloride (Kuvan®, BioMarin Pharmaceutical Inc., Novato, CA, USA), a synthetic formulation of tetrahydrobiopterin (6R-BH4), commonly referred to as 'BH4' (Feillet et al., 2014; Muntau et al., 2019). BH4 is an essential co-factor of the phenylalanine hydroxylase (PAH), tyrosine hydroxylase, and tryptophan hydroxylase enzymes, meaning that all three of these enzymes require BH4 for catalytic activity to occur (Manta-Vogli et al., 2020). Between 20% and 60% of people with PKU are thought to be BH4-responsive, meaning that oral administration of BH4 in these individuals will result in a decrease in blood-Phe levels of 30% or more within 8-14hrs (Blau, 2013; Blau, Hennermann, Langenbeck, & Lichter-Konecki, 2011; Steinfeld et al., 2002; Trefz, Burton, Longo, Casanova, Gruskin, Dorenbaum, Kakkis, Crombez, Grange, Harmatz, & Et Al., 2009). Dietary Phe intake in BH4-responsive patients taking supplements, therefore, can be increased by up to 4x whilst still maintaining target blood-Phe levels in some individuals, although others have reported reaching target Phe levels after an increase in dietary Phe of less than 2x baseline intake levels (Singh, Jurecki, & Rohr, 2008; Trefz, Burton, Longo, Casanova, Gruskin, Dorenbaum, Kakkis, Crombez, Grange, Harmatz, Lipson, et al., 2009).

Parental control of children's nutrition often means that dietary adherence is good in infants and children with PKU (Blau, Bélanger-Quintana, et al., 2010). As children with PKU reach adolescence, however, adherence often begins to waver, leading to decreased metabolic control in this age group (Crone et al., 2005; Walter et al., 2002; Walter & White, 2004). Whilst pharmaceutical treatments such as BH4 are able to reduce some of the restrictions placed on individuals following a low-Phe diet, most BH4-responsive patients are those with mild PKU or HPA, who retain some residual PAH enzyme activity (Fiege & Blau, 2007). Responsiveness in patients with classical PKU, who have little to no residual PAH activity, remains low (Blau, 2013). Furthermore, there remains a lack of uniformity across countries with regards to BH4-responsiveness testing and availability guidelines (Muntau et al., 2017). Even in instances where pharmaceutical supplementation is available, many adults with PKU find dietary adherence particularly difficult to maintain, citing reasons such as social pressures, family cohesion issues, lack of knowledge of the disease, and lack of reimbursement for supplements as the cause of dietary abandonment (Bekhof et al., 2003; Bowersox & Panel, 2001; Crone et al., 2005; Olsson et al., 2007; Rocha & MacDonald, 2018).

Impact of PKU and ageing on cognition

Whilst the PKU treatment diet has been proven to protect people with PKU from profound intellectual disability, early-treated individuals with PKU have been found to still demonstrate some level of impairment in a range of cognitive domains. Research is still required, therefore, to better understand how cognition is impacted in both adults and children with PKU who have received treatment from birth, and how it may be affected by metabolic control throughout the lifespan.

A recent review of cognition in PKU was conducted by Canton et al. (2019) considering performance of early-treated children and adolescents with PKU on measures of IQ, perception, visuo-spatial processing, visual construction, motor skills, memory, executive functions, and processing speed across 54 studies. The ages of participants included in this review varied significantly, ranging from 5 to 20 years. Cognitive measures, however, were largely consistent across studies, with a variety of around three to four different assessments in each domain. Deficits were reported in the domains of IQ, visuo-spatial attention, motor tasks, memory, and speed of processing, with mixed findings emerging in language performance, and no impairment observed in measures of visuo-spatial and visual construction abilities. Hofman et al. (2018), meanwhile, reviewed cognitive performance in early-treated AwPKU aged 18+ years across 22 studies and reported impairments in speed

of performance across tasks, as well as deficits in the domains of sustained attention, working memory, visuo-spatial attention, and some measures of executive function across the same assessment measures as reported by Canton et al., plus additional measures that were not included in Canton's review. A similar review of cognition in early-treated AwPKU was carried out by Burlina et al. (2019) who also reported significant deficits in executive function, speed of processing, and attentional outcomes across 28 studies, using a similar range of assessment measures to Hofman et al.'s (2018) reviewed studies. It is important to note, however, that Burlina and Hofman's reviews included 13 of the same studies, therefore a significant overlap in findings is to be expected.

The effect of high blood-Phe levels on cognition during childhood has also been well documented, with multiple between-participant studies establishing an increasing impairment with decreasing metabolic control in cognitive domains including sustained attention, executive function, speed of processing, IQ and visuo-motor coordination in CwPKU (e.g. Azen et al., 1991; Holtzman et al., 1986; Huijbregts, de Sonneville, van Spronsen, et al., 2002; Jahja et al., 2013; Leuzzi et al., 2004; Schmidt et al., 1996; Smith et al., 1990; Waisbren et al., 2007). Findings of increased cognitive impairment in CwPKU with mean blood-Phe levels above 620µmol/L, compared to those with mean blood-Phe levels of 250µmol/L and below, have been reported (Schmidt et al., 1996) as well as poorer performance on the Mental Development Index from CwPKU with blood-Phe levels above 360µmol/L during their first year of life, compared to those with levels below 360µmol/L (de la Parra et al., 2017).

Metabolic control has also been found to influence cognition in adulthood (e.g. Bik-Multanowski & Pietrzyk, 2011; Brumm et al., 2004; Jahja, van Spronsen, et al., 2017; Palermo et al., 2017). Burlina et al. (2019) systematically reviewed 30 studies investigating the relationship between metabolic control and cognition in AwPKU in terms of performance on executive function and attention tasks. Included papers demonstrated deficits in sustained and selective attention, as well as working memory, inhibition, and verbal fluency in comparison to healthy controls. Despite variation in findings across studies, an overall link between concurrent blood-Phe levels and executive function, attentional control, inhibition, and some memory abilities was identified, with AwPKU who maintained good adulthood dietary control performing better across tasks in these domains than those with poor control. Working memory was also found to be affected by metabolic control in childhood, however the relationship between dietary control and working memory in adulthood was conflicting across included studies.

Beyond the need to increase our understanding of the role of metabolic control in mantaining cognitive health in PKU, a further key consideration for research is the notable similarity between the profile of cognitive impairment observed in young AwPKU and that observed in healthy adults as a natural product of ageing. Older adults have been found to demonstrate impairments in the domains of visuo-spatial attention (Kramer et al., 1996), processing speed (Bucur & Madden, 2010; Der & Deary, 2006), visuo-motor coordination (Hamilton et al., 2018; Volkow et al., 1998) and executive function (De Beni & Palladino, 2004; Delaloye et al., 2009), suggesting that there may be some commonalities between the mechanisms underlying cognitive impairment in these two population groups. The following narrative literature review considers evidence regarding the impact of PKU and ageing on performance across cognitive domains.

Intelligence

One key cognitive domain affected by PKU in childhood is intelligence. Studies investigating IQ impairments in CwPKU have suggested that impairments are significantly mitigated by metabolic control at this stage in life, with negative outcomes found to occur once blood-Phe levels reach somewhere between 800µmol/L and 1300µmol/L (Azen et al., 1991; Holtzman et al., 1986; Smith et al., 1990; Waisbren et al., 2007). The relationship between metabolic control and IQ in early-treated CwPKU was demonstrated by Smith et al. (1990) whose regression analyses indicated that age at start of treatment, average lifetime Phe levels, and duration of blood-Phe levels <120µmol/L are all significant predictors of IQ at 4 years old. Meanwhile, Holtzman et al. (1986) found that the age at which dietary control was lost (defined as blood-Phe levels reaching 1326µmol/L) was a significant predictor of the magnitude of differences between CwPKU aged 8 to 10 years and their unaffected family members (Aged 8: $r=0.51$, $p<.001$; Age 10: $r=0.47$, $p<.001$). Between-participant t-test comparisons of CwPKU also found that IQ scores increased progressively as dietary control lasted longer, with the greatest differences between CwPKU and unaffected siblings found in children who lost dietary control the earliest (control lost<= 71months, $p<.001$), and the smallest differences in those who maintained control the longest (control lost >=96 months, p=non-significant).

Metabolic control in childhood has also been found to impact intelligence in AwPKU. A review by Burgard (2000) suggested that IQ is strongly affected by dietary control in early childhood, with each 300µmol/L increase in blood-Phe levels resulting in IQ decreasing by approximately half a standard deviation. A review by Waisbren et al. (2007) further reported significant correlations between concurrent IQ and both mean blood-Phe levels within the

first 12 years of life, and concurrent Phe. This review found that each increase in blood-Phe level of 100µmol/L during this age range was predictive of an average decrease in IQ of 1.9 to 4.9 points (within the range of 394-666µmol/L covered by the included studies) while each 100µmol/L increase of concurrent blood-Phe levels predicted a drop in IQ of 0.5 to 1.4 points. Burlina et al. (2019) found no such impact of dietary control at time of testing or current Phe levels on IQ in AwPKU, when comparing reports of IQ across 47 papers. A relationship between adult IQ and childhood and adolescent blood-Phe levels was identified, however, suggesting that, whilst dietary control in early life plays a key role in IQ development, intelligence remains stable once adulthood is reached, regardless of metabolite levels.

Whilst a slight decrease in IQ is commonly associated with PKU, findings that vocabulary appears to remain unaffected, or in some cases to even improve with age, suggest that crystallised intelligence, at least, is not negatively impacted by increasing age. A study by Kaufman and Horn (1996) used the Kaufman Adolescent and Adults Intelligence Test (Kaufman & Kaufman, 1993) to investigate the effects of ageing on fluid processing and crystallised knowledge in 1,500 participants aged 17 to 94 years. Fluid intelligence was assessed by collating scores from the Rebus Learning, Logical Steps, and Mystery Codes tasks, whilst crystallised intelligence scores were a combination of Auditory Comprehension and Double Meanings scores. Covariance analyses were used to ensure that any impact of education and gender were accounted for. ANOVA analyses of results from this study found a main effect of age on both fluid and crystallised intelligence scores, however the pattern of decline differed significantly between the two. Crystallised intelligence was found to increase slightly from late adolescence to middle age, and then to be maintained through older adulthood until around 70 years, after which it began to decline. Fluid intelligence scores, however, peaked with the age group 20-24, then dropped and continued to decline steadily from age group 25-29 through to group 50-54 years. After 50-54 years, the decline became more rapid, with performance dropping through the remainder of the lifespan.

Similar findings regarding the impact of age on intelligence have been found by studies using the WAIS measures for Verbal and Performance IQ. Averages for Verbal IQ have been reliably found to increase up to 60 years of age, before plateauing and then gradually dropping towards the end of the lifespan, whilst Performance IQ averages are found to peak in the early 20s and then drop steadily through the remainder of adulthood (Botwinick, 1977; Dixon et al., 1985; Matarazzo, 2013).

Executive function

Executive function refers to a set of top-down cognitive skills that play a key role in setting and maintaining goals, resisting temptations and/or primed responses, and preserving concentration in situations where 'automatic' functioning needs to be overridden (Miller & Cohen, 2001). Miyake et al. (2000) established that there are three core domains of executive function; inhibition (including inhibitory control and interference control), working memory (including updating), and cognitive flexibility (including task switching and set shifting).

In addition to impaired intelligence, a comparison of 33 studies investigating cognitive function in both early-treated children and adults with PKU vs. healthy controls found significant differences in performance on measures of executive function (DeRoche & Welsh, 2008). These differences were found to be more marked in measures of executive function than in tests of IQ, suggesting a greater impairment in executive function in PKU than in overall intelligence. Moderate effect sizes were produced for tests of working memory and planning, whilst larger effect sizes were evident in tasks requiring flexibility and inhibition. The heterogeneity found in effect sizes across measures of executive function suggest that PKU affects some areas of executive function more significantly than others. This variance, however, was also found to be moderated by the type of test used, with significant heterogeneity in outcomes found across the nine different tools being employed to measure executive function. The variance in impairment in this area, therefore, could reflect test sensitivity, rather than necessarily reflecting differential levels of impairment within the domain of executive function.

A study by Antshel and Waisbren (2003) found no differences between 46 CwPKU and healthy controls on measures of IQ but did find impairments in tasks sensitive to executive functions including word learning and semantic clustering, both thought to be related to executive function. This cohort of CwPKU also demonstrated impaired performance in the Rey-Osterreich Figure Copying and Stroop tasks, also found to measure executive function in terms of inhibition and planning and organisation. This study further found that parents of CwPKU reported more difficulties with executive function related tasks, such as working memory, self-monitoring, and planning and organising, than parents of control children. Correlation analyses found that executive function performance and general processing speed were significantly correlated with concurrent Phe levels. Meanwhile, Janos et al. (2012) found that performance on executive function tasks in 42 CwPKU was significantly mediated by processing speed and variability, however, once these impairments

were controlled for, working memory and inhibitory control deficits remained apparent in this cohort as compared to performance of matched controls.

Studies investigating the impact of PKU on executive function in early-treated AwPKU have found tasks requiring planning and task-switching abilities, such as the Wisconsin Card Sorting Task and the Tower of Hanoi, are significantly impaired (Brumm et al., 2004; Nardecchia et al., 2015; Romani et al., 2018), however, no reliable impairment has been found in measures of inhibition such as Stroop and Trail Making B tasks (Jahja, Huijbregts, et al., 2017; Jahja, van Spronsen, et al., 2017; Moyle et al., 2006; Palermo et al., 2017). Meanwhile, studies investigating working memory have frequently reported poorer accuracy in working memory tasks from AwPKU compared to controls (Bartus et al., 2018; Bik-Multanowski et al., 2011; Brumm et al., 2004; Channon et al., 2004; Jahja, van Spronsen, et al., 2017; Palermo et al., 2017), whilst concurrent Phe levels have also been found to result in increased RTs in measures of working memory such as feature integration and memory search tasks (Bik-Multanowski et al., 2011; Jahja, Huijbregts, et al., 2017).

A similar pattern of impairment has also been observed in older adults. The impact of ageing upon executive function was investigated by Bucur and Madden (2010), using verbal fluency, Trail Making, Stroop and digit symbol coding tasks. Participants were split into 3 groups: younger adults (aged 19-31ys, N=47), middle-aged adults (aged 41-58yrs, N=44), and older adults (60-79yrs, N=43). Task completion times were found to significantly increase across each adjacent age group, suggesting a continual slowing throughout adulthood, however a significant main effect of age upon composite processing speed scores (comprised of reaction time data from Digit Symbol Coding, Stroop, and Trail-Making A tasks) was only evident between older adults and younger adults. A significant main effect of age upon composite executive function scores (comprised of Stroop interference score, Trail-Making B-A, and verbal fluency scores) was also observed between younger adults and both middle-aged and older adults, suggesting a drop in executive function abilities, specifically in middle-age. It is key to note, however, that both digit symbol coding and Trail making tasks also require a significant visuo-motor control component, therefore a decline in fine visuo-motor control with age may also have contributed to the age differences observed in this study.

A meta-analysis by Verhaeghen and De Meersman (1998) examined 20 studies comparing older adults (aged >=60 years) and younger adults (aged <=30 years) on Stroop task performance. Their analyses found no significant differences in Stroop interference effect between age groups, beyond the general slowing of RTs that is normally associated

with ageing. The lack of any moderating impact of Stroop interference on RTs between groups indicates a lack of impact of ageing on inhibition, supported by the absence of significant differences found between middle-aged and older adults in performance on the Flanker Inhibitory Control and Attention Test by Hamilton et al. (2017). Similar results were found in a study conducted by Delaloye et al. (2009), with 81 participants in groups aged 20-34yrs (N=21), 35-47 (N=21), 50-64yrs (N=21), and 65+ (N=18). This study found no linear relationship between age and performance on measures of inhibition (Stroop and Hayling tasks), but a significant age-related decline on task-switching measures (number-letter shifting and global-local shifting tests) and updating measures (letter updating and cube updating tasks). It is key to note that affected tasks in this study were specific to the visuo-spatial domain, and required a working memory component, therefore it is possible that task-switching and updating abilities may only be impacted in circumstances where concurrent memory and visuo-spatial processing is required.

Speed of processing

Beyond the impact of PKU on IQ and executive function, a key area found to be impaired in this population is speed of processing (Moyle, Fox, Arthur, et al., 2007; Ris et al., 1994; Weglage et al., 2013). Huijbregts, de Sonneville, Licht, et al. (2002) reported impaired speed of processing in early-treated CwPKU, with speed of response found to improve as participants' Phe levels decreased. Meanwhile, Hofman et al. (2018) carried out a systematic review of 16 studies which compared performance of AwPKU to that of either healthy controls (n=14 studies) or standardised/normative data (n=2 studies). This review reported that impairments in AwPKU were primarily found in measures of speed rather than accuracy, with simple reaction time task performance found to be specifically poor in off-diet participants. Heterogeneity across studies, however, limits the reliability of this review, as cut-off levels for "high" vs. "low" Phe groups varied across studies, and no clear definitions of what "off-diet" means were available.

Similarly, a meta-analysis of 11 neuropsychological studies investigating cognitive function in early-treated adults and adolescents with PKU compared to that of healthy controls, found differences between groups to be the largest in measures of speed of processing, including simple and choice reaction time tasks, speed trail making, continuous performance RTs, and dot-pattern completion times (Moyle, Fox, Bynevelt, et al., 2007). The attribution of slowed RTs to a specific deficit in processing speed, however, remains contentious, as a number of studies have failed to find PKU-related impairments in

measures of simple processing speed including the Stockings of Cambridge (Bartus et al., 2018), simple detection and detection with distractors (Palermo et al., 2017).

Studies with healthy older adults have reported similar findings, with this cohort of participants also demonstrating slowed response times across a number of different tasks. Der and Deary (2006) reanalysed data from 7,130 British adults (aged 18-94yrs) who completed a simple reaction time (simple RT) and a choice reaction time (choice RT) task assessed as part of the United Kingdom Health and Lifestyle Survey conducted between 1984 and 1985. When this data was plotted on a line graph, there was clear evidence of slower responses to simple RT tasks from the age of 50 onwards. Choice RT, meanwhile, was found to both slow and become more variable throughout the ageing process.

Sliwinski and Buschke (1999) conducted a longitudinal study with 302 older adults (aged 66-92yrs) who were assessed across the domains of memory, verbal fluency, processing speed, verbal comprehension, and short-term memory span. A cross-sectional comparison of results at baseline found significant age effects across domains, with only verbal comprehension spared. An accelerating decline with increasing age was further found in performance on sentence repetition, letter fluency, digit symbol coding, and vocabulary tasks, with slowed RTs with age across tasks indicating an increasingly severe decline in speed of processing in these domains with age. Notably, when baseline results were compared longitudinally within-participants, age effects were found to be larger than those observed in cross-sectional analyses, with significant age effects apparent across all cognitive domains. Furthermore, processing speed at baseline was found to be a significant predictor of performance across all cognitive domains at baseline, whilst within-participant changes on measures of processing speed were a significant predictor of changes across all other measured domains, with the exception of verbal comprehension. It is important to note, however, that processing speed had a considerably smaller mediating effect on overall longitudinal age effects than on cross-sectional age effects. This suggests that age alone may not impact processing speed as dramatically as cross-sectional studies may suggest. Age was, in fact, not the only impacting factor on scores across domains, as level of education also had a significant effect on all scores, with higher education associated with better performance on tasks.

Visuo-spatial attention

Slowed processing has been found to be noticeably evident in tasks tapping visuo-spatial attention in AwPKU, including choice reaction time and visual search tasks. Romani

(2018) found no significant differences between early-treated AwPKU and control participants in terms of accuracy on tasks tapping visuo-spatial attention, such as simple detection, choice reaction, detection with distractors, and visual search tasks. AwPKU, however, were significantly slower than controls across all tasks, supporting some level of impaired processing speed in tasks requiring visuo-spatial attention. Interestingly, this study found no difference in RTs in simple detection tasks with and without distractors, with peripheral motor speed appearing to be unaffected in PKU. Choice RTs, however, were significantly slower in AwPKU, indicating that the increased cognitive burden of making a choice in these tasks had a significant impact upon visuo-spatial processing speeds. The lack of accuracy impairment in visual search tasks suggests that slowed RTs are likely due to a specific processing speed deficit in the domain of visuo-spatial attention, rather than being due to an impairment in general visuo-spatial processing ability. Brumm et al. (2004) assessed performance on the block design, picture arrangement, and picture completion subtests of the WAIS-R, and found no evidence of impaired visuo-spatial ability in AwPKU. This study further found no effect of concurrent Phe on performance in visuo-spatial processing tasks, a finding supported by Nardecchia et al. (2015) and Ris et al. (1994).

CwPKU have been found to be unimpaired in visual search tasks, with no correlations apparent between concurrent Phe and visuospatial abilities in this cohort (e.g. Antshel & Waisbren, 2003; Leuzzi et al., 2004).

Similarly to AwPKU, age has been found to have a particularly pronounced effect on RT performance in the domain of visuo-spatial ability, in particular with regards to visuo-spatial attention, as measured through visual search tasks. It has been suggested, however, that poor performance in tasks measuring these abilities in older adults could be due to a combination of a decrease in processing speed and poorer visual acuity in this population rather than due to specific impairments in these cognitive domains (Salthouse, 1996). However, whilst both RTs and task accuracy have been found to be negatively impacted by age in conjunction visual search tasks (Harpur et al., 1995; Kramer et al., 1996), accuracy in feature search tasks appears to remain unaffected by age (Folk & Lincourt, 1996; Foster et al., 1995; Greenwood & Parasuraman, 1999; Plude & Doussard-Roosevelt, 1989). This suggests that poor performance on visuo-spatial tasks does not result primarily from a deficit in visual acuity, however nor does it result from a deficit in processing speed alone. Specifically impaired performance in conjunction search tasks suggests that the extra requirement for executive function in this task, such as monitoring and updating, in combination with visuo-spatial ability, may be responsible for the observed impairment in conjunction search performance.

Sustained attention

The performance of early-treated CwPKU and AwPKU on measures of sustained attention have consistently been found to be impaired across a number of studies (e.g. Burgard et al., 1997; Channon et al., 2004; Channon et al., 2007; Jahja, Huijbregts, et al., 2017; Palermo et al., 2017; Weglage et al., 2013) with low scores on concentration tasks reported in CwPKU, and sustained attention found to increase in CwPKU as Phe levels decrease (Holtzman et al., 1975; Huijbregts, de Sonneville, Licht, et al., 2002).

Brumm et al. (2004) looked at performance on a continuous performance task in early-treated AwPKU and found no impairment in accuracy scores when compared to normative data. However, AwPKU with high concurrent Phe levels were found to perform significantly poorer than those with low concurrent Phe on attentional measures. Romani et al. (2017), meanwhile, reported significant correlations between concurrent Phe and sustained attention performance in the same group of AwPKU (r=0.49), supporting arguments for the impact of Phe levels on sustained attention suggested by previous studies (e.g. Bik-Multanowski et al., 2011; Jahja, van Spronsen, et al., 2017). A recent study by Palermo et al. (2017), however, measured sustained attention using a rapid visual information processing (RVP) task in AwPKU with an average Phe level below 850μmol/L and found significant accuracy impairments in a sample of 37 early-treated AwPKU when compared to 30 healthy controls, suggesting that poor metabolic control may not be the only cause of sustained attention impairments in PKU.

Contrary to findings in AwPKU, recent studies have suggested that sustained attention appears to be spared, or even improved, in older compared to younger adults. Robison et al. (2022) measured speed of processing and sustained attention in 60 younger (mean age 19.9), and 62 older (mean age 75.4) adults using the Psychomotor Vigilance Task (PVT) in combination with thought probes (in which participants were asked to "characterise [their] current conscious experience") and pupillometry. Thought probes found fewer task-unrelated thoughts in older compared to younger participants, as well as increased levels of motivation and alertness, and larger task-evoked pupillary responses which remained stable throughout the task. Additionally, whilst older adults were overall slower in their responses than younger adults, their RTs also remained stable throughout the task, rather than slowing over time as demonstrated by the younger cohort. Taken in combination, these findings suggest that older adults' sustained attention was superior to that of younger adults in this study.

This finding was supported by a meta-analysis of age differences in performance on the go-no-go Sustained Attention to Response Task (Vallesi et al., 2021) which found that older adults performed overall more slowly than younger adults on this task in both go and no-go conditions, but demonstrated fewer no-go errors, suggesting a more prudent allocation of attentional resources in the older cohort. Some negative impact of increasing age on sustained attention, however, has been demonstrated by McAvinue et al. (2012), who assessed 133 participants aged 12 to 75 on a short continuous performance task. This study found that, whilst attentional capacity demonstrated a significant decline with age, sustained attention and attentional selectivity demonstrated a relatively small decline, with attentional control appearing to be relatively preserved in older adults (although still demonstrating some level of decline).

Visuo-motor coordination

Visuo-motor coordination has also been found to be significantly impacted by PKU in both child and adult participants. A study of 20 CwPKU found these participants performed significantly more poorly than matched controls in fine motor subtests of the Motor Performance Series (Weglage et al., 1995). This study further found that poorer performance in tasks requiring fast and precise fine motor movements were significantly correlated with high blood-Phe levels in children with PKU. Similarly, Romani et al. (2019) found significant differences between performance of 56 AwPKU and 30 healthy controls on measures of visuo-motor coordination and fine motor control, including the grooved pegboard, digit symbol coding, and trail making A tasks. Significant correlations were also found between pegboard and digit symbol coding performance, and concurrent blood-Phe levels in 37 early-treated AwPKU (Romani et al., 2017). Finally, Moyle, Fox, Arthur, et al. (2007)'s meta-analysis compared motor control performance between 104 AwPKU and 100 matched controls across three studies. Motor control was measured using the Peg Transfer Task (Griffiths et al., 1995), a finger motor speed exercise (Pietz et al., 1995), and a task requiring participants to touch visual targets quickly and accurately on a computer screen (Luciana et al., 2001). Analysis of performance across these three studies found an overall Hedge's g effect size of 0.35, indicating a notably poorer performance on these tasks from AwPKU compared to matched controls.

Declining visuo-motor coordination has further been observed in ageing populations, with older adults again demonstrating a similar pattern of impairment to AwPKU. A meta-analysis of differences in performance on the Digit Symbol Substitution subtest of the WAIS and WAIS-R between younger and older adults was conducted by Hoyer et al. (2004)

including data from 141 studies, with 3,731 younger adults (aged <30 years), and 3,876 older adults (aged >60 years). A significant predictive effect of age group on digit symbol completion times was found, and this effect was further found to be independent of education level (overall model R^2=.85, predictive value of age β=-0.95, p<.05). Similarly, a study by Joy et al. (2000) assessed motor control in 177 older adults (aged 55-90yrs) and found significant main effects of both age and education on digit symbol and symbol copy scores, with response times increasing as age increased and years of education decreased. Finally, a more recent study by Hamilton et al. (2017), measured motor coordination in 25 middle-aged (44-58 years) and 25 older adults (66-80 years) using the 9-hole pegboard and grooved pegboard tests. They found that older adults took significantly longer to complete both tasks, with older participants taking on average 13.3% longer than middle-aged participants to complete the 9-hole pegboard, and 25.7% longer to complete the grooved pegboard. Completion times were also found to be significantly correlated with age.

Language

Language abilities in CwPKU have been reported to be impacted to some extent, with Koch et al. (1982) finding significantly lower WRAT reading scores as well as lower spelling scores in 50 CwPKU when compared to their matched sibling controls. When participants were split into on-diet and off-diet groups, however, significant differences in reading and spelling were only apparent between off-diet CwPKU and their sibling controls, with on-diet CwPKU performing at equivalent, or improved, levels compared to their siblings, indicating that language may be unimpacted by PKU when treatment adherence in childhood is good.

Language in AwPKU has been found to be unimpaired in terms of accuracy in spelling, reading, and vocabulary tasks (Brumm et al., 2004; De Felice et al., 2018). Verbal IQ scores, however, have been found to be impaired in AwPKU, in particular in verbal reasoning tasks such as the similarities subtest of the WAIS (Palermo et al., 2017). Furthermore, Brumm et al. (2004) reported significant differences in performance on verbal fluency and naming tasks between AwPKU with current blood-Phe levels >1000μmol/L and those with levels <1000μmol/L, as well as a significant correlation between performance on these measures and current metabolic control in AwPKU. This suggests that, whilst lexical access is unaffected by PKU, tasks requiring interactions with executive function show some level of impairment.

Similarly to findings with AwPKU, lexical abilities have consistently been found to be spared in older adults. In Park et al.'s (2002) study, language was assessed using the vocabulary section of the Shipley scale, as well as synonym and antonym vocabulary tasks. Performance on both tasks demonstrated a significant improvement in knowledge-based verbal ability with increasing age. These findings were supported by Salthouse et al.'s (2004) meta-analysis which reported a positive influence of age on measures of vocabulary across included studies.

A meta-analysis by Moers et al. (2017) looked at the effects of age on reading aloud. The effects of older age were assessed by comparing latency data from 64 older adults (aged 62-95) with that of 41 adolescents (aged 12-18) who were presented with the same materials to read aloud. When the data for older adults was considered in isolation, ageing effects were apparent within this population, with the length of time taken to read words found to increase in line with increasing age. Frequency effects were also observed, with frequent words produced more quickly than infrequent words. No interaction was found, however, between age and frequency. When older adults' performance was compared to that of adolescents, older adults were found to read words more slowly overall than adolescents, and frequent words were produced faster than infrequent words. When predictability of words (i.e., how predictable a word is from the prior text) was removed from analyses as a predictor of reading latency, an interaction between age and frequency did emerge, with older adults found to show less of an impact of frequency on reading latency than adolescents. These findings, therefore, suggest that any increasing impact of lexical difficulty with age is marginal. Consequently, age-related differences in response times in reading aloud tasks appear to be due to generalised slowing, either of perceptual or motor processes, in older adults, rather than lexical processing.

Memory & learning

Studies investigating memory and learning abilities in CwPKU have found some impaired performance with Holtzman et al. (1975) reporting low scores on memory tasks in a cohort of CwPKU. Smith et al. (2000) also found that CwPKU performed more poorly than control participants on a verbal memory task. However, similar to findings with language abilities, once children on- and off-diet were considered separately, these differences only remained for off-diet CwPKU, with on-diet CwPKU performing at a level equivalent to controls.

A number of studies with AwPKU have found no impact of PKU on performance in memory tasks. Channon et al. (2004) compared the performance of 20 early-treated AwPKU with that of 20 matched controls on the Rey Auditory Verbal Learning Test and Complex Figure Test. No significant differences between groups were found in delayed recall or recognition performance on either task. Similarly, Palermo and colleagues' (2017) study found no significant difference between performance of 37 early-treated AwPKU, and 30 healthy controls in immediate or delayed recall on the Rey Auditory Verbal Learning Test, or in a paired associated verbal learning test. Neither were significant differences evident in visuospatial delayed matching to sample or paired associate learning tasks. Comparisons between AwPKU with good and poor metabolic control, have reported differences in performance on verbal memory and learning tasks, however, with Romani et al. (2017) reporting a significant correlation between current Phe levels and verbal memory and learning performance in AwPKU (Brumm et al., 2004; Romani et al., 2017).

In contrast to findings with AwPKU, a clear decline in performance in the domain of LTM and learning has been reliably observed with increasing age. Park et al. (2002) investigated memory abilities in 345 participants aged 20-92 years. Participants completed a series of tasks investigating visuo-spatial and verbal short-term memory (STM) (visuo-spatial tasks: forward and backward Corsi span and spatial span task; verbal tasks: forward and backward digit span), long-term memory (LTM) (visuo-spatial tasks: Rey Visual Design Learning and Benton Visual Retention Test; verbal tasks: free and cued recall) and working memory (visuo-spatial tasks: line span and letter rotation; verbal tasks: reading span and computation span). A continuous age-related decline was found across all memory tasks, with no accelerated decline in later years. In particular, patterns of decline in processing-intensive memory tasks (i.e., those tapping processing speed, LTM and working memory abilities) were found to be notably similar to one another. No significant differences were found between performance in visuo-spatial and verbal domains, although age-related decline was more evident in visuo-spatial than verbal STM. Age-related decline was evident in STM span measures; however, it was more apparent in LTM tasks, with visuo-spatial LTM performance declining at a significantly faster rate than digit span.

Similarly, Gorbach et al. (2017) measure episodic memory in 155 adults (aged 55-80 years) over a 15-year period. Episodic memory was measured using five different free recall tasks, along with processing speed (digit-letter coding, letter-string comparison, and figure comparison), fluid intelligence (block design) and word fluency. Significant decline with age was found on all cognitive measures, with processing speed and fluid intelligence found to

decline steadily throughout the study period. Episodic memory also showed a clear decline with age, and the rate of this decline was found to accelerate from the age of 65 years.

Summary

Table 1.1. – *Summary of cognitive domains impaired and impacted by metabolic control in children and adults with PKU and in healthy older adults. X's marked in bold indicate a high density of available literature supporting impairment for a particular domain. One vs. two x's indicate high vs. low consistency of findings in the literature respectively.*

Cognitive Domain	CwPKU		AwPKU		Older Adults
	Impaired	Impacted by metabolic control	Impaired	Impacted by metabolic control	Impaired
Intelligence	**x**	**xx**	**x**		
Executive function*	**xx**	**x**	**x**	**x**	**x**
Speed of processing	**xx**	**xx**	**xx**	**xx**	**xx**
Visuo-spatial attention			x	x	**x**
Sustained attention	**xx**	**xx**	**xx**	**xx**	
Visuo-motor coordination	x	x	**xx**	xx	**xx**
Language	x	x			
Memory & Learning	x	x		x	**xx**

**excluding inhibition*

Impact of PKU and ageing on neuropathology

PKU has been closely linked, not only to cognitive performance, but also to neuropathological abnormalities. Early animal studies have found that elevated Phe levels during early development can result in abnormal gene expression in oligodendroglia cells, such that these cells are unable to produce myelin (Dyer et al., 1996). These findings have been supported by human neuropathological investigations, which have found evidence of hypomyelination and gliosis in white matter of patients with late-treated PKU (Alvord Jr et al., 1950; Bauman & Kemper, 1982; Dyer et al., 1996). Whilst early treatment of PKU can drastically reduce the neuropathological trauma caused by excess Phe in the developing brain, some cortical abnormalities are still apparent, with studies with this cohort of participants reporting findings of white matter lesions and demyelination in periventricular parietal-occipital and frontal regions (Cleary et al., 1994; Hawks et al., 2019; Jaulent et al., 2020; Lou, 1994; Mastrangelo et al., 2015; Nardecchia et al., 2015; Rubin et al., 2013; Thompson et al., 1993; Villasana et al., 1989; Wesonga et al., 2016; White et al., 2013).

Studies investigating the association between neuropathological abnormalities and metabolic control have found associations between elevated Phe levels and increased white

matter abnormalities in individuals with PKU of all ages (Cleary et al., 1994; Mastrangelo et al., 2015; Nardecchia et al., 2015; Thompson et al., 1993). Thompson et al. (1993) conducted a cross-sectional study with 25 early-treated children and adults with PKU aged 8 to 30 years, all of whom were on a low-Phe diet until at least 7 years old. The study completed MRI scans of participants 1 to 17 years after they had discontinued their low-Phe diet and identified cortical white matter abnormalities in 22 of the 25 participants. No significant associations were found between white matter abnormalities and mean lifetime blood-Phe levels in this study, however, significant associations were found with number of years off diet, and concurrent blood-Phe levels. Regression analyses further found that, when other factors in the model were controlled for, the likelihood of more severe MRI abnormalities increased 1.5 times for every 100μmol/L rise in concurrent blood-Phe levels, and 1.3 times for every additional year off a low-Phe diet.

A similar study by Cleary et al. (1994) investigated neuropathological abnormalities in adults and adolescents with PKU, aged 10 to 50 years using MRI. 71 of the 74 participants scanned demonstrated white matter abnormalities, in particular in the posterior occipital-parietal and frontal lobes. These abnormalities were found to be significantly more severe in older participants who had abandoned their treatment diet than in younger participants who were still following a low-Phe diet. Abnormalities were also found to correlate significantly with both concurrent blood-Phe levels, and Phe levels over the last 5 years, suggesting that age alone does not explain the differences in abnormalities between these two groups. Neurophysiological abnormalities were also investigated, with visual evoked potentials of participants found to differ significantly from controls in terms of latencies, and 8/77 participants with PKU producing latencies above the upper limit of normal.

A more recent longitudinal study was carried out by Nardecchia et al. (2015) who compared MRI abnormalities in 14 participants with PKU scanned twice over a 14-year period. Participants were children during their first scan (age range 7.8-13.5 years) and adults during their second scan (age range 22.2-27.7 years) therefore the authors noted that, not only was age an influencing factor between scans, but also most participants had relaxed their diet during upon reaching adolescence, and their blood-Phe levels had increased accordingly. MRI scans of participants detected no white matter abnormalities during their first scan, however, second scans identified moderate to severe white matter alterations, with worse abnormalities exhibited by participants with worse life-long metabolic control. Similar findings were reported by Mastrangelo et al. (2015) who retrospectively analysed serial MRIs from participants with PKU aged 12 to 37 years. Again, participants were children during their first scan (mean age 9 +/- 4.4 years) and mostly adults during their

last scan (mean age 20 +/- 6 years). Regression analyses of white matter severity scores (WMSS - Leuzzi et al., 1993) found that change in severity of white matter abnormalities was significantly influenced by both dietary control and age, with each increase in blood-Phe of 100µmol/L associated with an 0.46 increase in WMSS, and each increase of 1 year in age associated with a 0.29 increase in WMSS. Changes were not so apparent over a smaller time period, however, with a long-term study by Weglage et al. (2013) failing to find any difference in MRI scans of 57 early-treated AwPKU (aged 19 to 41 years) who were scanned twice within a 5-year period, despite a significant increase in blood-Phe levels occurring during this time.

The relationship between cerebral white matter abnormalities and metabolic control in PKU has been recognised by a number of scoping reviews (Anderson & Leuzzi, 2010; Ferreira et al., 2020; van Spronsen et al., 2011). Anderson and Leuzzi (2010)'s review of the literature provides an effective summary of neuropathological findings in PKU, noting consistent reports of increased signal intensity in periventricular white matter in early-treated participants with PKU, which appear to reliably increase with increasing age and blood-Phe levels. They also note that severe cases have demonstrated extensions of abnormalities to subcortical regions, the brain stem, and the cerebellum. This review also suggests that white matter pathologies may be reversible through strict dietary control of Phe intake, with the authors postulating that a minimum of 2 months may be required for this reversal of neuropathological changes to occur.

Recent systematic reviews of the literature surrounding neuropathological changes in PKU have been carried out by Jaulent et al. (2020) and Burlina et al. (2019). Burlina et al.'s systematic review of neuroimaging results in early-treated AwPKU compared MRI and DTI findings from 35 publications, finding consistent reports of increased MRI signal in periventricular white matter, the brain stem, and the cerebellum, as well as an overall reduced diffusion coefficient. Similarly, Jaulent et al. reviewed neurological complications in 22 cases of AwPKU, 9 of which were early-treated. 20 of the 22 participants reported abnormal MRI results, with extensive demyelination occurring in periventricular regions. 17 participants also underwent a second MRI scan after initiating a low-Phe diet, with 3 showing a complete regression and 5 showing a partial regression of imaging abnormalities. This paper indicates that neuropathological abnormalities may be reversible through increased metabolic control in some circumstances.

Similarly to AwPKU, white matter lesions (WMLs) have consistently been found to be significantly more prevalent in elderly populations than in younger populations, with their

occurrence ranging from 5% to ≥90%, and increasing exponentially with age (Galluzzi et al., 2008; Grueter & Schulz, 2012; Hopkins et al., 2006; Launer et al., 2006; Vernooij et al., 2007; Wen & Sachdev, 2004). These WMLs have been reliably associated with cognitive impairment, as well as with an increased risk of stroke and dementia, and are believed to be indicative of altered interstitial fluid mobility and water content, a possible precursor to demyelination and axonal damage (Wardlaw et al., 2015). Further to this, it has been estimated that, after 60 years of age, total cerebral volume loss increases by approximately half a percent of the total brain volume per year (Enzinger et al., 2005; Hedman et al., 2012; Sigurdsson et al., 2012).

With regards to associations between brain atrophy and cognition with increasing age, associations have been found between total white and grey matter volume and visuo-spatial function, as well as between total white matter volume and speed of processing, with periventricular WMLs found to play a particular role in cognitive impairment with age (Bracco et al., 2005; Jiang et al., 2018) a finding similar to that observed in AwPKU (see Anderson and Leuzzi, 2010; Burlina et al., 2019 for reviews). Grey matter atrophy specifically, meanwhile, has been associated with impaired language, visuo-spatial, and global cognitive function, as well as with the classification of impaired or cognitively normal status of older adults (Jiang et al., 2018). Additionally, grey matter changes in the hippocampus, parahippocampal gyrus, and lateral occipital cortex have been associated with episodic memory decline in participants aged 65+ (Gorbach et al., 2017).

The cerebellum has also been found to play a key role in age-related cognitive decline, in particular with regards to visuo-motor coordination and processing speed impairments (Archibald et al., 2004; Eckert et al., 2010; Genova et al., 2009; Koppelmans et al., 2015). Eckert et al. (2010) conducted MRI scans with 42 healthy adults, aged 19-79, who also completed the Connections Test (a variant of the better-known Trail-making tasks, measuring both perceptual and motor processing speed) as well as a number of tasks from the Woodcock-Johnson Tests of Achievement and Cognitive Abilities. Processing speed was found to decrease with age for all tasks except the Picture Vocabulary subtest of the Woodcock-Johnson battery, with perceptual and motor processing speed decline occurring in participants aged 40 and above. MRI scanning found grey matter atrophy within a specific frontal network to be more predictive of declining performance on simple processing speed tasks with age, than global grey matter volume was. This network specifically included the anterior cingulate cortex and dorsolateral prefrontal cortex. Additionally, a separate cerebellar network was found to play a significant predictive role in simple processing speed performance, with cerebellar volume found to significantly decrease with increasing age.

The role of the cerebellum in cognitive impairments with age has been further supported by Koppelmans et al. (2015) who investigated neuropathological associations between increasing age and performance on a number of motor coordination tasks. 217 healthy older adults aged 64-87 complete the Digital Tapping and Two-hand Coordination subtests from the Vienna Test System, and the grooved pegboard test, as well as completing MRI scans. MRI results found that cerebellar left and right, grey, and white matter volume significantly decreased with age. A significant association with performance on the motor coordination task was only found with left cerebellar grey matter volume, however, whilst grooved pegboard and digital tapping performance was significantly associated with right cerebellar white matter volume. These findings further support the role of both white and grey cerebellar matter in modulating cognitive performance in the domains of visuo-motor coordination and fine motor skills.

Impact of PKU and ageing on neurophysiology

Beyond the structural damage caused by excess Phe in the brain, dysfunction of the dopaminergic and serotonergic neuronal systems due to PKU has also been reported. PKU is known to have a significant impact on dopamine production in the brain due to a lack of metabolism of Phe into tyrosine (a necessary substrate of dopamine). Whilst a low-Phe treatment diet is effective in dramatically reducing blood-Phe levels in PKU, it does very little to address the plasma-tyrosine reduction caused by a lack of metabolisation of Phe, meaning that AwPKU are largely reliant upon tyrosine consumed directly through the diet to produce dopamine (Diamond et al., 1997). The prefrontal cortex is known to be particularly sensitive to tyrosine fluctuations, with even the smallest reduction in tyrosine having a significant effect on dopamine production in this region (Bradberry et al., 1989).

Additionally, early studies found notably decreased blood serotonin levels in people with PKU (Pare et al., 1957), as well as decreased brain serotonin levels in PAH-deficient mice (Yuwiler et al., 1965). Phe and other large neutral amino acids, including tyrosine and tryptophan (a precursor of serotonin), are known to use the same transport proteins to cross the blood-brain barrier, and these transport proteins are known to have a higher affinity for Phe than both tyrosine and tryptophan (Oldendorf, 1973). Reduced levels of dopamine and serotonin, therefore, have been argued to be the result of increased levels of Phe in the blood competing with available tyrosine and tryptophan for passage across the blood brain barrier (Curtius et al., 1972; McKean, 1972; Pardridge, 1998; Pietz et al., 1999). Additionally, it has been argued that high Phe values inhibit tyrosine hydroxylase and tryptophan hydroxylase activities, resulting in decreased availability of dopamine and serotonin as Phe

levels increase (de Groot et al., 2010; Ogawa & Ichinose, 2006; van Spronsen et al., 2009). These reduced levels of dopamine and serotonin have been implicated in memory deficiencies exhibited by PAH-deficient mice (Bruinenberg et al., 2016; Zagreda et al., 1999), as well as in impaired executive function (see Christ et al., 2010; de Groot et al., 2010 for reviews) and decreased well-being (Scala et al., 2020) in AwPKU, whilst observed deficiencies in FDOPA influx and distribution in the striatum of early-treated AwPKU with blood Phe levels of 1260μmol/L (Landvogt et al., 2008) may contribute to observed visuo-motor deficits in PKU.

Impaired cognition in older adults has also been linked to altered neurotransmission in the ageing brain, mirroring neurotransmitter depletions associated with PKU. Ageing has been associated with a loss of striatal and extrastriatal dopamine biomarkers from early through to later adulthood (Kaasinen et al., 2000; Wang et al., 1998), with healthy ageing found to be associated with a notable decline in dopaminergic neuromodulation (Backman et al., 2010; Bannon & Whitty, 1997; see Reeves et al., 2002 for a review). A review of three studies by Kaasinen and Rinne (2002) found age-related decline in dopamine receptor availability to be significantly associated with impaired performance on a range of cognitive functions, in particular those requiring executive functions (such as the Wisconsin Card Sorting Task, and the Stroop test), similar to that observed in AwPKU. Increased striatal dopamine release has also been found to occur during card-sorting tasks (Monchi et al., 2006). Conversely to observations in PKU, age-related striatal dopamine depletion has also been linked to episodic memory impairments (Crosson, 1992), a domain known to be spared in AwPKU. Erixon-Lindroth et al. (2005) found significant links between age-related loss of striatal dopamine transporter density and memory impairments in 12 adults aged 34-81, with loss of dopamine transporter binding in later adulthood found to be associated with poorer performance in tasks tapping LTM (namely word and figure recall) and face recognition tasks. Transporter density was also found to be associated with performance in executive function tasks, such as visual working memory and verbal fluency, with performance on these tasks decreasing as transporter density decreased with age. This suggests that there may be some difference in how dopamine depletion translates to cognitive ability (especially with regards to memory) with increasing age compared to due to PKU.

Age-related memory deficits have not only been associated with striatal dopamine depletion, but also with reductions in hippocampal dopamine and glutamate levels with increasing age. N-methyl-D-aspartate (NMDA) glutamate receptor efficacy has been found to decline with age in both human and animal studies (see Magnusson et al., 2010 for a review), with blockage of these receptors in humans found to impair learning and memory

abilities in healthy adults (Morgan et al., 2004). Prefrontal serotonin levels have also been suggested to play some role in LTM impairments with age, as the serotonin system has been found to play a key role in performance in long-delay recall tasks (de Quervain et al., 2003; Koppel & Goldberg, 2009) and PET previously found a reduced number of prefrontal cortical 5-HT2A serotonin receptors in older adult participants (Sheline et al., 2002). It could, therefore, be the interaction between reduced levels of multiple neurotransmitters across a range of brain areas that results in impaired memory in older adults that is not observed in AwPKU.

Table 1.2. – *Summary of neuropathological and neurophysiological impairments in healthy older adults and adults with PKU. X's marked in bold indicate a high density of available literature supporting impairment for a particular domain. One vs. two x's indicate high vs. low consistency of findings in the literature respectively.*

	AwPKU		**Older adults**
Neuropathology	**Observed**	**Impacted by metabolic control**	**Observed**
White matter atrophy	**xx**	**xx**	**xx**
Grey matter atrophy			**xx**
Neurophysiology			
Frontal/striatal dopamine depletion	**xx**		**xx**
Hippocampal dopamine depletion			xx
Frontal serotonin depletion	**xx**		xx
Hippocampal glutamate depletion			xx

Impact of PKU on well-being

Beyond the effects of PKU on cognition and neurology, untreated PKU has been found to elicit severe emotional and behavioural differences, including psychotic and autistic conditions, aggression, erratic behaviour, irritability, severe temper tantrums, affective lability and hyperkinesis (Bjornson, 1964; Hackney et al., 1968; Jervis, 1954; Koch et al., 1964; Wood Jr et al., 1967; Wright & Tarjan, 1957). Whilst the introduction of the low-Phe diet alleviated many of the characteristics of psychosis and autism associated with PKU (Sullivan & Chang, 1999), early-treated patients with PKU still exhibit a number of emotional and behavioural difficulties, including high levels of sadness, fear, anxiety, phobic reactions, hyperactivity, inattention, feelings of being different, poor self-image, a sense of isolation, withdrawal, and a lack of autonomy and drive in comparison to healthy controls (see Bilder et al., 2016; Smith & Knowles, 2000 for reviews; van Spronsen et al., 2011).

Although less well-established than with cognition, the role of childhood metabolic control in emotional and behavioural development has been documented by a number of cross-sectional studies. One such study conducted by Holtzman et al. (1986) measured well-being in 82 early-treated CwPKU, aged 8 years, using the Louisville Behaviour Checklist, and found a significant correlation between the age at which participants lost dietary control, and higher behaviour problem scores. Similarly, Smith et al. (1988) measured atypical behaviour in 544 early-treated, 8-year-old CwPKU, using teacher's responses on the Rutter Behaviour Questionnaire and found significantly more abnormal behaviours in CwPKU than in matched controls, in particular in those with poor metabolic control. Furthermore, whilst CwPKU with good metabolic control were found to be 1.6 times more likely to show behavioural difficulties than healthy controls, this increased to a level of 2.2 times more likely in CwPKU with poor metabolic control. More recently, emotional problems associated with PKU in childhood have been reported by Landolt et al. (2002), who found a reduction in positive emotions and psychological adjustment reported by caregivers of early-treated CwPKU compared to controls, and by Ford et al. (2018), who surveyed 293 parents and caregivers of early-treated CwPKU (89% of whom were still on their treatment diet) and reported a range of behavioural and emotional difficulties including difficulty maintaining focus, anxiety, depression, social exclusion, and difficulty maintaining relationships.

The relationship between childhood metabolic control and well-being is further supported by Matthews et al. (1986) who found that scores on the Vineland Social Maturity Scale dropped by an average of 10 points in CwPKU who had been off their treatment diet for 2-4 years, with scores dropping in particular in the domains of self-help behaviours, socialisation, and communication skills. Matthews and colleagues further found significant correlations between scores and blood-Phe levels in CwPKU off-diet, whilst no such correlation was observed in CwPKU still on diet. Childhood metabolic control was also demonstrated to impact adulthood well-being by Koch et al. (2002) who found increased difficulties, including phobias, depression, hyperactivity, and lethargy, in AwPKU who had discontinued their low-Phe diet aged 10 compared to on-diet AwPKU, with AwPKU who had stopped their diet before the age of 6.5 showing more predominant symptoms.

Associations between adult well-being and childhood Phe were also reported by Waisbren and Zaff (1994) who investigated responses to the Minnesota Multiphasic Personality Inventory in 28 adult and adolescent women with PKU, aged between 11 and 35 years. They found that participants who had not continuously followed a diet (either starting the diet late or discontinuing the diet for a period of at least 5 years) scored higher on scales of depression, rebellion (defined on scale as "psychopathic deviance"), paranoia, obsessive-

compulsive disorder, schizophrenia, hypomania, and social introversion, than those who had not deviated from their treatment diet. Furthermore, of 9 participants who had sought professional help for emotional problems at some point in their life, 8 of these had lost dietary control at some period in their life.

With regards to well-being in adulthood, Bilder et al. (2013; 2017) and Manti et al. (2016) have reported increased clinical indications of depression, anxiety, and obsessive-compulsion in early-treated AwPKU compared to healthy controls and population norms. More recently, Burton et al. (2013) reported higher levels of psychiatric distress (including obsessive-compulsion, anxiety, depression, phobic anxiety, somatization, interpersonal sensitivity, psychoticism, paranoid ideation, and hostility) in 57 children, 34 adolescents, and 72 adults with PKU with higher median (582µmol/L) and concurrent (683µmol/L) Phe levels, than in those with lower levels (median: 354µmol/L, concurrent: 442µmol/L). Jahja, Huijbregts, et al. (2017) also measured depression, anxiety, somatic problems, avoidant personality traits, ADHD, and antisocial personality traits in 57 early-treated AwPKU using the Adult Self Report measure of the Achenbach System of Empirically Based Assessment, finding significantly more depressive and avoidant personality traits in AwPKU than controls, in particular in AwPKU with a mean lifetime blood-Phe level of >=360µmol/L, although no significant correlation with concurrent blood-Phe levels was found. This group of participants also reported more depressive and somatic problems than either AwPKU with low childhood Phe (<360µmol/L) or healthy controls.

Findings of decreased well-being in adulthood, and their association with metabolic control, however, have been contested (e.g. Aitkenhead et al., 2021; Brumm et al., 2004; Palermo et al., 2020; Pietz et al., 1997). Palermo et al. (2020) measured well-being in terms of quality of life relating to emotional, mental, and physical health in 26 early-treated AwPKU, 6 of whom were on an unrestricted diet and 20 of whom were following a low-protein diet. Well-being was measured using the Beck Depression Inventory-II (BDI-II), the Mental and Physical Health Quality of Life Scales from the SF-36, the Interpersonal Reactivity Index (IRI; measuring empathy), and the Comprehensive Affect Test System (CATS; measuring emotion recognition). Scores from AwPKU did not differ from matched controls on any of the assessed well-being measures, other than speed of response on the CATS test of emotional recognition. Only two correlations were found between well-being scores and metabolic control. These correlations were in the opposite to expected direction, with better perspective taking significantly associated with higher childhood and adolescent blood-Phe levels and reduced personal distress scores associated with higher adolescent and

concurrent blood-Phe levels. A positive correlation was found, however, between better metabolic control and more Vitality. The negative correlations between metabolic control and well-being in the study may reflect a direct negative impact of Phe levels on well-being, however they may also be reflective of the severe restriction imposed by following a low-Phe diet, and the social and emotional difficulties associated with such restrictions. Similarly, Aitkenhead et al. (2021) found no differences in well-being between AwPKU and control participants in terms of anxiety and depression as measured by the HADS, although they did find that AwPKU demonstrated a more anxious relationship style than controls. With regards to the relationship between metabolic control and well-being, Aitkenhead et al. found significantly lower quality of life, both physical and mental, was demonstrated by AwPKU who only partially adhered to a treatment diet, as compared to those who were either completely on, or completely off, diet. This further supports the assertion that poor well-being in AwPKU may be more reflective of the lifestyle elements surrounding trying to maintain a low-Phe diet, than of the specific impact of blood-Phe levels.

Summary

Overall, evidence regarding the impact of dietary control on well-being in AwPKU is less clear than that with CwPKU, with studies finding differing levels of emotional and behavioural difficulties in participants who have abandoned a low-Phe diet post-childhood. The effects of losing metabolic control on cognition in childhood are well-established, with impairments found in the domains of sustained attention, speed of processing, IQ, visuo-motor coordination, and executive function (working memory, planning, task-switching, and inhibition) in CwPKU. Some of these difficulties persist into adulthood, with high adulthood Phe levels associated with deficits in sustained attention, speed of processing (in particular in tasks requiring visuo-spatial attention), and some executive functions (namely working memory, planning, and task-switching). Inhibition and IQ, however, seem to be fairly robust to altering Phe levels once people with PKU reach adulthood, whilst language and memory and learning appear to be largely spared, with decreased accuracy on measures in these domains only appearing in tasks requiring interactions with executive function. This pattern of cognitive impairment in AwPKU is notably similar to that observed in healthy ageing, with both populations demonstrating reduced speed of processing, deficits in flexibility and sustained attention, and impaired visuo-motor coordination, but exhibiting milder deficits in reasoning and inhibitory control, and spared lexical access.

Cognitive abilities have proven to be sensitive to changing blood-Phe levels, and as such can be an effective indicator of the effects of metabolic control, and interactions with

ageing, on brain health in people with PKU. Neuropathological studies have found interactions between impaired cognition and brain abnormalities. PKU has been consistently associated with white matter atrophy, in particular in periventricular regions, which has been found to increase with age and blood-Phe levels. These abnormalities have also been associated with slower processing speeds due to a combination of hypomyelination and demyelination caused by the toxic effects of Phe in the brain. Additionally, neurophysiological investigations have reported that increased Phe levels are linked to depletions of dopamine and serotonin in the frontal lobes, leading to impaired cognition (including memory and executive function deficits) and decreased well-being. Increasing age has been similarly associated with white matter atrophy as well as frontal and striatal dopamine depletion and frontal serotonin depletion. However, ageing has further been linked to depletion of dopamine and glutamate in the hippocampus, as well as to cerebellar and hippocampal atrophy, areas not typically impacted by PKU.

Early-treated PKU has been associated with a range of emotional and well-being difficulties, with high childhood Phe levels linked to behavioural difficulties (including hyperactivity, inattention, and a difficulty maintaining relationships), and to emotional difficulties (such as anxiety and depression). Links between increased Phe and emotional problems are also apparent in adulthood, with higher depression, anxiety, and psychiatric distress reported in adults with higher Phe levels, as well as increased personality problems. However, findings with well-being are more heterogenous than those with neurophysiology, neuropathology, and cognition, with numerous reports of increased well-being after diet termination, and partially-adherent AwPKU suggested to be the most at risk of well-being difficulties.

More knowledge is required to inform guidelines about what levels of Phe are safe for people with PKU to maintain, especially into adolescence and adulthood, when many people with PKU choose to relax or discontinue their low-Phe diet. Importantly, use of the same assessments to measure impairment across the lifespan (i.e. in childhood, adolescence, early adulthood, and middle-age) must be ensured, to allow clinicians to fully understand the differential effects of PKU on outcomes at critical ages, without results being impacted by differential assessment sensitivities. As the first cohort of early-treated AwPKU move into middle- and older-age, an increased knowledge of the effects of prolonged, high Phe levels on the brain, along with the possible interactions that this may have with ageing, is required, as well as a consideration of the effect that the socially restrictive nature of following a low-Phe diet can have on well-being, and the trade-off between these two effects of maintaining metabolic control throughout the lifespan.

Thesis aims and objectives

This thesis has two main aims. The first is to increase our understanding of the impact of early-treated PKU on cognition and well-being in middle-aged AwPKU, whilst also considering the possible interactions between increasing age and PKU-related impairments as these adults move into older-age. Whilst considerable research has taken place investigating outcomes in children, adolescents, and young adults with early-treated PKU, the relatively recent introduction of treatment from birth has limited the research that can take place with older adults with early-treated PKU. We will combine data gathered directly from middle-aged AwPKU with information extrapolated from comparisons between young AwPKU and healthy older controls to investigate the potential effects that PKU may have on cognitive health as early-treated AwPKU continue to age. The second aim of this thesis is to increase our understanding of the direct impact of metabolic control on cognitive, neurological, and well-being outcomes, in within-participant studies where the impact of external influences (such as socioeconomic background and Phe tolerance) is mitigated. This includes investigating the potential reversibility of damage caused by high Phe levels through increased metabolic control in adulthood.

In Chapter 2, the methodology for delivering a comprehensive cognitive assessment battery will be detailed. The assessment battery used was intended to mirror assessments given to young AwPKU and controls in previously published studies. Due to the impact of the Covid-19 pandemic, however, cognitive assessments had to be modified in some instances for remote delivery. This chapter will outline the parameters of, and differences between, tasks delivered face-to-face and online to control and clinical participants.

In Chapter 3, existing comparisons between cognition in young AwPKU and matched controls are extended to include a comparison of performance with healthy older controls, aged 53-88. Impairment profiles of older controls and young AwPKU are contrasted to identify similarities and differences between deficits caused by ageing and those caused by PKU. These comparisons, along with considerations of impairment severity, allow for a better understanding of how cognitive impairments due to increasing age may interact with those caused by PKU as early-treated AwPKU reach later life.

In Chapter 4, the results of Chapter 3 are built upon by carrying out in-depth investigations of the impact of task difficulty on performance across the domains of visuo-spatial processing and language. The relative impact of task difficulty across these domains is compared to identify whether a generalised speed of processing impairment is apparent,

or whether impairments are domain-specific, in one or both cohorts. Comparing the outcomes of these investigations for both cohorts further allows us to understand whether the mechanisms underlying slowed responses due to ageing and PKU are the same, and therefore likely to interact as AwPKU age, or whether they are different, and therefore unlikely to impact upon one another.

In Chapter 5, an online version of the cognitive assessment battery is delivered to middle-aged AwPKU and healthy controls to investigate how differences between these two cohorts may mirror, or differ from, those between younger adults with and without PKU. These comparisons provide an opportunity to assess the cognitive health of the first cohort of early-treated AwPKU as they move from young adulthood to middle-age. It also aims to increase our understanding of how, and when, age-related cognitive deficits may interact with those caused by PKU, as well as how these interactions may be mitigated by metabolic control in later life. In addition to cognitive investigations, the impact of PKU and metabolic control on well-being and quality of life in this cohort is assessed.

In Chapter 6, a systematic review is conducted to assess the impacts of metabolic control on cognition, well-being, and neurology in children and adults with PKU. Within-participant studies are reviewed to control for the significant individual differences that are apparent between individuals with PKU. This review further examines the reversibility of negative outcomes in individuals with PKU with high Phe levels if metabolic control is tightened, and concurrent Phe levels decreased. These findings provide valuable insight into the potential for older adults with PKU to reverse impairments caused by relaxing or terminating their low-Phe diet during adulthood (and to prevent further damage being caused through interactions between PKU and ageing) by resuming a low-Phe diet in later life. This chapter also emphasises the importance of maintaining tight metabolic control during childhood.

In summary, this thesis aims to increase our understanding of how early-treated PKU affects cognition, well-being, and neurology in adults with PKU, how these outcomes are likely to interact with increasing age and metabolic control, and whether impairments caused by high Phe levels are reversible through decreasing Phe levels in later life.

Chapter 2: Cognitive Testing Methodology

Participants for Chapters 3, 4 and 5 of this thesis were assessed using a comprehensive cognitive assessment battery. As this battery was extensive, detailed descriptions of all included tasks are laid out here, allowing more concise methodological descriptions to be included in individual chapters. Participants in Chapters 3 and 4 (older controls, young adult controls, and young AwPKU) were assessed by previous members of the research group, using a face-to-face version of this assessment battery. Analyses comparing performance in young adult controls and young AwPKU have previously been published in Palermo et al. (2017) and Romani et al. (2018). These comparisons were then extended upon to include older controls in Chapters 3 and 4. Participants in Chapter 5 (middle-aged controls and middle-aged AwPKU) were assessed for the purposes of this PhD, and comparisons were then made between data collected from these participants, and that collected from participants in previous chapters. Initial study designs for this chapter intended to use the same cognitive assessment battery as was used previously by Palermo et al. (2017) and Romani et al. (2018), to allow direct comparisons of data collected from participant groups before and during this PhD. Due to the outbreak of the global pandemic during the execution of this research, however, some tasks within the assessment battery required re-designing to allow testing to continue during periods of national lockdown and social distancing. Eight tasks, therefore, were moved online whilst 13 tasks were conducted remotely via video call (two of which also included an online component to deliver task materials). Unfortunately, seven assessments (plus one subtest of the WAIS) were not possible to complete remotely, therefore they were necessarily removed from the testing batteries for these studies. Adaptations of tasks for remote delivery are also laid out within this chapter. Some tasks were also redesigned to allow for redelivery of matched assessments to participants on 3 occasions during a longitudinal study however, due to the pandemic, the study requiring redelivery of assessments was placed on indefinite hold.

Summary of Procedures

Face-to-face testing was conducted over two separate testing sessions, each lasting 2.5 to 3 hours, including a break in the middle of each session. Remote assessment sessions were completed over 2 or 3 sessions, the first of which lasted 45 minutes. The second session was then completed in one 1.5 to 2hr session, or across two 1hr sessions. All participants were asked to complete a uniform battery of neuropsychological tests investigating cognitive abilities in the following areas: executive function, memory and learning, visual attention, sustained attention, language, and motor control. All face-to-face

tasks were table-top or computerised. Computerised tasks were created using E-Prime and Python Experiment Builder. Cognitive assessment was completed over two sessions, approximately 1 week apart. Data included in comparisons between young adult controls, young adults with PKU, and older adult controls (Chapters 3 and 4) was collected through face-to-face assessment.

Online materials were created using the PsychoPy3 Experiment Builder and hosted online by Pavlovia. Remotely delivered tasks were designed to match face-to-face assessments as closely as possible, but some differences were unavoidable due to software and programming constraints. Cognitive assessment was completed over two or three sessions, approximately one week apart. The first session took approximately 45mins to 1hour to complete and consisted entirely of online-delivered tasks hosted on Pavlovia. The remaining assessments were then conducted as 'pen and paper tasks' via video call, where researchers delivered and scored assessments during the video call. These tasks were delivered either in a single 1.5 to 2hour session, or over two 45min to 1hr sessions, dependent on participant availability/preference. Data from remote assessments was included in comparisons between middle-aged adults with and without PKU (Chapter 5). In all studies, clinical and control groups within the same age group were administered the same tasks.

For the purposes of the methods description below. Original designs of cognitive tasks (delivered in a face-to-face environment) are described. Any adaptations for remote delivery are then detailed.

Cognitive Tasks Across Domains

General intelligence

***Weschler Abbreviated Scale of Intelligence (WASI; Weschler, 1999)*.**

The vocabulary, block design, and similarities subtests from the WASI were administered to all participants being assessed face-to-face.

Remote delivery – Included only the vocabulary and similarities subtests. The block design subtest was not possible to complete in remote assessments as it requires participants to interact with physical blocks and arrange them in different patterns.

Visuo-spatial attention and processing speed

Choice reaction time*.* An arrow was presented on the screen pointing either to the left or to the right. Side of presentation and direction of pointing were always congruent. Stimuli were presented for a maximum of 3100ms. Trials where no response was given were recorded as errors. Participants had a response box in front of them and were asked to press the button on the left of the press pad when they saw the arrow pointing to the left, and the button on the right of the press pad when they saw the arrow pointing to the right. Inter-stimulus intervals ranged from 1 to 750ms, with intervals randomly selected.

Remote delivery – This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. Instead of a response box, participants were asked to respond to arrows by pressing the corresponding arrow key on their keyboard. Inter-stimulus intervals ranged from 750 to 1000ms, with intervals randomly selected. Inter-stimulus intervals differed from face-to-face conditions due to set limitations within PsychoPy Experiment Builder.

Simple reaction time. The screen was divided into quadrants, with an empty box in each quadrant. Participants were asked to press the space bar as quickly as possible whenever a ladybird appeared in one of the boxes. The ladybird appeared after a variable delay of 500 to 3000 msec.

Remote delivery – This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. All elements were identical to the original version of the task.

Detection with distractors *(go-no-go).* The screen was divided into quadrants, with an empty box in each quadrant. In different trials, a fixation cross was presented in the middle of the screen, followed by a ladybird and/or green bugs presented in 1 or 2 of the boxes. Participants were asked to press the space bar as quickly as possible when the ladybird was presented on the screen and not to respond when the green bug appeared. In the “go” condition, the ladybird was presented on its own, with a second ladybird, or with a green bug. In the “no-go” condition, the green bug was presented either on its own, or with a second green bug. The ladybird was present in 32 of 64 trials. The presentation order of trials was randomised. Of the trials without the ladybird, 16 had one distractor, and 16 had two distractors. After all 64 trials had been completed, the task was run again with the required responses reversed (i.e., participants were asked to respond when the green bug was present and withhold responses when only the ladybird was present. **This is discussed in further detail in the ‘Executive function’ tasks section below**). Trials were presented for 3000ms or until a response was made. Inter-stimulus intervals were 1500ms. During the

first 500ms only a fixation cross was present in the centre of the screen. This was followed by 1000ms of the fixation cross as well as 4 empty boxes in target locations.

Remote delivery – This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. During inter-stimulus intervals, the fixation cross was present in the centre of the screen as well as with 4 empty boxes at target locations (rather than the boxes appearing later) due to the researcher's programming limitations with PsychoPy Experiment Builder.

Visual search*.* In visual search tasks, participants were asked to search for a red ladybird (the target) amongst several other bugs on the screen. Participants were asked to press the 'm' key on their keyboard when the target was present, and the 'z' key on their keyboard when the target was absent (this was switched in cases where participants were left-handed). In a feature search condition, distractors were all green bugs differing from the target both in colour and shape. In a conjunction search condition, distractors also included red bugs, which differed from the target only in shape. In conjunction trials, equal numbers of green and red distractors were presented. Each search-type (feature and conjunction) consisted of 36 trials. Reaction time to find the target was measured, along with accuracy. A fixation cross appeared at the beginning of each trial and was displayed for 1sec before the stimulus array was presented. Stimulus arrays consisted of 4, 8, or 12 distractor items in target-absent trials, with one distractor replaced by the target in target-present trials. In target-absent conjunction search trials, equal numbers of green and red distractors were presented. In target-present conjunction search trials, one red bug distractor was replaced by the target. The array was displayed for 10secs or until a response was made.

Remote delivery - This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. Due to the researcher's programming limitations with PsychoPy Experiment Builder, stimulus arrays consisted of 4, 8, or 12 distractor items in all trials, with an additional ladybird presented in target-present trials. In conjunction trials, equal numbers of green and red distractors were presented in all conditions. The array was displayed until a response was made.

Sustained Attention

Rapid Visual Information Processing *(RVP).* This task assessed the ability to maintain attention over time. Digits ranging from 2 to 9 appeared on the screen, one at a time, at a rate of 100 digits per minute. Participants were asked to detect sequences of 3 digits (i.e., 2-4-6, 3-5-7, 6-4-8) by pressing the response key when the last number of the sequence appeared on the screen. There were 3 different target sequences to watch for. An error was

scored for each incorrect answer (e.g., pressing the response key at the wrong time) or omission (i.e., not pressing the response key for the last number of a sequence) within a window including the following 2 digits (1800ms). There were 76 'trials' in total, split into four 'blocks' of 19 trials in E-Prime. Each 'trial' contained a variable number of stimuli (from 3 to 10). For each 'block', 9 trials contained a target sequence and 10 did not. 36 target sequences were included in total.

Remote delivery - This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. Due to the different set-up of the experiment in PsychoPy and E-Prime, the number of total trials and target trials differed slightly between the face-to-face and online conditions but were kept as close as possible. As PsychoPy does not allow for 'blocks' of trials in the way that E-Prime does, stimuli were instead made up of 134 individual sequences of 3 digits between 1 and 9. 29 target sequences were included in total. The order of the 29 target trials within the presentation of 134 trials was randomised initially. All participants were then presented with the same order of trials.

Visuo-motor coordination

Grooved Pegboard Task*.* The pegboard task involves both visual-motor coordination and fine motor control. Participants were required to put 25 pegs into 25 holes in a board as quickly as possible. Each of the pegs have fins so they must be rotated to match slots at different angles. Participants were asked to fill the board with pegs as quickly as possible using only one hand. Dominant and non-dominant hands were used for two trials each. Number of seconds to complete the tasks and number of dropped pegs were recorded.

Remote delivery - This task was not included in the remote assessments, as it requires participants to physically manipulate pegs and place them in the board.

Digit symbol-coding*.* Participants were given a key that matches numbers to symbols. They were asked to write symbols in boxes arranged below numbers on a test page, filling each box with the corresponding symbol in serial order. They were given 90 seconds to fill as many boxes as possible. Each incorrectly matched number–symbol and each empty box was counted as an error (max. number of errors = 93).

Remote delivery - This task was not included in the remote assessments, as it requires participants to have a physical key and response sheet in front of them on which they can write their responses.

Trail Making Task A *(Reitan & Wolfson, 1985)*. This task required participants to draw a trail with a pencil to join a number of circles. In version A of the task, the circles only contained numbers. Participants were instructed to connect the circles in ascending order as quickly as possible. The time taken to complete the task was recorded.

Remote delivery - This task was not included in the remote assessments, as it requires participants to have a physical response sheet in front of them on which they can draw with a pen/pencil.

Short-term and working memory tasks

Digit span*.* Participants were asked to repeat back sequences of digits read to them by the experimenter at a rate of approximately one per second. Ten lists were presented starting at a length of 4 digits. If an error was made within the presentation of 4 digits, sequences of 3 digits were presented. If the participant recalled more than half of the sequences correctly, a longer sequence was presented (up to a maximum length of 8 items). Otherwise, the test was stopped. To calculate span, each correctly repeated sequence was assigned a value of 0.1. Each length, therefore, was scored one point if all sequences of that length were correct. Three points were added for any lengths not administered (i.e., 1, 2, and 3 digits) as these were assumed correct.

Remote delivery - This task was delivered via video call. All elements were identical to the original version of the task.

Non-word repetition. Participants were asked to repeat several sequences of non-words. There were 10 sequences each of 2, 3, and 4 non-words. All non-words respected the phonotactic constraints of the English language. The task was discontinued if more than half of the sequences of a given length were repeated incorrectly. Each sequence incorrectly repeated was scored as 1 error (max. number of errors = 30). Performance was also scored by individual item, resulting in scores out of 20, 30, and 40, for lists of 2, 3, and 4 non-words respectively.

Remote delivery - This task was delivered via video call. All elements were identical to the original version of the task.

Corsi Block Span*.* Assessed participants' ability to remember spatial sequences.

Design A - The examiner tapped a sequence of blocks at a rate of one per second. Immediately afterwards, the participant attempted to reproduce the sequence in the same order. Sequences of increasing length (from 2 to 9) were presented with three trials for each length. The task was stopped when the participant failed to reproduce any sequences of a

given length. Points were awarded for the number of individual sequences repeated correctly. The number of correct sequences were then totalled and divided by three to calculate Corsi span.

Remote delivery - This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. Participants were presented with 9 white squares on their computer screen. The array of squares was designed to bear as close a resemblance as possible to the distribution of blocks on the Corsi board used in the face-to-face version of this task. A sequence of blocks would flash red for 1 second each. Immediately afterwards, the participant attempted to reproduce the sequence in the same order by clicking on the appropriate boxes on the screen.

Language tasks

Picture naming *(lexical access)*. Participants were asked to do a picture naming task similar to that originally used by Howard et al. (2006), measuring both speed of lexical access and effect of semantic interference **(see Executive Function tasks below)**. A total of 165 pictures were displayed on a computer screen, one at a time. 120 images were split into 24 semantic categories, whilst the remaining 45 were filler images. Participants were asked to name each picture as quickly as possible. Errors and self-corrections were noted, and RTs were recorded via a Cedrus Voice-key. The overall reaction time was taken as a measure of speed of lexical access. Each picture remained on screen until a verbal response was made.

Remote delivery – This task was delivered via video call. It was redesigned to be shorter, therefore containing less images. This was originally to allow for redelivery of this task with matched images at a later date as described above. A total of 120 pictures, split in 24 semantic categories, were displayed on a computer screen, one at a time. Participants were presented with a fixation cross followed by a stimulus image in the centre of the screen. An auditory beep was played at the onset of each stimulus and recorded through the computer speakers along with the participant's responses. Each picture remained on screen until the space bar was pressed. Sound files were then uploaded to Audacity to compute RTs between the beep (i.e., stimulus onset) and participant's responses (Figs.2.1. and 2.2.).

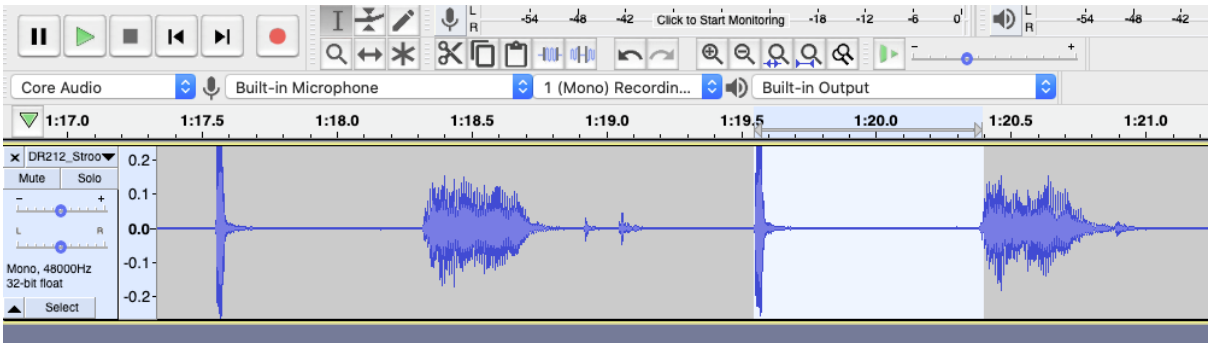


Figure 2.1. – *an example of a sound file demonstrating the space between the stimulus onset beep and the participant's verbal response to the stimulus.*

Figure 2.2. – *an example of the output given when the space between the stimulus onset beep and the participant's verbal response to the stimulus is highlighted. The final figure represents the timeframe of a given selection, and therefore provides the participant's response time.*

Word and non-word spelling. Participants were given a list of words to spell to dictation. Homophones were presented with a disambiguating sentence. All non-words were derived from real words by substituting one or two phonemes. 140 regular and irregular words of various frequencies and lengths were presented as well as 40 non-words. For each word, the participant was asked to repeat the word and then write it down on a blank sheet of paper.

Remote delivery - This task was delivered via video call. It was redesigned to be shorter, therefore containing fewer words. This was originally to allow for redelivery of this task with matched words at a later date as described above. 30 regular and irregular words were presented as well as 40 non-words. For each word, the participant was asked to repeat the word and then type it into an online Qualtrics form which was submitted at the end of the task. Participants were asked to ensure that spell checker was turned off. For participants who were unable to disable their spell checker, they were asked to write the word on a piece of paper and then hold the paper up to the camera for the researcher to transcribe their response.

Word and non-word reading. Participants were asked to read aloud, and as fast an accurately as possible, 140 individually presented words. Words were varied along measures of regularity, frequency, and length. An additional 40 non-words were created by changing one or two letters in corresponding high-frequency words. Words appeared one at

a time at the centre of a computer screen following a fixation cross. RTs were recorded using a Cedrus Voice-key. The same procedure was then used with 40 non-words, constructed by changing one or two letters in corresponding high frequency words. Non-words were split into the same length categories as real words. Stimuli were presented one at a time in the centre of a computer screen, following a fixation cross lasting 1000ms. Stimuli disappeared 500ms after a response was recorded. Stimuli remained on screen indefinitely until a response was made. RTs were recorded via a Cedrus Voice-key.

Remote delivery - This task was not included in remote assessments. It was not possible to employ the Cedrus Voice-Key technology via video call. Whilst a similar method to that used in the picture naming and Stroop tasks could have been employed (playing an audible beep then manually extracting reaction times from soundwave data), the additional processing time that this would add in terms of task redesign and analysis led to researchers agreeing to remove this task from the remote assessment battery.

Phonological awareness*.* In a phoneme deletion task, participants heard a spoken English word (e.g., table) and one sound (/t/). Participants were asked to repeat back the word, but without the sound (e.g., able). There were 40 trials in total. Taking away the sound resulted in a real word in half of the trials, and in a nonsense word for the other half (e.g., powder; /d/ >power; cabbage; /k/ > abbage). In a spoonerisms task, participants heard two words. They were asked to exchange the first sounds of the pair. For example, for the pair "bad sin," the correct response is "sad bin." Both sounds in each pair had to be exchanged correctly for participants to score a point. There were 24 trials in total. Switching the first sounds of the given words resulted in 2 new, real words for half of the trials, and in 2 nonsense words for the other half. Participants were informed in advance whether the correct switching of sounds would produce real, or nonsense, words.

Remote delivery - This task was delivered via video call. All elements were identical to the original version of the task.

Sentence completion*.* **This task was not included in original assessments. It was added to remote assessments as an additional measure of language and inhibition.** Two sets of 15 sentences were read aloud to participants, each with a missing final word. In the logical condition of the task, the participant was asked to complete the sentence with a logical final word as quickly as possible. The researcher timed responses using a stopwatch. In the illogical condition, sentences were read aloud again, but this time participants were asked to complete a sentence with a word NOT connected to the rest of the sentence in any way, providing a measure of response time and inhibition **(see Executive Function section below)**. Two types of errors were recorded; a) direct completion (where the obvious missing

word was given) or b) indirect completion (where a different word was chosen, but the chosen word still related to the content of the sentence such that the sentence still made sense). Sentences in logical and illogical completion conditions were matched on length and Cloze completion probability scores (Block & Baldwin, 2010).

Stroop Task *(colour naming).* Participants were asked to name the colour of stimuli presented on the screen. Participants were presented with a fixation cross, followed by a stimulus from one of 3 different congruency conditions: Congruent (e.g., 'GREEN' written in green ink), Incongruent (e.g., 'GREEN' written in blue ink), and Neutral (i.e., "XXXX" written in coloured ink). Stimuli appeared in the centre of the computer screen and participants were required to say, out loud, the colour that the word/letters were printed in as quickly as they could. Each condition consisted of 24 trials. RTs for congruent and neutral conditions, measured using a Cedrus Voice-key, were reported as a language speed measure. Responses to incongruent trials were used as a measure of inhibition **(see Executive Function section below).** Stimuli remained on the screen indefinitely until a verbal response was made.

Remote delivery - This task was delivered via video call. An auditory beep was played at the onset of each stimulus and recorded through the computer speakers, along with the participant's responses. Stimuli remained onscreen until the space bar was pressed. Sound files were then uploaded to Audacity to computer RTs between the beep (i.e., stimulus onset) and participant's responses (Figs.2.1. and 2.2.).

Long-term memory and learning tasks

Delayed Matching to Sample*.* Participants were shown a complex visual pattern (the target) on a computer screen and asked to select the target when it was presented along with three similar patterns. The patterns to choose from were either shown at the same time as the target (10 trials), after a delay of 4 seconds (10 trials), or after a longer delay of 12 seconds (10 trials).

Remote delivery – This task was not included in remote assessments. The original task was designed in E-Prime. The complexity and specificity of the task meant that it could not be recreated in PsychoPy/Pavlovia by the researcher within the given timeframe for this project.

The Rey Auditory Verbal Learning Test. Participants were orally presented with a list of 15 nouns (List A) and asked to recall as many words as they could, in any order, immediately after presentation. List A was repeated five times to assess learning. A second list (List B)

was then presented for recall, to assess interference. Recall of list A was assessed again immediately after recall of list B. After a 20-minute delay, participants were asked to recall as many items as possible from list A. In a delayed recognition task, participants were presented with a grid of 50 words as asked to circle any of the words that were included in lists A and/or B, testing both long-term verbal recall and recognition.

Remote delivery – This task was delivered via video call. Participants were only presented with words from list A in recall and recognition tasks. This was to allow for repeated assessments (as discussed previously). Researchers had access to matched versions of list A but only one version of list B, therefore list B was removed from the repeated assessment battery. Removal of this additional measure also ensured that the remote assessment battery did not become excessively time-consuming to complete. The recognition portion of this task was delivered using an online form hosted by Qualtrics. All recognition words were presented in a list, and participants were asked to select all words that they recognised from list A.

Visual Paired Associate Learning *(PAL).* Participants were shown a number of white boxes on a computer screen which were 'opened', one at a time in a random order, to reveal visual patterns. Each 'open' box was presented for 1500ms. Participants were instructed to remember which pattern belonged to which box. After the last box is opened, the patterns were displayed in the middle of the screen, one at a time, and participants were asked to point to the box where the pattern was originally located. If the participant made an error, the patterns were re-presented (maximum 10 attempts at each game size). There are 2 trials with 3 patterns, 1 with 6 patterns, and 1 with 8 patterns (N = 80 patterns; max. number of errors = 200).

Remote delivery – This task was not included in remote assessments. The original task was designed in E-Prime. The complexity and specificity of the task meant that it could not be recreated in PsychoPy/Pavlovia by the researcher within the given timeframe for this project.

Verbal Paired Associate Learning*.* In an initial phase, participants were presented with nine pictures, each associated with a written novel word, and asked to write down the words. They were then asked to write the correct word on presentation of the picture alone (testing phase). In case of errors, the correct word was presented, and the participants asked to copy it down. The task was discontinued when all the words in the list were recalled correctly or after a maximum of five attempts at the whole list. Where testing finished after a completely correct list, but before all five attempts were necessary, all subsequent attempts were assumed to be correct (maximum score = 45). After 20 minutes (delayed recall)

participants were presented with the nine pictures again and asked to write the correct words (maximum score = 9). Participants were presented with a piece of paper and pen to write down their responses for the researcher to see.

Remote delivery – The task was delivered via video call. Images and novel words were presented to participants by the researcher via the video chat camera. Participants wrote down their responses, and then held these up to the camera for researchers to see.

Executive function tasks (Inhibition, planning, and task-switching, and reasoning).

WCST – 64 Card Version. Participants were presented with four place-holding cards depicting symbols differing in colour (green, red, blue, and yellow), number (1-4), and shape (circles, triangles, crosses, and stars). Participants were then given a deck of standardised WCST-64 cards, organised in numerical order (as per the printed numbers on the backs of the cards). Each of these cards depicted symbols differing in colour (green, red, blue, and yellow), number (1-4), and shape (circles, triangles, crosses, and stars). Participants were asked to match each card with the corresponding place-holding card by placing each card below the place-holding card representing the category that they had chosen. The order of the cards in the deck was the same for each assessment. Participants were not told what dimension (colour, number, or shape) to use to match the cards, but feedback was provided after each choice. Once the participant made 10 consecutive correct matches in one category (e.g., colour), the sorting rule was changed (e.g., shape) without telling the participant. The participant then had to discover the new rule, assessing both their ability to derive rules, and use of feedback to shift a cognitive set. Participants were presented with 64 cards in total. Performance was scored in terms of number of correct responses, number of errors (cards not sorted according to the current rule), categories completed, and number of perseverative responses. Perseverative responses refer to participants continuing with the same response strategy after a rule has changed and were assessed using the standardised WCST scoring manual.

Remote delivery - This task was re-created in PsychoPy and delivered to participants via a Pavlovia link. A publicly available ‘WCST demo’ programme hosted on Pavlovia was used as a coding template for this task. In the online version of this task, participants were presented with images of four place-holding ‘cards’ on their screen. Participants were then presented with images of different ‘cards’ and asked to match each one with the corresponding place-holding card by clicking on the relevant place-holding image. The order that the ‘card’ images were presented to participants was randomised.

Trail Making Tasks B-A. This task requires participants to draw a trail with a pencil to join a number of circles. In version A of the task, the circles only contain numbers (**see visuo-motor tasks above**). In version B, circles contain either numbers (1 – 13) or letters (A – L) and participants are instructed to alternate between the numbers and the letters (i.e., 1-A-2-B, etc.). Time taken to complete the task was recorded. Completion time for task A was then subtracted from completion time for task B, providing a difference score representative of the increased planning and task-switching abilities required for task B.

Remote delivery - This task was not included in the remote assessments, as it requires participants to have a physical response sheet in front of them on which they can draw with a pen/pencil.

Tower of Hanoi puzzle. Participants were asked to move a tower of rings (i.e., 3, 4 or 5 rings) stacked on one peg in a pyramid of decreasing sizes. They were required to follow rules that allowed only certain moves to move the stack to a different peg (i.e., a larger ring can never be placed on a smaller ring and only one ring can be moved at a time). Problems of increasing complexity were presented, defined by the minimum number of moves necessary to solve the problem (from a 7-move to a 31-move problem). A trial was ended if, after 6 minutes, the participant was not able to reach the goal configuration. Performance was scored in terms of total number of unresolved trials, rather than in terms of number of moves (max. errors = 9).

Remote delivery – This task was not included in remote assessments. The original task used a physical, wooden version of the task. The complexity and specificity of the task meant that it could not be coded into an online format by the researcher within the given timeframe for this project.

Verbal Fluency. In a letter fluency condition, participants were required to produce as many words as possible beginning with a given letter in a minute of time. Participants were instructed not to use proper nouns or words sharing the same root. In a semantic fluency condition (Rosen, 1980), participants were required to generate as many words as possible belonging to a specific semantic category in a minute of time. Scores were based on the number of acceptable words produced. Participants were presented with 3 different letter conditions (‘c’, ‘f’, and ‘l’), and one semantic category (animals).

Remote delivery – This task was delivered via video call. Assessment stimuli differed from the original task, to allow for matched repeated assessments (as previously discussed). Participants were presented with 2 different letter conditions (‘c’ and ‘p’), and 2 semantic categories (animals and clothing).

Detection with distractors *(go-no-go).* In addition to its role as a visuo-spatial attention assessment (**see visuo-spatial attention and processing speed tasks above**), the detection with distractors task was employed as a measure of inhibition. In the "go" condition, the ladybird was presented on its own, with a second ladybird, or with a green bug. In the "no-go" condition, the green bug was presented either on its own, or with a second green bug. After all 64 trials had been completed, the task was run again with the required responses reversed (i.e., participants were asked to respond when the green bug was present and withhold responses when only the ladybird was present). Differences in responses times and accuracy between the original task, and the reversed task, were then taken as a measure of inhibition, as previously learned responses had to be suppressed.

Sentence completion. **This task was not included in original assessments. It was added to remote assessments as an additional measure of language and inhibition.** In addition to its role as a language assessment (**see language tasks above**), the sentence completion task was employed as a measure of inhibition. In the logical condition of the task, the participant was asked to complete the sentence with a logical final word as quickly as possible. In the illogical condition, sentences were read aloud again, but this time participants were asked to complete the sentence with a word NOT connected to the rest of the sentence in any way, therefore requiring participants to suppress a logical response. Differences in response time between the logical and illogical conditions were calculated as a measure of inhibition.

Picture naming *(semantic interference)*. In addition to its role as a language assessment (**see language tasks above**), the picture naming task was employed as a measure of inhibition. 120 picture naming stimuli were presented, belonging to 24 different semantic categories (5 items in each category). Previous results have shown that RTs increase with the ordinal position of a stimulus in a set of semantically related pictures (Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2007). Reaction time increases and accuracy differences between the first two presented stimuli and the last two presented stimuli in each semantic category, therefore, were taken as measures of inhibitory control.

Stroop Task. In addition to its role as a language assessment (**see language tasks above**), the Stroop task was employed as a measure of inhibition. The difference in accuracy and RTs between congruent and incongruent trials of the Stroop task was taken as a measure of inhibitory control. To succeed in the incongruent condition, participants were required to suppress their instinctive tendency to read the word rather than describe the ink colour.

Assessment delivery

Face-to-face assessment sessions were delivered in the Psychology labs on campus at Aston University. Participants were asked to switch their phones to silent, and only the participant and the researcher were present during the assessment. Participants were reminded that they could take a break at any time. Assessments were conducted in two separate testing sessions, each lasting 2.5 to 3 hours, conducted approximately one week apart. Tasks were delivered in the same order to all participants, and all efforts were made to administer all tasks to all participants.

Online assessment sessions took place remotely via video call, using either Zoom or Microsoft Teams. All participants were asked to find a quiet place to complete the assessments, where they would not be disturbed, and to switch their phones to silent for the duration of the session. All participants with PKU completed assessments in their homes. Most control participants completed assessments in their home, but two participated at their place of work. Whilst every effort was made to ensure that participants were completing assessments in an appropriately controlled environment, remote assessment sessions were occasionally interrupted (e.g., by landlines ringing, pets/people entering the room etc.,). In these instances, the researcher would pause the session if possible and wait until the participant was able to return their full attention to the assessment. Assessments were conducted in two separate testing sessions. Session one lasted approximately 1 hour. Session two was either conducted over a single 2-hour session or split further into two 1-hour sessions, dependent on the participant's schedule. Sessions were conducted approximately one week apart. Where possible, tasks were administered in a fixed order to all participants, however some flexibility in task order was required dependent on participants' schedules and engagement levels (e.g., if time restraints meant that the next task in the assessment order would take too long to complete in the current session, this may be swapped out for a shorter task, and the longer task then completed in the second session). All efforts were made to administer all tasks to all participants.

Data Analysis

It is recognised that, whilst every effort was made to keep face-to-face and remote assessments as similar as possible, direct comparison of raw data gathered through differing testing modalities would be subject to a number of confounding influences (such as differing levels of environmental control, differences in test procedures etc,). To address the impact of these factors, z-scores were calculated for each group ([individual score] – [average of

control group]/[control group SD]). This provided a standardised measure of impairment relative to controls (tested using the same modality) for each task which could then be compared between participant groups even if they were tested using different modalities. Z-scores were then directly compared between groups using t-test comparisons, and comparisons of Cohen's D measures of effect size.

Summary

Taken together, these tasks form a comprehensive cognitive assessment battery, assessing abilities in the domains of general intelligence, executive function, memory and learning, visual attention, sustained attention, language, and visuo-motor coordination. Due to the global pandemic, we were not able to deliver identical tasks to all participants included in this thesis. Remotely delivered assessments, however, were designed to match those delivered face-to-face as closely as possible, allowing comparisons across participant groups assessed at different times to be carried out. Unfortunately, the limitations of remotely-delivered assessments meant that fewer tasks were included in the remote version of the assessment battery than in the original, face-to-face version. Therefore, data for certain domains (in particular visuomotor control and some measures of executive function) could not be obtained for participants assessed using this testing modality.

In the next two chapters, data gathered using the original, face-to-face assessment battery is used to conduct secondary analyses, adding results from older controls to previously published comparisons between younger adults with and without PKU. In Chapter 5, data gathered using the remote assessment battery is then added to analyses. Importantly, control vs. PKU participant cohorts in all age groups were assessed using identical materials, therefore standardised scores indicating levels of impairment in AwPKU compared to matched controls should remain relatively unaffected by the changes in assessment modality.

Chapter 3: Investigating the Impacts of PKU and Healthy Ageing on Cognition

Introduction

Since the 1950s, PKU has been treated through pharmacological intervention and/or a low-Phe diet, started immediately after diagnosis in the neonatal period at approximately 5 days old (Moyle, Fox, Arthur, et al., 2007). This treatment approach has proven to be an effective means of reducing blood-Phe levels in people with PKU (Scriver & Kaufman, 2001), however, some cognitive impairment is still evident in individuals following the diet. Furthermore, whilst the life expectancy of individuals with untreated PKU has previously been estimated to be around 56 years (Jancar, 1998), the introduction of the low-Phe treatment diet has resulted in people with PKU now living past this age and maintaining comparatively good neurological health. The relatively recent introduction of the treatment diet, however, means that little is known about the impact that old age will have upon the cognitive function of early-treated individuals with PKU.

Patterns of impairment associated with healthy ageing and PKU

As discussed in the introductory chapter of this thesis (Chapter 1), AwPKU have been shown to demonstrate impaired cognitive function compared to healthy populations in the domains of general intelligence, executive function, speed of processing, visuo-spatial attention, sustained attention, and visuo-motor coordination. Abilities in the domains of language, and long-term memory and learning, however, have been reported to be spared in this population. The profile of cognitive impariment in PKU also bears a striking similarity to that observed in healthy older adults as a natural product of ageing. Significant impairments have been found to be associated with both age and PKU across a range of cognitive functions, including visuo-spatial attention (for PKU see Brumm et al., 2004; Nardecchia et al., 2015; Romani et al., 2018; for ageing see Kramer et al., 1996), processing speed (Bucur & Madden, 2010; Der & Deary, 2006), visuo-motor coordination (for PKU see Moyle et al., 2007; Weglage et al., 1995; for ageing see Hamilton et al., 2018; Volkow et al., 1998) and executive function (for PKU see Brum et al., 2004; DeRoche & Welsh, 2008; Nardecchia et al., 2015; for ageing see De Beni & Palladino, 2004; Delaloye et al., 2009). There are, however, some key differences between the profiles of cognitive impairment in these two populations, with older adults frequently demonstrating impaired memory, whilst IQ in this population (a domain known to be impacted by PKU, e.g., Burgard, 2000; Burlina et al.,

2019; Waisbren et al., 2007) appears to remain unaffected (Craik, 1994; Salthouse 2004; Park et al., 2002; see Chapter 1, pgs.14-27 for a detailed review of the literature).

Cognitive impairment in PKU has been linked to neuropathological damage caused by PKU, including white matter damage such as hypomyelination/demyelination and decreased diffusivity in periventricular regions of the brain (Cleary et al., 1994; Hawks et al., 2019; Jaulent et al., 2020; Lou, 1994; Mastrangelo et al., 2015; Nardecchia et al., 2015; Rubin et al., 2013; Thompson et al., 1993; Villasana et al., 1989; Wesonga et al., 2016; White et al., 2013). The role of myelin in speed of neural transmission has led researchers to hypothesise that the toxic impact of excess Phe on myelin integrity may be a cause for the slow speed of processing demonstrated by AwPKU. White matter lesions have further been consistently linked with ageing, with significant associations found between white matter atrophy and cognitive impairments in ageing populations, and associations reported between decreasing white matter volume and slowed speed of processing (Bracco et al., 2005; Galluzzi et al., 2008; Grueter & Schulz, 2012; Hopkins et al., 2006; Jiang et al., 2018; Launer et al., 2006; Vernooij et al., 2007; Wen & Sachdev, 2004). In contrast to the neuropathological profile typically observed in AwPKU, however, cognitive impairments in older populations have further been linked to grey matter atrophy, in particular in the hippocampus and cerebellum, with atrophy in these areas associated with deficits in the domains of memory, visuo-motor coordination, and speed of processing (Eckert et al., 2010; Gorbach et al., 2017; Koppelmans et al., 2015).

Beyond the structural damage caused by excess Phe in the brain, PKU has been found to result in neurophysiological dysfunction of the dopaminergic and serotonergic neuronal systems (de Groot et al., 2010; Ogawa & Ichinose, 2006; Pare et 1957; van Spronsen et al., 2009; Yuwiler et al., 1965). Reduced levels of dopamine and serotonin in PKU, as well as deficiencies in striatal FDOPA influx and distribution, have been implicated in impaired memory, executive function, and visuo-motor function in PKU (Bruinenberg et al., 2016; Christ et al., 2010; de Groot et al., 2010; Landvogt et al., 2008; Zagreda et al., 1999). Similar effects of age on serotonergic and dopaminergic transmission have also been reported, with neurotransmitter depletion due to ageing found to occur in striatal and prefrontal regions (Kaasinan et al., 2000; Sheline et al., 2002; Wang et al., 1998), but with additional reduced hippocampal serotonin, dopamine, and glutamate transmission, not apparent in PKU, reported in older populations (see Magnusson et al., 2010 for a review). Age-related decline in dopamine receptor availability has been reported to have a similar negative impact on executive function as that observed in AwPKU (see Kaasinen & Rinne, 2002 for a review). In contrast to AwPKU, however, neurotransmitter depletions due to age

have also been significantly associated with impaired learning and memory abilities (Crosson, 1992; Erixon-Lindroth et al., 2005; Morgan et al., 2004), suggesting that there may be some difference in how neurotrasmitter depletion translates to cognitive ability with increasing age compared to due to PKU.

Summary

In summary, there are a number of overlaps between both the neurological and the cognitive impacts of ageing and PKU (Table 3.1.). In particular, the domains of processing speed, visuo-spatial attention, visuo-motor coordination, and executive function (planning, updating, task-switching), show evidence of impairment in both these populations, as do the physiological substrates thought to underly them. Despite this clear overlap, however, a number of differences are also present between these two populations, with PKU studies finding no evidence of impaired memory and learning, an area consistently found to be impaired in ageing populations. Differences are also apparent in the opposite direction. General intelligence, in particular areas of crystallised knowledge such as vocabulary knowledge, have often been found to be unaffected, or even improved by normal ageing, whilst evidence of some impairment compared to matched controls is apparent in multiple studies with PKU participants.

Table 3.1. – *Summary of cognitive, neuropathological, and neurophysiological impairments in healthy older adults and adults with PKU. X's marked in bold indicate a high density of available literature supporting impairment for a particular domain. One vs. two x's indicates high vs. low consistency of findings in the literature respectively.*

Cognitive Domain	**AwPKU**	**Older adults**
Intelligence	**x**	
Executive function	**x**	**x**
Speed of processing	**xx**	**xx**
Visuo-spatial attention	x	**x**
Sustained attention	**xx**	
Visuo-motor coordination	**xx**	**xx**
Language		
Memory & Learning		**xx**
Neuropathology		
White matter atrophy	**xx**	**xx**
Grey matter atrophy		**xx**
Neurophysiology		
Frontal/striatal dopamine depletion	**xx**	**xx**
Hippocampal dopamine depletion		xx
Frontal serotonin depletion	**xx**	xx
Hippocampal glutamate depletion		xx

Objectives and hypotheses

Whilst a large amount of research has been conducted investigating the impact of both age and PKU on cognition individually, a direct comparison has yet to be made. Furthermore, the heterogeneity of tasks used to assess cognition across studies prohibits an effective comparison of performance of participants in different study cohorts. A comparison such as this is imperative to allow clinicians and individuals with PKU alike to prepare for the possible impact that older age will have upon cognitive function in PKU as the first generation of early-treated AwPKU reach middle- and old-age.

Data from AwPKU and young adult controls has previously been compared in this manner by members of our research group, and these comparisons can be found in a number of publications (published in Palermo et al., 2017, Romani et al., 2018, 2019). For this study, previously gathered data from young adult controls and AwPKU was compared with data gathered from older control participants, using the same battery of cognitive tests, to assess the impact of age and PKU on performance across cognitive domains. Performance of both young AwPKU and older adult controls were compared to that of younger adult controls on the same cognitive assessments, to identify any influences of age or PKU on performance in these populations. This research provides a unique opportunity to directly compare the impact of ageing and PKU on cognition across domains, and to ascertain the areas of overlapping impairment that are likely to be significantly impacted by ageing in older adults with PKU.

Hypotheses:

1. Cognitive impairment associated with white matter abnormalities (namely speed of processing) and with prefrontal or striatal cortical neurotransmitter depletion (namely executive function, and visuo-motor coordination) will be visible in both AwPKU and older control populations.

2. Cognitive impairment associated with hippocampal neurotransmitter depletion will be observed in older controls only. We expect older controls to perform poorly on memory and learning tasks, but that this domain will be spared in AwPKU.

3. We expect to see some evidence of impaired general intelligence in AwPKU, relative to younger controls, whilst general intelligence in older controls will be spared.

4. We expect to see impaired performance from AwPKU in measures of sustained attention, but that this domain will be spared in older controls.

5. We expect to find no impairment in tasks measuring inhibition and vocabulary knowledge in either population, due to crystallised knowledge remaining intact, and the use of compensatory strategies.

Methods

PKU Participants.

Thirty-seven early-treated young adults, aged 18 to 41 years (mean=27.5, SD=7.3), with classical PKU were recruited from the Department of Inherited Metabolic Disorders at the University Hospitals Birmingham. All participants were diagnosed through new-born screening conducted 5–7 days after birth and were continuously treated with a low-Phe diet from diagnosis. Data on historical Phe levels was gathered via the PKU database at The Clinical Chemistry Department at Birmingham Children's Hospital. These participants were recruited and assessed prior to the start of this PhD. At time of testing, seven participants were on an unrestricted diet and 30 were following a low-Phe diet. All early-treated PKU individuals attending the clinic were invited to participate, as well as a number of individuals who were not currently attending clinic follow-up appointments but were still contactable. All individuals who responded to the invitation were tested.

Participants for this group were recruited and assessed prior to the undertaking of this PhD. Data from this cohort has previously been published in Palermo et al. (2017) and Romani et al. (2017, 2019).

Control Participants

Controls consisted of a group of 30 young, healthy adults aged 18-41 (mean=27.6, SD=7.4), matched with PKU participants for age, gender, and educational status, and a group of older, healthy adults aged 53 to 88 years (mean=69.2, SD=7) (Table 3.2.). Older and younger controls were matched for years of education and gender. Healthy volunteers were recruited through the Aston University volunteering website.

Participants for this group were recruited and assessed prior to the undertaking of this PhD. Data from the cohort of young controls has previously been published in Palermo

et al. (2017) and Romani et al. (2017, 2019). Data from the cohort of older controls has not previously been analysed or published.

Table 3.2. – *Demographic information for each participant group*

		Sex		Age (yrs.)			Education (yrs.)	
Participant Type	**N**	**% Male**	**% Female**	**Mean**	**SD**	**Range**	**Mean**	**SD**
AwPKU	37	35	65	27.5	7.3	18-41	14.4	2.0
Younger control	30	33	67	27.6	7.4	18-41	15.2	1.7
Older control	56	37.5	62.5	69.2	7.0	53-88	16.2	3.5

Assessments

Cognitive assessments were conducted in two separate testing sessions, each lasting 2.5 to 3 hours. This 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 in the same order; however, some data points are missing due to data collection errors, and participant attrition between the first and second sessions.

All participants were asked to complete the same battery of neuropsychological tests previously delivered by Palermo et al. (2017) and Romani et al. (2017, 2019), investigating cognitive abilities in the following areas: IQ, executive function, memory and learning, visual attention, sustained attention, language, and visuo-motor coordination. All tasks were table-top or computerised, and adjustable to suit the age of each participant. Cognitive assessment was completed over 2 sessions, approximately 1 week apart. For a detailed descriptions of task design, delivery, and scoring see the methodology chapter of this thesis (Chapter 2).

IQ

WASI – The vocabulary, block design, and similarities subtests from the WASI were administered. Verbal IQ, Procedural IQ, and Full-Scale IQ were scored.

Executive function

WCST (64-card version) – Participants were asked to correctly match 64 cards with symbols on them with corresponding place-holding cards. Sorting rules (colour, form, or

number) were not disclosed to participants, and were changed following 10 correct responses. Number of correct responses was scored.

Trail making tasks B-A – Participants joined a number of circles in order using a pencil. In task A, circles contained only numbers. In task B, circles contained numbers or letters and participants were asked to alternate between numbers and letters in ascending order. Time taken to complete task A was then subtracted from time taken to complete task B.

Tower of Hanoi – Participants were asked to move a stack of rings from one peg to another following a set of rules (i.e., a larger ring can never be placed on a smaller ring and only one ring can be moved at a time). Increasing numbers of rings were included in the stack (3, 4, or 5 rings), increasing the minimum number of moves required to solve the problem for each trial. Number of unsuccessful trials was scored.

Verbal fluency – Participants were asked to name as many words as possible, either starting with a particular letter or belonging to a specific category, within 1 minute. The task included three letter fluency conditions (with letters ‘c’, ‘f’, and ‘l’) and one semantic fluency condition (with category ‘animals’). Number of correct responses was scored.

Picture naming – Inhibitory control was measured through comparing response times to images presented later in a semantic category to those presented earlier in the same category to assess the effect of semantic interference on speed of response in the picture naming task (see language tasks).

Stroop task – Inhibitory control was measured through measuring differences in response times to congruent and incongruent stimuli in the Stroop task (see language tasks).

Memory and learning

Digit span – Participants were asked to repeat back sequences of digits read to them. Lists started at a length of four digits. If the participant recalled more than half of the 10 sequences presented for a given length, they were presented with a longer sequence (up to a maximum of eight digits). Number of correct responses was scored.

Corsi span – Participants were presented with nine square, wooden blocks. The examiner tapped the blocks in a particular sequence, immediately after which participants were asked to tap the blocks in the same order. Three trials were presented for each sequence length (from 2 to 9). If the participant responded correctly to at least one trial of a given length, they were presented with the next length, otherwise the task was stopped. Number of correct responses was scored.

Non-word repetition – Participants were asked to repeat sequences of non-words. Ten trials were presented for each sequence-length (2, 3, and 4 non-words). The task was stopped if more than half of the sequences for a given length were incorrect. Number of correct responses was scored.

Delayed matching to sample – Participants were shown a complex visual pattern (the target) and asked to select the target when it was presented along with three similar patterns either at the same time as the target, after a delay of 4 seconds, or after a delay of 12 seconds. Number of correct responses was scored.

Rey Auditory Verbal Learning Task – Participants were read a list (List A) of 15 nouns and asked to recall as many words as they could, in any order. The list was repeated five times and participants were asked to repeat what they could remember after each presentation. A second list (List B) was then presented for recall, to assess interference. Recall of list A was assessed again immediately after recall of list B. After a 20-minute delay, participants were asked again to recall as many items from List A as they could remember.

Visual paired associated learning – Participants were presented with white boxes on a computer screen which ‘opened’ to reveal visual patterns. After all patterns were revealed, patterns were presented again in isolation ad participants were asked to point to the box where the pattern was originally located. Trials were made up of 3 patterns, 6 patterns, then 8 patterns. A maximum of 10 attempts was allowed for each trial. Number of correct responses was scored.

Verbal paired associate learning – Participants were presented with nine pictures, each associated with a written non-word. They were then presented with the images again and asked to write down the associated non-word. The task ended when all nine words were recalled correctly, or after five presentations of the whole list. Number of correct responses was scored.

Visual attention

Choice reaction time – Participants were presented with an arrow pointing to either the left or the right of the computer screen. They were asked to respond by pressing the corresponding arrow on their keyboard as quickly as possible. Reaction time and number of correct responses was scored.

Detection with distractors – Participants were asked to respond as quickly as possible when a target (ladybird) was presented on a screen, with or without a distractor (green bug), and not to respond if only distractors (green bugs) were presented. Upon completion of 64 trials, the task was run again but required responses were reversed (i.e., participants were asked to respond for the green bug and not respond for the ladybird). Reaction time and number of correct responses was scored.

Visual search – Participants were asked to search for a target (the red ladybird) among a number of distractors (4, 8, or 12 other bugs) on their screen. They were asked to press the ‘m’ key on their keyboard if the target was present, or the ‘z’ key if it was absent (responses were reversed for left-handed participants). In a feature search condition, distractors were green bugs only. In a conjunction search condition, distractors were both green and red bugs. Reaction time and number of correct responses was scored.

Sustained attention

Rapid Visual Information Processing (RVP) – Participants were asked to detect target sequences of three digits (3-5-7, 2-4-6, or 4-6-8) in a rapidly presented string of digits between ‘1’ and ‘9’. Reaction time and number of correct responses was scored.

Language

Picture naming – Participants were asked to name 120 images, split into 24 semantic categories, presented on the screen, one at a time. They were asked to respond as quickly as possible. Reaction time and number of correct responses was scored.

Word and non-word spelling – Participants were given a list of words to spell to dictation. 140 regular and irregular words of various frequencies and lengths were presented, as well as 40 non-words. Number of correct responses was scored.

Word and non-word reading – Participants were asked to read aloud, as quickly as possible, individually presented words on a screen. 140 regular and irregular words of various frequencies and lengths were presented, as well as 40 non-words. RT and number of correct responses was scored.

Phoneme deletion – Participants were read a word (e.g., table), and then one sound from within that word (e.g., /t/). They were then asked to repeat back the word without the given sound. The task included 20 trials which resulted in real words (e.g., powder, /d/ = power) and 20 which resulted in non-words (e.g., cabbage, /k/ = abbage). Number of correct responses was scored.

Spoonerisms – Participants were read two words. They were asked to exchange the first two sounds and repeat the words back (e.g., bad, sin = sad, bin). The task included 12 trials which resulted in two real words, and 12 trials which resulted in two nonsense words. Number of correct responses was scored.

Stroop task – Participants were asked to name the colour of stimuli presented on a screen. In a neutral, condition they were presented with a string of X's. In a congruent condition, they were presented with a word in the same colour text as the word (e.g., the word 'Green' in green text). In an incongruent condition, they were presented with a word in a different colour text than the word (e.g., the word 'Blue' in red text).

Visuo-motor coordination

Grooved Pegboard – Participants were asked to put 25 pegs with fins in 25 slots in a board as quickly as possible using one hand, rotating the pegs to match slots at different orientations. Each participant completed two trials with their dominant hand, and two trials with their non-dominant hand. Number of drops, and time to complete the task was scored.

Digit-symbol coding – Participants were asked to write symbols in boxes below numbers using a key. They were given 90 seconds to fill as many boxes as possible. Number of correct responses was scored.

Trail Making Test A – Participants were asked to join numbered circles in ascending order into a trail. Time taken to complete the task was scored.

Data handling and analysis

Whilst the data for this study had been gathered prior to commencement of this PhD, all data required checking and cleaning by the thesis author before analyses could be conducted. Outcomes on all measures were scored and input into one database. Z-scores were calculated for individual scores by taking the young-adult control group as a baseline for both the PKU participants and the older controls ([individual score] – [average of control group]/[control group SD]), allowing direct comparisons between PKU participants (AwPKU), younger controls (YC) and older controls (OC) to be made across different tasks in different cognitive domains. Due to this study's interest in group, rather than individual, z-scores, as well as the particular focus on differences in z-scores between groups, a relatively low threshold for identifying mild impairments was applied. As such, for this study, $z>-0.5$ = no or minimal impairments, $z=-0.5$ to -2 = mild to moderate cognitive impairments, and $z\leq-2$ = severe cognitive impairment. Similar z-score impairment categories have previously been applied in studies investigating mild cognitive impairments in clinical cohorts (Bohnen et al., 2015; Rowe et al., 2013). Data was cleaned across participants to remove any individuals with ≥20% non-valid responses from all reaction time (RT) and error analyses (see Table 3.A1 in Appendix F). Individual participant RTs were also cleaned to exclude error responses and non-valid responses (responses <100ms or ±3SD from participant mean; see Tables 3.A2 and 3.A3 in Appendix F).

Composite scores were calculated for accuracy and RT performance of AwPKU and OC in each cognitive domain using the task categories outlined in the method section, apart from executive function which was split into two categories: inhibition and executive functions (i.e., planning, reasoning, and task-switching). These two categories for executive function were created based on the current literature, which has consistently found inhibition to be spared in AwPKU and older adults, whilst the remaining executive functions show evidence of impairment in these cohorts.

Between-subjects t-tests were then carried out to investigate the differences in individual task performance between OC and YC to assess the pattern of impairment in OC across tasks. These were then compared to previous t-test comparisons between AwPKU and YC (published in Palermo et al., 2017) to identify similarities and differences in patterns of impairment in these two cohorts compared to healthy younger adults. Cohen's D effect sizes have also been computed, to contextualise the strength of results. It must be noted that the large number of individual assessments puts these analyses at risk of a Type I error, where seemingly significant differences may emerge due to chance. We have reported

significance with a standard p-value, without applying corrections. Significant differences in individual tests may warrant follow up in further research. However, it is important to stress that individual significant results are of limited value. What is important is whether patterns of impairment can be identified, where impairments in one task are corroborated by impairments in similar tasks. For this reason, composite z-scores have been computed so that scores across heterogenous tasks in different cognitive domains can be directly compared using a single value representing 'level of impairment'. Finally, comparisons between z-scores of AwPKU and OC were carried out, to assess the relative severity of impairment in these two cohorts.

Results

Composite scores

To address hypotheses 1 to 4 of this chapter, comparisons of composite scores across cognitive domains were initially carried out. These can be seen in Table 3.3. Older controls (OC) were moderately to severely impaired in measures of verbal and visuo-spatial RTs, LTM and learning, and visuo-motor coordination. Conversely, they performed significantly better than younger controls (YC), in measures of general intelligence, and at the same level in accuracy measures across domains of language, visuo-spatial attention, and inhibition. Contrary to hypothesis 4, however, OC also showed moderate impairment in the domain of sustained attention.

AwPKU also demonstrated expected mildly to moderately impaired performance in the domains of verbal and visuo-spatial RTs, sustained attention, and visuo-motor coordination, along with unimpaired performance in the domains of verbal and visuo-spatial accuracy, and inhibition. STM was found to be impaired to some extent in both cohorts, but differences in performance from YC did not reach significance.

AwPKU, however, differed from OC with regards to full-scale IQ (in which they demonstrated a mild impairment whilst OC performed at an equal or higher level than YC), executive function (in which both cohorts were mildly to moderately impaired, but comparisons between OC and YC did not reach significance) and LTM and learning (in which AwPKU performed well and OC were moderately impaired). Comparisons of z-scores in impaired domains between OC and AwPKU suggest that, although both cohorts demonstrated moderate to severe impairments in the domains of verbal and visuo-spatial

RTs and visuo-motor coordination, these impairments were significantly more pronounced in OC than in AwPKU, suggesting that ageing more severely impacts these domains than PKU.

To further investigate the nature and severity of impairments in affected domains (as demonstrated by composite score comparisons), and to investigate the impact of age and PKU on the specific abilities set out in hypothesis 5 (namely crystallised intelligence, inhibition, and vocabulary knowledge), performance on individual tasks was compared between participant groups (Table 3.4.).

Table 3.3. - *Comparison of composite z-scores across cognitive domains. Z-scores have been reversed in some instances so that lower z-scores indicate worse performance across tasks. Z-scores of 0.6 or above in either direction have been highlighted in bold. Unless indicated otherwise, p-values indicate significance of t-tests between each population and Younger Controls. Differences in z-scores are calculated by subtracting AwPKU from OC z-scores.*

Cognitive domain	Older Control (OC)		Adult with PKU (AwPKU)		OC vs. AwPKU	
	Mean z-score	p-value	Mean z-score	p-value	Difference in z-scores	p-value
Full Scale IQ	0.9	<.001**	-0.9	<.001**	-1.8	<.001**
Language accuracy	0.0	.46	-0.2	.69	-0.2	.20
Language RT	-2.1	<.001**	-0.9	<.001**	1.2	<.001**
Visuo-spatial accuracy	0.0	.65	0.0	.79	0.0	.83
Visuo-spatial RT	-3.0	<.001**	-0.8	<.001**	2.2	<.001**
Sustained attention	-1.4	<.001**	-0.6	.04*	0.8	.02*
STM	-0.9	.09	-0.6	.50	0.3	.31
Inhibition	-0.1	.36	0.2	.46	0.3	.17
Executive Function	-0.9	.75	-1.4	.02*	-0.5	.04*
LTM & learning	-1.3	<.001**	-0.1	.36	1.2	<.001**
Visuo-motor coordination	-2.5	<.001**	-1.1	<.001**	1.4	<.001**

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Composite scores:
Visuo-spatial attention and processing speed = choice RT, simple RT, detection with distractors, visual search
Sustained attention = RVP
Visuo-motor coordination = grooved pegboard, digit symbol-coding, Trail Making Task A
STM and Working memory = digit span, non-word repetition, Corsi block span
Language = picture naming, word and non-word spelling, word and non-word reading, phoneme deletion, spoonerisms, Stroop neutral and congruent conditions, letter fluency
LTM and learning = delayed matching to sample, visual and verbal paired associate learning and delayed recall, Rey delayed recall
Inhibition = Stroop incongruent-congruent conditions, Trail Making Tasks B-A, detection with distractors bug-ladybird
Executive function = WCST, Tower of Hanoi, semantic fluency.

Older controls vs. younger controls

Results of individual task performance supported hypotheses 1, with OC demonstrating expected moderately to severely impaired performance on all executive function tasks and visuo-motor tasks, as well as consistently slower RTs across domains. Hypothesis 2 was also supported, with OCs performance found to be moderately impaired on all LTM and learning tasks. With regards to hypotheses 3 and 5, comparisons of individual tasks demonstrated superior vocabulary knowledge in OC compared to YC (as demonstrated by VIQ, reading word accuracy, and picture naming accuracy scores) as well as higher full-scale IQ and unimpaired inhibition (as demonstrated by Stroop incongruent accuracy), supporting expected findings in these domains. Hypothesis 4, however, was not supported, with OC demonstrating significantly poorer accuracy scores on the RVP measure of sustained attention than YC. Scores on individual tasks can be seen in Table 3.4.

AwPKU vs. younger controls

Hypothesis 1 (that AwPKU and OC would both demonstrate visuo-motor, speed of processing, and executive function performance) was supported, with AwPKU demonstrating mild to moderately impaired z-scores and significantly slower RTs than YC across domains. AwPKU also demonstrated some impairment across visuo-motor coordination tasks, although not to the same level as that exhibited by OC participants. Impairments were also observed in measures of executive function, in particular in tasks such as the Tower of Hanoi, which require a significant amount of goal-oriented planning.

Hypothesis 2 was also supported, with a similar distribution of scores found between AwPKU and YC in the domain of LTM and learning, suggesting that this area of cognitive function is unaffected by PKU in early adulthood. Hypothesis 3 was similarly supported by AwPKU's full-scale IQ score which, although remaining within the range of unimpaired scores, was significantly lower than those exhibited by the matched controls.

Hypothesis 4 was supported in terms of AwPKU performance, with this cohort demonstrating expected poorer accuracy on the RVP task than YC. Finally, hypothesis 5 was only partially supported, with inhibition scores found to be unimpaired in AwPKU. Vocabulary knowledge, however, was found to be lower than that of YC with regards to the vocabulary and similarities subtests of the WAIS, although word reading and spelling accuracy was at the same level of controls. Scores on individual tasks can be seen in Table 3.4.

Older controls vs. AwPKU

A number of similarities were observed between the cognitive profiles of OC and AwPKU, with a significant correlation found between z-scores for the two cohorts (Pearson r=.39; p=.002). Differences between OC and AwPKU, however, were present in both directions, thus excluding the possibility that differences simply reflected different severity levels. As hypothesised, both groups demonstrated moderate to severe impairments in the domains of speed of processing, executive function, and visuo-motor coordination. OC were significantly more impaired than AwPKU in the domains of speed of processing and visuo-motor coordination, however, performing significantly worse than AwPKU across RT measures (excluding reading tasks), and visuo-motor tasks (excluding trail-making A). Both cohorts also demonstrated equal or superior performance to YC in tasks tapping vocabulary knowledge and inhibition abilities, although this was more evident in OC than AwPKU.

AwPKU, however, performed worse than OC in terms of IQ, verbal accuracy, and executive function, whilst OC performed worse than AwPKU in terms of verbal and visuo-spatial IQ, LTM and learning, visuo-motor coordination. Contrary to hypotheses, both cohorts demonstrated impaired performance in the RVP measure of sustained attention, with OC also performing significantly less accurately than AwPKU on this task. Scores on individual tasks can be seen in Table 3.4.

Table 3.4. - *Comparison of scores on all tasks across cognitive domains. Z scores reflect differences from younger controls. They have been reversed for some tasks so that lower z-scores always indicate worse performance. Z-scores of 0.5 or more have been highlighted in green. Z-scores of -0.5 or less have been highlighted in red. For the comparison between AwPKU and OC, positive z scores (in yellow) indicate better performance of OC; negative z-scores (in grey) indicate better performance of the AwPKU. Effect sizes of 0.8 or above have been highlighted in bold text. YC = Younger Controls, OC = Older Controls, AwPKU = Adults with PKU.*

Cognitive task	YC		OC					AwPKU					AwPKU vs. OC	
	Mean	SD	Mean	SD	z-score	p-value	Cohen's d	Mean	SD	z-score	p-value	Cohen's d	diff in z-scores	p-value
IQ General Knowledge														
Procedural IQ	112.0	11.7	121.5	14.9	**0.8**	**<.001****	**1.0**	104.5	15.1	**-0.6**	**.02***	**0.8**	**1.4**	**<.001****
Verbal IQ	112.2	10.2	119.2	10.2	**0.7**	**<.001****	**1.0**	102.3	12.9	**-1.0**	**<.001****	**1.2**	**1.7**	**<.001****
Vocabulary subtest	64.2	7.1	68.5	6.2	**0.6**	**.008****	**0.9**	58.4	8.7	**-0.8**	**.004****	**1.0**	**1.4**	**<.001****
Similarities subtest	39.2	3.6	39.0	5.3	**0.0**	.86	0.1	36.0	5.8	**-0.9**	**.008****	**0.9**	**0.9**	**.013**
Full-Scale IQ	113.8	10.9	123.3	12.5	**0.9**	**<.001****	**1.1**	103.9	14.3	**-0.9**	**<.001****	**1.1**	**1.8**	**<.001****
Spelling accuracy														
Word spelling % errors	4.8	5.7	5.7	19.7	**-0.2**	.75	0.1	4.0	4.6	**0.2**	.52	0.2	**-0.4**	0.52
Non-word spelling % errors	11.9	7.6	22.4	20.1	**-1.4**	**<.001****	**1.0**	13.5	8.2	**-0.2**	.43	0.3	**-1.2**	**<.001****
Reading accuracy														
Reading words total % errors	0.6	0.8	0.1	0.4	**0.7**	**<.001****	**1.1**	0.7	1.1	**-0.1**	.56	0.1	**0.8**	**<.001****
Reading regular words % errors	0.6	1.6	0.1	0.7	**0.3**	.15	0.6	0.2	0.9	**0.3**	.26	0.4	**0.0**	.64
Reading irregular words % errors	1.3	2.3	0.1	0.5	**0.5**	**.01***	**1.0**	2.3	3.7	**-0.4**	.22	0.5	**0.9**	**<.001****
Reading non-words % errors	5.8	5.7	3.9	6.8	**0.3**	.19	0.4	7.7	10.6	**-0.4**	.88	0.3	**0.7**	.08
Language/lexical access														

Cognitive task	YC		OC					AwPKU					AwPKU vs. OC	
	Mean	SD	Mean	SD	z-score	p-value	Cohen's d	Mean	SD	z-score	p-value	Cohen's d	diff in z-scores	p-value
Verbal fluency (letters) total	41.8	12.9	47.3	11.3	**0.4**	.06	0.6	36	11.4	**-0.4**	.06	0.7	**0.8**	**<.001****
Stroop congruent % errors	0.0	0.2	0.0	0.2	**0.1**	.76	0.0	0.0	0.0	**0.2**	.33	0.0	**-0.1**	.32
Stroop incongruent % errors	0.8	1.0	0.1	0.3	**0.6**	**<.001****	**1.3**	0.7	1.1	**0.1**	.76	0.1	**0.5**	**.02***
Stroop neutral % errors	0.1	0.3	0.0	0.0	**0.3**	.08	0.7	0.1	0.4	**-0.1**	.76	0.0	**0.4**	.1
Picture naming % errors	6.5	4.0	4.6	4.3	**0.5**	.08	0.6	7.4	3.8	**-0.2**	.45	0.3	**0.7**	**.01***
Phoneme deletion % errors	11.2	8.7	12.9	15.5	**-0.2**	.51	0.2	14.4	13.0	**-0.4**	.24	0.4	**0.2**	.64
Spoonerism % errors	6.8	6.6	9.9	16.2	**-0.5**	.25	0.4	10.8	14.5	**-0.6**	.18	0.5	**0.1**	.8
Visual search accuracy														
Simple detection % errors	1.3	2.2	1.5	3.0	**-0.1**	.75	0.1	1.5	2.3	**-0.1**	.84	0.1	**0.0**	.91
Choice RT % errors	0.4	0.8	0.9	1.3	**-0.6**	.07	0.7	0.5	0.8	**-0.1**	.68	0.2	**-0.5**	.12
Detection with distractors (Ladybird) % errors	1.0	0.9	0.8	1.2	**0.3**	.33	0.3	1.4	2.1	**-0.3**	.45	0.4	**0.6**	.19
Attentional switch (Bug) % errors	0.2	0.5	0.5	1.3	**-0.5**	.21	0.4	0.3	0.6	**-0.1**	.76	0.3	**-0.4**	.3
Detection with distractors overall % errors	0.6	0.6	0.7	1.0	**0.0**	.71	0.2	0.8	1.1	**-0.3**	.44	0.3	**0.3**	.66
Feature search overall % errors	2.0	2.3	0.7	1.4	**0.6**	**.01***	**1.0**	1.6	2.6	**0.2**	.5	0.2	**0.4**	.1
Conjunction search overall % errors	3.1	4.4	5.6	11.6	**-0.6**	.21	0.4	2.4	2.7	**0.2**	.44	0.3	**-0.8**	.12

Cognitive task	YC		OC					AwPKU					AwPKU vs. OC	
	Mean	SD	Mean	SD	z-score	p-value	Cohen's d	Mean	SD	z-score	p-value	Cohen's d	diff in z-scores	p-value
Short Term Memory														
Digit span	6.5	0.9	6.3	1.1	**-0.3**	.18	0.3	6.1	1.0	**-0.5**	.05	0.6	**0.2**	.45
Non-word repetition % err.	39.3	10.4	52	17.2	**-1.2**	**<.001****	**1.3**	48.2	12.5	**-0.9**	**<.001****	**1.1**	**-0.3**	.27
Corsi span	5.7	0.9	4.6	1.3	**-1.2**	**<.001****	**1.4**	5.3	0.9	**-0.4**	.1	0.6	**-0.8**	**<.001****
Inhibition														
Stroop incongruent-congruent % errors	0.7	1.0	0.1	0.3	**0.6**	**.01***	**1.1**	0.7	1.1	**0.0**	.87	0.0	**0.6**	**.01***
Stroop incongruent-congruent RT	92.4	51.6	130.8	122.9	**-0.7**	.07	0.6	117.8	80.1	**-0.5**	.15	0.5	**-0.2**	.58
Trail Making B-A	20.7	13.3	26.2	20.6	**-0.4**	.14	0.4	19.0	10.8	**0.1**	.58	0.2	**-0.5**	**.03***
Detection with distractors bug-ladybird % errors	-0.8	0.9	-0.3	1.6	**0.6**	.08	0.5	-1.1	2.2	**-0.3**	.52	0.3	**0.9**	.09
Detection with distractors bug-ladybird RT	4.2	38.4	13.9	58.1	**-0.3**	.39	0.3	13.8	33.0	**-0.2**	.3	0.4	**-0.1**	.99
Executive Function														
WCST Total errors	10.8	4.9	15.2	10.3	**-0.9**	**.01***	**0.8**	14.0	8.2	**-0.7**	.05	0.7	**-0.2**	.56
Tower of Hanoi solved trials	8.7	0.7	8.2	1.6	**-0.7**	.08	0.6	7.3	2.5	**-2.8**	**<.001****	**1.1**	**2.1**	.07
Tower of Hanoi % errors	2.9	7.9	8.5	18.3	**-0.7**	.08	0.6	24.7	31.6	**-2.8**	**<.001****	**1.3**	**2.1**	**.02***

Cognitive task	YC		OC					AwPKU					AwPKU vs. OC	
	Mean	SD	Mean	SD	z-score	p-value	Cohen's d	Mean	SD	z-score	p-value	Cohen's d	diff in z-scores	p-value
Verbal fluency (semantic)	25.0	5.0	20.1	4.8	**-1.0**	**<.001****	**1.4**	21.4	6.2	**-0.7**	**.01***	**0.9**	**-0.3**	0.3
Sustained attention														
RVP % errors	13.2	8.8	25.7	16.1	**-1.4**	**<.001****	**1.4**	18.9	11.3	**-0.6**	**.02***	**0.8**	**-0.8**	**.04***
LTM & Learning														
Picture non-word paired associates % errors	47.0	24.9	64.6	19.5	**-0.7**	**<.001****	**1.1**	43.5	22.9	**0.1**	.58	0.2	**-0.8**	**<.001****
Picture non-word paired associates delayed recall % errors	29.3	24.5	59.4	34.0	**-1.2**	**<.001****	**1.4**	20.4	23.7	**0.4**	.16	0.5	**-1.6**	**<.001****
Rey A learning % errors	20.5	8.4	36.0	11.8	**-1.8**	**<.001****	**2.1**	24.1	11.2	**-0.4**	.17	0.5	**-1.4**	.05
Rey A immediate recall % errors	15.8	16.1	35.1	22.6	**-1.2**	**<.001****	**1.4**	18.3	15.6	**-0.2**	.54	0.2	**-1.0**	**<.001****
Rey A delayed recall % errors	15.1	15.1	30.9	23.4	**-1.6**	**<.001****	**1.1**	16.3	15.9	**-0.1**	.76	0.1	**-1.5**	**<.001****
Rey B % errors	52.2	11.6	67.0	12.7	**-1.3**	**<.001****	**1.7**	56.6	11.8	**-0.4**	.15	0.5	**-0.9**	**<.001****
Delayed match to sample % errors	10.5	8.3	20.5	11.5	**-1.2**	**<.001****	**1.4**	14.1	8.5	**-0.4**	.08	0.6	**-0.8**	**.01***
Picture location learning % errors	1.9	1.9	5.1	4.8	**-1.7**	**<.001****	**1.2**	3.1	3.8	**-0.6**	.14	0.6	**-1.1**	**.04***
Visuo-Motor Control														
Grooved pegboard	62.7	5.9	84.2	19.2	**-3.6**	**<.001****	**2.1**	69.4	12.7	**-1.1**	**.01***	**1.0**	**-2.5**	**<.001****
Digit symbol % errors	27.1	10.2	44.1	12.0	**-1.7**	**<.001****	**2.2**	37.3	11.3	**-1.0**	**<.001****	**1.3**	**-0.7**	**.01***
Trail Making A	19.9	4.7	29.2	8.7	**-2.0**	**<.001****	**1.9**	25.1	9.9	**-1.1**	**.01***	**0.9**	**-0.9**	.05

Cognitive task	YC		OC					AwPKU					AwPKU vs. OC	
	Mean	SD	Mean	SD	z-score	p-value	Cohen's d	Mean	SD	z-score	p-value	Cohen's d	diff in z-scores	p-value
Language RTs														
Reading words overall RT	519.5	93.9	622.4	122.7	**-1.1**	**<.001****	**1.3**	576.8	123.2	**-0.6**	.08	0.7	**-0.4**	.12
Reading regular words RT	522.4	89.7	625.6	133.2	**-1.2**	**<.001****	**1.3**	582.8	133.6	**-0.7**	**.05***	**0.8**	**-0.3**	.64
Reading irregular words RT	551.3	128.8	641.2	129.3	**-0.7**	**.01***	**1.0**	626.2	180.2	**-0.6**	.08	0.7	**-0.1**	.07
Reading non-words RT	624.3	104.3	840.1	248.0	**-2.1**	**<.001****	**1.6**	818.9	312.8	**-1.9**	**.005***	**1.2**	**-0.2**	.77
Stroop neutral RT	604.5	92.0	836.1	174.3	**-2.5**	**<.001****	**2.4**	701.6	129.6	**-1.2**	**<.001****	**1.2**	**-1.3**	**<.001****
Stroop congruent RT	639.2	111.0	913.7	217.4	**-2.5**	**<.001****	**2.2**	758.9	151.7	**-1.1**	**<.001****	**1.3**	**-1.4**	**<.001****
Stroop incongruent RT	731.6	134.9	1044.5	226.0	**-2.3**	**<.001****	**2.4**	876.7	206.1	**-1.1**	**<.001****	**1.2**	**-1.2**	**<.001****
Picture naming RT	825.7	125.3	981.6	171.1	**-1.2**	**<.001****	**1.5**	890.0	130.8	**-0.5**	**.02***	0.7	**-0.7**	**.01***
Visual search RT														
Simple detection RT	315.9	57.0	453.0	114.1	**-2.4**	**<.001****	**2.1**	331.7	52.9	**-0.3**	.27	0.4	**-2.1**	**<.001****
Choice reaction RT	281.3	31.3	387.5	108.5	**-3.4**	**<.001****	**1.9**	306.9	420.0	**-0.8**	**.01***	0.1	**-2.6**	**<.001****
Detection with distractors (Ladybird) RT	405.1	83.6	513.0	98.3	**-1.3**	**<.001****	**1.7**	431.2	68.4	**-0.3**	.19	0.5	**-1.0**	**<.001****
Attentional switch (Bug) RT	409.3	64.6	526.9	110.0	**-1.8**	**<.001****	**1.8**	445.0	76.3	**-0.6**	.05	0.7	**-1.2**	**<.001****
Detection with distractors overall RT	407.2	72.2	519.4	99.1	**-1.6**	**<.001****	**1.8**	438.1	70.6	**-0.4**	0.1	0.6	**-1.2**	**<.001****
Feature search RT	498.4	71.7	818.9	245.3	**-4.5**	**<.001****	**2.5**	606.0	167.6	**-1.5**	**<.001****	**1.2**	**-3.0**	**<.001****
Conjunction search RT	840.5	126.3	1318.7	383.0	**-3.8**	**<.001****	**2.4**	1007.9	239.7	**-1.3**	**<.001****	**1.2**	**-2.5**	**<.001****

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Discussion

This study aimed to increase our understanding of the likely impact of increasing age on cognition in AwPKU. Based upon the existing literature surrounding cognition, neuropathology, and neurophysiology in older adults and AwPKU, we posited 5 key hypotheses:

1. Cognitive impairment associated with white matter abnormalities (namely speed of processing) and with prefrontal or striatal cortical neurotransmitter depletion (namely executive function, and visuo-motor coordination) will be visible in both AwPKU and older control populations.

2. Cognitive impairment associated with hippocampal neurotransmitter depletion will be observed in older controls only. We expect older controls to perform poorly on memory and learning tasks, but that this domain will be spared in AwPKU.

3. We expect to see some evidence of impaired general intelligence in AwPKU, relative to younger controls, whilst general intelligence in older controls will be spared.

4. We expect to see impaired performance from AwPKU in measures of sustained attention, but that this domain will be spared in older controls.

5. We expect to find no impairment in tasks measuring inhibition and vocabulary knowledge in either population, due to crystallised knowledge remaining intact, and the use of compensatory strategies.

Overlap in impairment between populations

As predicted in hypothesis 1, a large overlap in impairments associated with either white matter abnormalities, or prefrontal neurotransmitter depletion, was observed in OC and AwPKU cohorts. Both AwPKU and OC showed significantly slower RTs than YC across tasks, consistent with previous studies (for OC see Der & Deary, 2006; Salthouse 2004; Sliwinske & Buschke, 1999; Volkow et al., 1998; for AwPKU see Moyle et al., 2007; Ris et al., 1994; Weglage et al., 2013). This may be due to a reduction in speed of processing, caused by white matter degradation which has been reported in both populations (for OC see Bracco et al., 2005; Jiang et al., 2018; for AwPKU see Dyer, 2000). Furthermore, AwPKU showed normal RTs in a simple detection task, where participants had to press a key when a stimulus appeared on the screen, while OC were impaired in this task,

suggesting that there is a difference in severity of processing speed impairments in these two populations. Alternatively, differences in simple RT performance could be due to differential patterns of atrophy, as PKU has been found to specifically impact white matter (the oligodendroglia) and peripheral axons (Schwann cells; Dyer, 2000), while ageing has more pervasive effects, impacting periventricular and cerebellar white matter, as well as cerebellar grey matter, found to be associated with processing speed impairments due to increasing age (Bracco et al., 2005; Eckert et al., 2010; Gorbach et al., 2017; Jiang et al., 2018; Koppelmans et al., 2015).

Both AwPKU and OC were impaired in tasks tapping visuo-motor coordination, as previously demonstrated (for OC see Hamilton et al., 2017; Hoyer et al., 2004; Joy et al., 2000; Volkow et al., 1998; for AwPKU see Moyle et al., 2007; Weglage et al., 1995), with both groups performing slower than YC on grooved pegboard and Trail Making A tasks, as well as having a significantly higher error rate in digit symbol coding. This slowing may, in part, reflect striatal dopamine depletion in both populations (for evidence of dopaminergic depletion in OC see Backman et al., 2010; Kaasinen et al., 2000; Wang et al., 1998; for evidence in AwPKU see Diamond et al., 1997; De Groot et al., 2010; Ogawa & Ichinose, 2006; Van Spronsen et al., 2009). This impairment, however, was significantly more severe in OC than in AwPKU across tasks (excepting Trail Making A, $p=.05$) suggesting these impairments in OC may also be partially attributable to cerebellar abnormalities associated with increasing age, which have previously been found to be linked to impaired motor coordination and fine motor skills in this population (Eckert et al., 2010; Koppelmans et al., 2015).

Finally, both AwPKU and OC were impaired in tasks tapping executive function. Impairments were evident in the WCST, Tower of Hanoi, semantic fluency and digit symbol coding tasks, which require planning, updating, and task-switching abilities (digit symbol coding was previously included among the tasks requiring visuo-motor coordination discussed above, however this is a complex task which also requires working memory and updating abilities) supporting existing findings of executive function deficits in both of these populations (for OC see Bucur & Madden, 2010; De Beni & Palladino, 2004; Delaloye et al., 2009; for AwPKU see Brum et al., 2004; DeRoche & Welsh, 2008; Nardecchia et al., 2015). This time, however, more severe impairments were apparent in AwPKU, possibly indicative of a more severe impact of prefrontal cortical dopamine depletion in AwPKU.

Differences between populations

Whilst there was a large overlap in the tasks spared and impaired in AwPKU and OC, there were also a number of significant differences. As predicted by hypotheses 2 and 3, certain cognitive domains were found to be impaired in one group only, with IQ only impaired in AwPKU, and LTM and learning only impaired in OC. These findings are consistent with results from previous studies which found no learning and memory impairment in AwPKU (Channon et al., 2004; Moyle et al., 2007; Nardecchia et al., 2015; Palermo et al., 2017). Instead, impairments in these domains have been consistently associated with increasing age (Craik, 1992; Gorbach et al., 2017; Joy et al., 2000; Park et al., 2002; Salthouse, 2002) and possibly linked to hippocampal abnormalities not typically found in individuals with PKU (Gorbach et al., 2017; Papenberg et al., 2014). It is noteworthy, however, that prefrontal serotonin depletion and striatal dopamine depletion have also been associated with LTM deficits in older adults (de Quervain et al., 2003; Erixon-Lindroth et al., 2005; Koppel & Goldberg, 2009), and similar depletions have also been observed in younger AwPKU. It may be, therefore, that prefrontal neurotransmitter depletions (found in both populations) interact with hippocampal depletions (including serotonin, dopamine, and glutamate) associated only with ageing, therefore resulting in LTM impairments not observed in PKU populations.

Contrary to our predictions in hypothesis 4, sustained attention was found to be significantly impaired in both OC and AwPKU in comparison to YC. An impairment in this domain in AwPKU is consistent with existing literature which has found that high Phe levels in AwPKU are associated with poorer performance on sustained attention tasks (Bik-Multanowski et al., 2011; Jahja, van Spronsen, et al., 2017; Romani et al., 2017). Existing literature with older adults, however, has suggested that sustained attention in older adults is spared, or even improved with age (Robison et al., 2022; Vallesi et al., 2021). These studies, however, utilised a psychomotor vigilance task, and a go-no-go task respectively, therefore slowed speed of processing did not impact upon participants ability to complete the task, just the speed with which they could respond to stimuli. Our RVP task consisted of quickly changing numbers presented on screen, wherein participants had to press a button when particular sequences of numbers appeared. Participants, therefore, had to not only sustain their attention to the task, but also quickly process the information being presented to them to allow them to press the response key before the window for identifying the sequence had passed. It could then be that OCs severely impaired speed of processing is largely responsbile for the high number of errors and missed responses on this task, rather than an impairment in sustained attention specifically. This theory is supported by findings by

McAvinue et al. (2012) reported a small but significant decline in sustained attention when pariticpants were assessed using a continuous performance task, similar to our RVP task.

Domains spared in both populations

As predicted by hypothesis 5, both AwPKU and OC generally demonstrated good accuracy across domains, as well as good inhibition abilities when compared to YC, consistent with previous results (for OC see Verhaeghen & De Meersman, 1998; Hamilton et al., 2017; Delaloye et al., 2009; for AwPKU see Brumm et al., 2004; Craik, 1994; Salthouse, 2004; Sliwinski & Buschke, 1999). Excellent performance from OC in the verbal IQ subtest of the WAIS (particularly in the vocabulary subscale), as well as in letter fluency (which has a strong lexical access component), word spelling, reading irregular words, and accurate picture naming indicate that OC maintain very good lexical knowledge as demonstrated by previous studies (Craik, 1994; Park et al., 2002; Salthouse, 2004). Similar results were also apparent in AwPKU, although to a lesser extent than in OC.

Summary

In summary, although, there was a large overlap in the domains and tasks spared and impaired in both groups, there were differences in severity, sometimes favouring OC (with better performance in IQ, executive function, and language accuracy in this group) and other times favouring AwPKU (with better speed of processing, learning and memory and visuo-motor coordination in this group). There were also domains where only one group was impaired. Double dissociations were apparent between IQ and memory tasks, with IQ impaired in AwPKU but not in OC, whilst memory was impaired in OC but not in AwPKU. These differences are important because they show that the pattern of spared and impaired abilities seen in AwPKU and in OC is not just a result of severity and task difficulty (with the task more impaired being the most difficult), but due to specific ways in which neurological circuits and associated cognitive functions are affected by PKU and ageing.

High levels of Phe in PKU seem to specifically affect frontal lobes and associated executive functions. Ageing, instead, seems to affect less the frontal lobes and more the hippocampal system involved in memory and learning, and the cerebellum involved in visuo-motor coordination. Both groups share a reduction in speed of processing, although this is more marked in OC, which may be related to white matter abnormalities in both populations.

Limitations

Whilst this study provides some insight into how increasing age may affect cognition in AwPKU, it is purely extrapolatory at this stage. Close monitoring of AwPKU as they reach middle and older age will provide a much clearer picture of how decreasing brain health due to old age may interact with existing abnormalities caused by PKU.

Additionally, results of exploratory analyses comparing the performance of different cohorts on individual tasks are also limited, due to the increased risk of a Type I error when t-test analyses are conducted across a large number of individual measures. Whilst this limitation was recognised, the potential to identify key differences within a cohort of participants with a rare clinical disorder was considered sufficiently important to conduct analyses across the entire assessment battery with a more liberal p-value than may otherwise have been applied to such a large number of analyses. Rather than being taken individually, these differences could then contribute towards a better understanding of the overall patterns of impairment demonstrated by each cohort.

Implications and future directions

These findings suggest that, whilst some domains are likely to remain unaffected by age in individuals with PKU, we can expect some interaction between increasing age and PKU with regards to impairments caused by white matter atrophy and both prefrontal and striatal neurotransmitter depletion. It may be the case that executive functions and processing speeds, in particular, are dramatically affected by ageing in AwPKU as already damaged white matter atrophies further. Similarly, peripheral abilities, such as visuo-spatial attention and motor control, may also prove to be particularly vulnerable to the effects of ageing in people with PKU. There is evidence, however, that white matter abnormalities in PKU can be reversed (to some extent), and dopamine depletion mitigated, by strict dietary adherence (Bick et al., 1993; Cleary et al., 1995; Palermo et al., 2017; Thompson et al., 1990), suggesting that it may be possible to attenuate the dramatic impairments that could otherwise result from ageing in PKU.

Future research is needed to better understand the differences in neurological atrophy and neurotransmitter depletions due to age and PKU as this may help clinicians to develop a clearer understanding of how these neurological abnormalities may interact with one another as AwPKU start to reach old age. Additional longitudinal studies will also be of paramount importance in the near future, to allow researchers to monitor cognition and brain

health in AwPKU as they move into older age, and to investigate the potential mitigating impact of resuming a low Phe diet in later life.

Chapter 4: Speed of Processing and Executive Function in Ageing and PKU

Introduction

In the previous chapter, our results indicated that there is a significant overlap between the patterns of cognitive impairment demonstrated by young AwPKU those associated with ageing. One key area of impairment in both of these cohorts was an apparent slowed speed of processing across tasks tapping visuo-spatial attention, such as choice reaction time and visual search tasks (plus Simple RT tasks in OC), as well as in tasks tapping language (i.e., picture naming, reading words and non-words, and Stroop congruent and neutral conditions) and visuo-motor coordination (i.e., digit symbol coding and grooved pegboard). Slower responses across these tasks were found to be more pronounced in OC participants than in AwPKU, with OC demonstrating large z-scores and effect sizes compared to YC across tasks in these domains, whilst AwPKU demonstrated moderate to large z-scores and effect sizes across tasks. In addition to an apparent slowed processing speed, both cohorts demonstrated an independent deficit in the domain of executive function, in particular in tasks requiring planning, reasoning, and cognitive flexibility (i.e., WCST, Tower of Hanoi, and semantic fluency), with both groups demonstrating large z-scores across these tasks. In contrast to findings with visuo-spatial attention, however, slowing in these tasks was more severe in AwPKU than in OC, with OC demonstrating moderate to large effect sizes across tasks, whilst AwPKU demonstrated large effect sizes throughout. Of particular importance, deficits in executive function were apparent in both cohorts, even in tasks that did not contain any timed element that may be affected by speed of processing impairments.

The prefrontal-executive hypothesis of age decline (West, 1996) posits that changes in frontal brain areas (such as the neurotransmitter depletions observed with age and PKU) can cause specific executive function deficits, which can then go on to cause more general deficits, including impaired speed of processing (see Albinet et al., 2012). It could be, therefore, that slowed RTs in AwPKU and OCs are due to deficits in planning and inhibitory control, rather than resulting from a generalised processing speed impairment. Slow response times in verbal fluency and visual search tasks reported in the previous chapter, then, may not be the result of an overall slowed speed of processing across domains, but rather due to an impaired ability in these populations to carry out an effective search of either

the mental lexicon or of a visual display, and/or a deficit in inhibitory control preventing participants from ensuring that they do not revisit locations that have already been explored.

Executive function and processing speed abilities are closely linked to one another, with both abilities developing from childhood through to adulthood, and then showing evidence of a decline in later life (for processing speed see: Cerella & Hale, 1994; Jenkins et al., 2000; Kail 1991a, 1991b; Nettlebeck & Burns, 2010; for executive function see: Anderson, 2002; Best & Miller, 2010; Romine & Reynolds 2005), as well as improvements in one ability often contributing to improvement in the other (e.g. Christ et al., 2001; Fry & Hale, 1996). Whilst deficits in processing speed and executive function often co-occur across populations, this does not necessarily mean that they are causally related to one another, nor that they depend on the same underlying neurological mechanisms. Disentangling the relative contribution of these two cognitive domains to impaired performance across difference tasks, however, is a difficult task.

The relative contribution of speed of processing and executive function to increased response times in visuo-spatial and language tasks in AwPKU was previously investigated by Romani et al. (2018). That study compared how the performance of AwPKU and YC changed as a function of task difficulty, in particular using variables which reflect the need to inhibit distractors while searching either a visual display or a lexical space. If AwPKU are slower at processing in a particular task, therefore, differences with controls would be expected to increase with variables of increasing difficulty. Thus, in visuo-spatial tasks, comparisons in performance were made considering the number of distractors and whether or not the target was present in the display. In a picture naming task, comparisons considered the number of semantic distractors that preceded the target. In reading tasks, the effects of frequency, regularity, and word length were considered.

In visual search tasks, AwPKU showed an enhanced effect of variables which are known to increase response times in a visual search (i.e., longer response times in displays with increased numbers of distractors and an absent target), with differences between AwPKU and controls found to increase exponentially in line with task difficulty. This is consistent with AwPKU having a reduced ability to allocate visual attention. In picture naming tasks, however, differences in RTs between AwPKU and controls remained stable across difficulty conditions (i.e., ordinal position of semantically related items) suggesting that no specific processing deficit exists for language tasks in AwPKU. Differences in reading times for high and low frequency, high and low regularity, and very short to long words were also stable for AwPKU compared to YC. In contrast, AwPKU showed increased reading

times with non-words of increased length, again consistent with a deficit in allocating attention (in this case to orthographic units).

Romani et al. (2018) suggested that slow RTs across tasks may reflect domain-specific impairments, rather than resulting exclusively from a generalised speed of processing deficit. In particular, slow RTs and good accuracy in language tasks from AwPKU may, therefore, be due to AwPKU exercising more caution in their responses but remaining largely unimpaired in their lexical processing abilities. This is in contrast to visuo-spatial abilities, where differences in RTs may reflect a specific impairment in the speed at which visuo-spatial attention is allocated.

Chapter 3 demonstrated that, similar to AwPKU, older adults demonstrate slowed responses in tasks tapping both the visuo-spatial and lexical domains, as well as showing impaired executive function abilities. If impaired performance in older adults is, therefore, caused by similar mechanisms as those observed in AwPKU, we should expect to see the same differences in the mechanisms underlying slowed performance in lexical vs. visuo-spatial tasks reported by Romani et al. (2018) in both cohorts.

Objectives

This chapter aims to provide further evidence of similarities between AwPKU and OC by carrying out comparisons similar to those carried out by Romani et al. (2018). In other words, we will expand on that study by adding analyses with OC to analyses comparing AwPKU with YC. We investigate distributions of RTs across trials for visuo-spatial and language tasks, to provide key information about the nature of potential speed impairments in these populations, whilst also comparing the similarities and differences in the underlying cause of slower response times in both groups.

If results from these analyses indicate similar effects of exponentially increasing impairment with increased difficutly across both language and visuo-spatial tasks in OC, this could indicate either a generalised speed of processing deficit in OC, or a supra-modal impairment in inhibitory control, affecting tasks across domains. This would then suggest that the mechanisms underlying slow RTs in OC and AwPKU are differential, and less likely to interact as AwPKU get older.

Alternatively, findings of cumulative impairment with task difficutly in the visuo-spatial domain only (as reported by Romani et al., 2018) would indicate that there may be a

domain-specific speed of processing impairment in this domain, affecting the ability to allocate attention and/or inhibit repeated searches of previously visited areas. If language RTs, meanwhile, show a fixed delay in RTs relative to younger controls, this would indicate that there is no specific impairment in the language domain, with slower RTs possibly reflective of an executive deficit, linked to the ability to weigh evidence and make decisions, and possibly leading to increased caution in responses of affected cohorts.

Hypotheses

We believe that the mechanisms underlying cognitive impairments in OC and AwPKU are strongly linked, therefore we hypothesise that:

1. Both AwPKU and OC will demonstrate a fixed delay in the lexical domain, with the difference in RTs from YC remaining stable across difficulty conditions in language tasks.

2. Both AwPKU and OC will demonstrate an exponential impact of task difficulty on RTs in the visuo-spatial domain, with differences in RTs from YC getting progressively larger as task difficulty in visual search tasks increases.

3. We expect to find correlations between performance in executive function tasks and RTs in tasks of increasing difficulty in both visual and language tasks, reflecting the fixed impact of executive function impairments on RTs across domains.

Method

PKU Participants.

Thirty-seven early-treated young adults, aged 18 to 41 years (mean=27.5), with classical PKU were recruited from the Department of Inherited Metabolic Disorders at the University Hospitals Birmingham. All participants were diagnosed through new-born screening conducted 5–7 days after birth and were continuously treated with a low-Phe diet from diagnosis. Data on historical Phe levels was gathered via the PKU database at The Clinical Chemistry Department at Birmingham Children's Hospital. These participants were recruited and assessed prior to the start of this PhD. At time of testing, seven participants were on an unrestricted diet and 30 were following a low-Phe diet. All early-treated PKU

individuals attending the clinic were invited to participate, as well as a number of individuals who were not currently attending clinic follow-up appointments but were still contactable. All individuals who responded to the invitation were tested.

Participants for this group were recruited and assessed prior to the undertaking of this PhD. Data from this cohort has previously been published in Palermo et al. (2017) and Romani et al. (2017, 2019). It has also been previously presented in Chapter 3 of this thesis.

Control Participants

Controls consisted of a group of 30 young, healthy adults aged 18-41 (mean=27.6), matched with PKU participants for age, gender, and educational status, and a group of older, healthy adults aged 53 to 88 years (mean=69) (Table 4.1.). Older and younger controls were matched for years of education and gender. Healthy volunteers were recruited through the Aston University volunteering website.

Participants for this group were recruited and assessed prior to the undertaking of this PhD. Data from the cohort of young controls has previously been published in Palermo et al. (2017) and Romani et al. (2017, 2019). Data from the cohort of older controls has not previously been published. Data from both control cohorts has previously been presented in Chapter 3 of this thesis.

Table 4.1. – *Demographic information for each participant group*

		Sex		Age (yrs.)			Education (yrs.)	
Participant Type	**N**	**% Male**	**% Female**	**Mean**	**SD**	**Range**	**Mean**	**SD**
AwPKU	37	35	65	27.5	7.3	18-41	14.4	2.0
Younger control	30	33	67	27.6	7.4	18-41	15.2	1.7
Older control	56	37.5	62.5	69.2	7.0	53-88	16.2	3.5

Assessments

Visuo-spatial attention

Visual search – Participants were asked to search for a target among a number of distractors on their screen. They were asked to press the ‘m’ key on their keyboard if the target was present, or the ‘z’ key if it was absent (responses were reversed for left-handed

participants). The target in all trials was a red ladybird, whilst distractors in feature search trials were green bugs, and distractors in conjunction search trials were made up of both green and red bugs. The target was presented on screen in 50% of trials, along with either 4, 8, or 12 distractors. A number of conditions were varied within the visual-search tasks, including target feature/conjunction search, target presence/absence, and the number of distractors present (4, 8, or 12). Reaction time and number of correct responses was scored.

Language

Picture naming – Participants were asked to name 120 images, belonging to 24 different semantic categories (5 items in each category). They were asked to respond as quickly as possible. Previous results have shown that RTs increase with the ordinal position of a stimulus in a set of semantically related pictures due to the need to inhibit previous responses from the same semantic category. This is known as 'semantic interference' (Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2007). Task difficulty was then measured in terms of ordinal position of presented images within their semantic category. Reaction time and number of correct responses was scored.

Word and non-word reading – Participants were asked to read aloud, as quickly as possible, individually presented words on a screen. Word stimuli included 140 regularly and irregularly spelled words. 60 words were contrasted in terms of orthographic regularity, 30 being regular, and 30 being irregular (i.e., containing at least one very uncommon grapheme-phoneme correspondence). The remaining words were all regular and split into groups of 40 high frequency and 40 low frequency words. Within each frequency category, there were 10 words of each length category: very short (4 letters), short (5 letters), medium (6 letters), and long (7-9 letters). An additional 40 non-words were created by changing one or two letters in corresponding high-frequency words. Non-words were then divided into length categories of very short, short, medium, and long, with 10 words in each category. RT and number of correct responses was scored.

Stroop task – Participants were asked to name the colour of stimuli presented on a screen. In a neutral, condition they were presented with a string of X's. In a congruent condition, they were presented with a word in the same colour text as the word (e.g. the word 'Green' in green text). In an incongruent condition, they were presented with a word in a difference colour text than the word (e.g. the word 'Blue' in red text). RT and number of correct responses was scored.

Data handling and analysis

This chapter carried out secondary analysis on data collected as part of the previous chapter (for data collection procedures, see Chapter 3 pgs 59-64).

Planned analyses

Brinley plot analyses

Brinley plot analyses were carried out to assess the relationship between performance in language and visuo-spatial domains. Brinley plots entail plotting the mean latencies of one population (Group A) vs. another (Group B) for different tasks/conditions. If measures vary from one another only because of a single dimension, for example task difficulty, that is the same across the two groups, then the dots should fall close to a linear line, because longer RTs in one group will correspond to longer RTs in the other group in a systematic way. Therefore, if there is no difference between the two groups, the plotted dots will fall close to a linear line with slope=1 and an intercept=0 (Fig.4.1.A). If there is a fixed difference, for example because one group has a fixed hesitation before deciding on a response or because of a fixed delay in the motor response, then the dots will fall close to a linear line with slope=1, but with an intercept >0 (Fig.4.1.B). Finally, if one group is faster than the other, then the dots will still fall along a linear line, but with a slope >1 (Fig.4.1.C). This is because the more difficult the task is (i.e., as more processing stages are required) the larger the difference between the two groups should be as delays cumulate.

The below examples all demonstrate Brinley plots wherein participant responses are affected differentially by a single factor. If RTs in different groups are affected by different factors, however, not all dots will fall close to a linear line. For example, as reading and picture naming are affected by language processing abilities, while visuo-spatial RTs are affected by visuo-attentional skills, these abilities will be different in the two groups, therefore we would expect Brinley plots of this data to demonstrate a non-linear function, with the different conditions of language tasks on one line and the different conditions of visuo-spatial tasks on a different line (Fig.4.1.D).

Linear plots, therefore, suggest that a single dimension/impairment (such as motor slowing or hesitation in carrying out a decision) is causing a proportionate increase in RT across conditions in a domain. Conversely, ‘fanned out’ plots suggest that an additional

speed of processing impairment is enhancing the impact of task difficulty on speed of response in a particular domain.

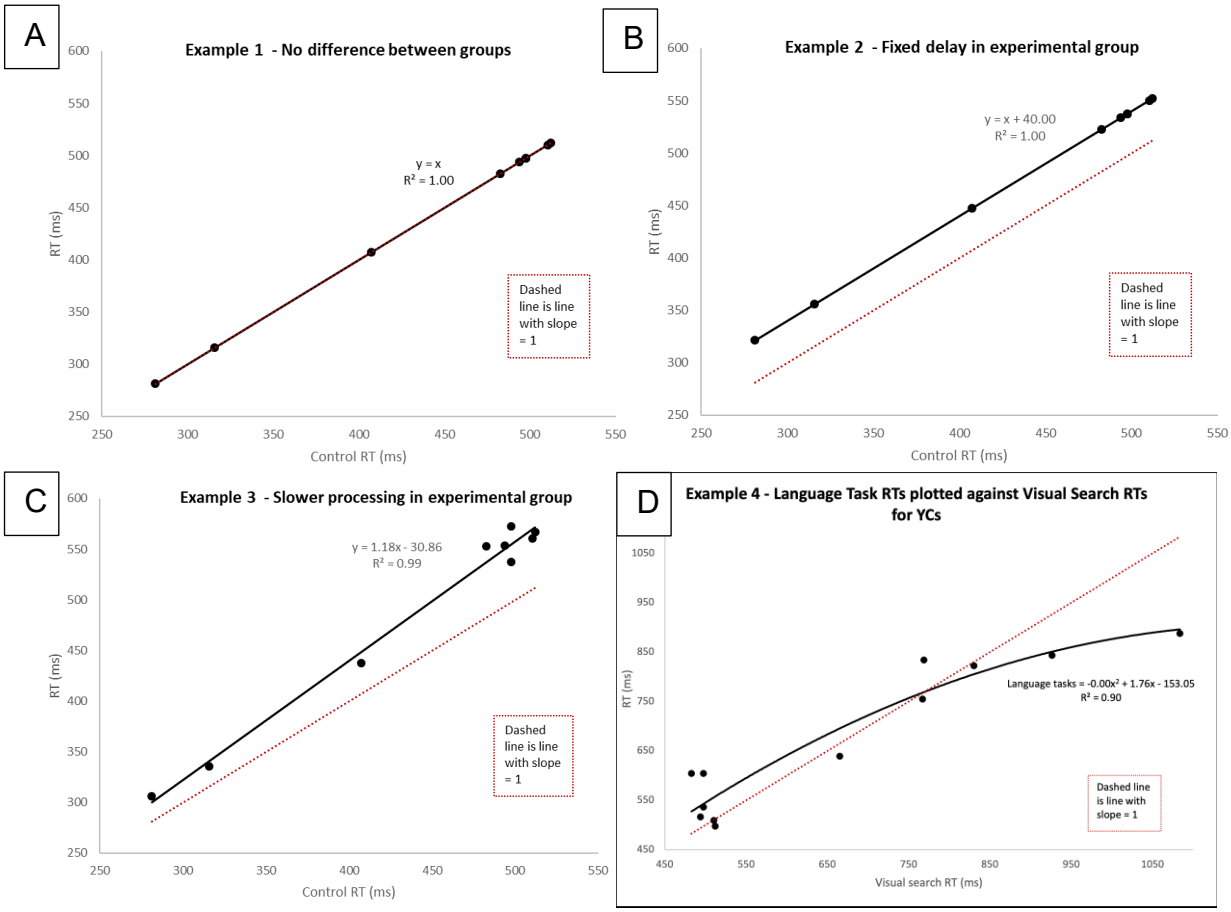


Figure 4.1. – Example Brinley plots demonstrating no difference in RTs between the control group and the experimental group (A), a fixed delay in the experimental group compared to controls (B), slower processing abilities in the experimental group (C) and increasing task difficulty in tasks tapping different domains (D).

Brinley plots for visuo-spatial conditions included simple detection, detection with distractors, choice reaction time, and target present and absent conditions with 4, 8, and 12 distractors for feature and conjunction search tasks. Brinley plots for language conditions included neutral and congruent Stroop conditions, semantic positions 1-5 and fillers for picture naming, reading of irregular and regular words of high and low frequency, and reading of non-words.

General linear analyses

For comparisons where difficulty conditions were categorical (i.e. for language tasks: word frequency, word regularity, word and non-word length; for visual search: target presence) ANOVA analyses were carried out, with difficulty condition as the independent

variable and either RT or error rate as the dependent variable. For comparisons where difficulty conditions were continuous (i.e. for language tasks: picture naming item position; for visual search: number of distractors), Mixed Model Linear (MML) analyses were carried out, with difficulty condition as the independent variable and either RT or error rate as the dependent variable. Where relevant, post-hoc ANOVA or MML analyses were then carried out to further investigate differences between individual group pairings. Line graphs were also plotted to visually represent the relationship between task difficulty and performance across conditions.

Correlation analyses

Finally, to further investigate the interaction between executive function deficits and slower responses across domains, correlation analyses were carried out between performance differences between high and low difficulty conditions in visual search and language tasks, and performance on tasks tapping executive function abilities including inhibition (incongruent vs. congruent Stroop RTs), task switching (WCST errors) and planning (Tower of Hanoi errors).

Supplementary analyses

If performance is affected by a reduced speed of processing in a particular task, slow participants should be more affected by difficulty of condition than fast participants for that task. Instead, if the participants do not have a particular difficulty in the task, but only display a fixed delay in responding, differences across conditions should remain stable across the different groups. Therefore, in line with analyses carried out in Romani et al. (2018), participants within the experimental groups were further split into ‘fast’ and ‘slow’ sub-groups using a performance-based median split. Fast and slow participants were identified by averaging scores across conditions within each task, and then sorting participants into the fastest 50% and the slowest 50%. To demonstrate appropriate rigour in analyses, general linear and Brinley plot analyses as described above were further carried out between conditions and sub-groups for each task to investigate the impact of increasing difficulty on performance. These additional analyses are presented in supplementary materials (Appendix G).

Results

Language tasks

To investigate hypothesis 1, that AwPKU and OC will both demonstrate a fixed delay in language tasks compared to YC, error rates and RTs across difficulty conditions in picture naming and spelling tasks were compared across participant groups. Accuracy and RT results across language tasks are presented in Tables 4.2. and 4.3. Significant differences in RT from YC were demonstrated across all tasks and conditions by OC and across both Stroop and reading task conditions by AwPKU. No significant differences in accuracy were found between AwPKU and YC, whilst only a few differences were apparent between OC and YC in task conditions requiring concomitant executive function abilities. A number of significant differences between OC and AwPKU were found across tasks and conditions both in terms of RT and accuracy.

Picture naming

The impact of semantic interference on performance in picture naming tasks was assessed by carrying out MML analyses with ordinal item position as a continuous within-subjects factor, group as a categorical between-subjects factor, and either RT or error rate as the dependent variable. Analyses of accuracy data found a significant effect of item position on error rate ($F(1,273)=22.7$, $p<.001$) but no main effect of group ($F(2,172)=0.92$, $p=.40$) and **no interaction between the two** ($F(2,273)=0.26$, $p=.77$) (Fig4.2.A). Analyses of RTs also found a significant main effect of both item position ($F(1,263)=41.7$, $p<.001$), and group ($F(2,198)=11.9$, $p<.001$), but **no interaction between these two factors** ($F(2,263)=0.52$, $p=.59$) (Fig4.2.B). These findings indicate that semantic interference did not disproportionally affect OC or AwPKU compared to YC in terms of either accuracy or speed.

The main effect of group on RTs was further investigated through post-hoc MML analyses between individual group-pairings. These analyses found no significant differences between YC and AwPKU ($F(1,124)=2.5$, $p=.12$), but significant differences were apparent between YC and OC ($F(1,137)=22.0$, $p<.001$), and between OC and AwPKU ($F(1,135)=9.8$, $p=.002$). **No significant interactions** between individual group pairings and item position were found, indicating that all cohorts were equally impacted by increasing task difficulty.

Table 4.2. - *Comparison of z-scores across RTs in language tasks from OC and AwPKU. Z-scores have been reversed in some instances so that higher z-scores indicate worse performance across tasks. Z-scores of 0.6 or above in either direction have been highlighted in bold. Differences in z-scores are calculated by subtracting AwPKU from OC z-scores.*

Language task RTs	**Younger Control (YC)**		**Older Control (OC)**			**YC vs. OC**	**Adults with PKU (AwPKU)**			**YC vs. AwPKU**	**OC vs. AwPKU**	
	Mean	**SD**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Diff in z-scores**	**p-value**
Reading words	507.1	93.3	622.4	122.7	**1.2**	<.001**	579.6	115.8	**0.8**	.007**	0.4	.11
Reading non-words	604.3	103.6	840.1	248.0	**2.3**	<.001**	803.3	296.0	**1.9**	.001*	0.4	.57
Stroop neutral condition	604.5	92.0	836.1	174.3	**2.5**	<.001**	701.6	129.6	**1.2**	.001**	**1.3**	<.001**
Stroop congruent condition	639.2	111.0	913.7	217.4	**2.5**	<.001**	758.9	151.7	**1.1**	.001**	**1.4**	.001**
Stroop incongruent condition	731.6	134.9	1044.5	226.0	**2.3**	<.001**	876.7	206.1	**1.1**	.002**	**1.2**	.002**
Picture naming	835.1	138.1	989.0	144.0	**1.1**	<.001**	901.8	139.1	0.5	.07	**0.6**	.02*
Naming position 1	755.0	131.7	925.1	147.5	**1.3**	<.001**	819.7	131.8	0.5	.06	**0.8**	.004*
Naming position 2	829.3	155.9	999.7	166.0	**1.1**	<.001**	903.7	152.9	0.5	.07	**0.6**	.02*
Naming position 3	838.1	142.2	1007.8	158.0	**1.2**	<.001**	896.5	154.6	0.4	.13	**0.8**	.006**
Naming position 4	862.0	145.6	973.7	146.1	**0.8**	<.001**	910.8	144.0	0.3	.20	0.5	.09
Naming position 5	899.5	150.7	1057.4	149.1	**1.0**	<.001**	992.6	168.1	**0.6**	.03*	0.4	.11
Filler items	827.0	145.0	980.4	143.6	**1.1**	<.001**	887.7	150.4	0.4	.012	**0.7**	.02*

* T-test is significant at the 0.05 level (2-tailed).

** T-test is significant at the 0.01 level (2-tailed).

Table 4.3. - *Comparison of z-scores across accuracy scores in language tasks from OC and AwPKU. Z-scores have been reversed in some instances so that higher z-scores indicate worse performance across tasks. Z-scores of 0.6 or above in either direction have been highlighted in bold. Differences in z-scores are calculated by subtracting AwPKU from OC z-scores.*

Language task error rates	**Younger Control (YC)**		**Older Control (OC)**		**YC vs. OC**		**Adult with PKU (AwPKU)**		**YC vs. AwPKU**		**OC vs. AwPKU**	
	Mean	**SD**	**Mean**	**SD**	**z-score**	**p-value**	**Mean**	**SD**	**z-score**	**p-value**	**Diff in z-scores**	**p-value**
Reading words	0.6	0.8	0.0	0.2	**-0.7**	<.001**	0.7	1.1	0.1	.75	**-0.8**	.001**
Reading non-words	5.8	5.7	3.1	4.2	-0.5	.03*	7.7	10.6	0.4	.36	**-0.9**	.02*
Stroop neutral condition	0.1	0.3	0.0	0.0	-0.3	.08	0.1	0.4	0.1	.76	-0.4	.10
Stroop congruent condition	0.0	0.2	0.0	0.2	-0.1	.76	0.0	0.0	-0.2	.33	0.1	.32
Stroop incongruent condition	0.8	1.0	0.1	0.3	**-0.6**	.004**	0.7	1.1	-0.1	.76	-0.5	.02*
Picture naming	7.1	5.1	4.6	4.3	-0.5	.04*	6.9	3.7	0.0	.90	**-0.5**	.02*
Naming position 1	1.1	1.3	0.6	0.9	-0.4	.07	1.3	1.2	0.2	.54	**-0.8**	.01*
Naming position 2	1.9	1.4	1.5	1.6	-0.3	.24	1.7	1.5	-0.2	.54	-0.1	.66
Naming position 3	1.8	1.4	1.4	1.4	-0.3	.23	2.0	1.5	0.1	.59	**-0.4**	.10
Naming position 4	2.0	1.6	1.5	1.8	-0.3	.27	1.8	1.4	-0.1	.67	-0.2	.46
Naming position 5	2.3	2.0	1.2	1.7	**-0.6**	.02*	2.5	2.0	0.1	.70	**-0.7**	.007**
Filler items	2.5	2.4	2.1	2.0	-0.2	.39	2.1	1.7	-0.2	.46	0.0	.87

* T-test is significant at the 0.05 level (2-tailed).

** T-test is significant at the 0.01 level (2-tailed).

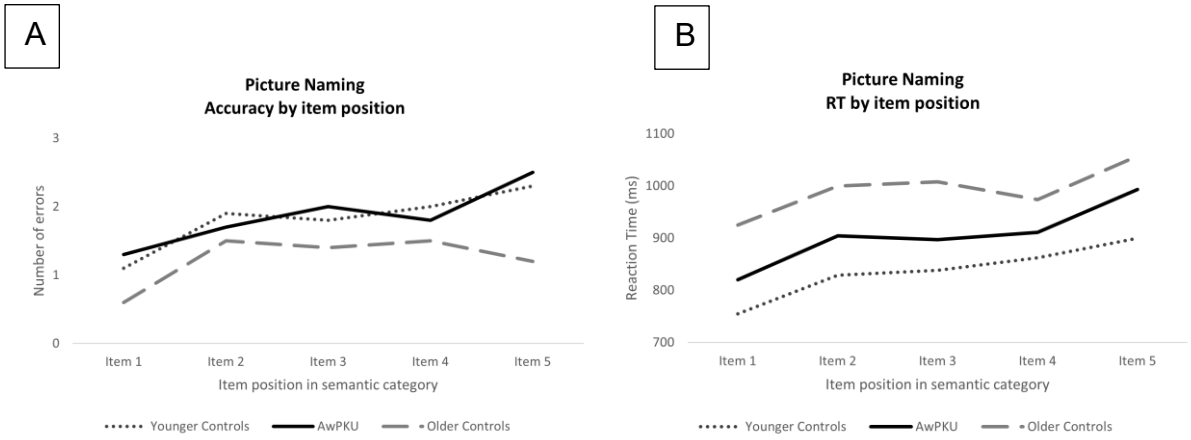


Figure 4.2. – Error rates by item position in semantic category for YC, OC, and AwPKU (A). RTs by item position in semantic category for YC, OC, and AwPKU (B) in picture naming task.

To investigate whether increased RTs for OC and AwPKU were due to an accuracy-speed trade off (wherein increased search times could be attributed to participants exercising more caution before giving their responses) Pearson correlational analyses were carried out between accuracy scores and RTs across groups for picture naming tasks. These analyses found no significant correlation between errors rates and RTs in YC ($r(150)=.15$, $p=.08$) although moderate correlations were found for AwPKU ($r(150)=.26$, $p=.001$) and OC ($r(170)=.28$, $p<.001$). In both instances of significant correlations, slower RTs were associated with higher error rates, demonstrating that slower responses from AwPKU and OC across tasks did not result from a speed-accuracy trade off.

Word Reading

The impact of task difficulty in word and non-word reading tasks was investigated by assessing the effects of regularity and frequency of words, as well as length of words and non-words on accuracy and response times. ANOVAs were carried out using regularity and frequency as within-subject factors, group as the between-subject factor, and RT or error rate as the dependent variable.

Regularity. ANOVA analyses of error rates found no significant main effects of regularity ($F(1,108)=0.14$, $p=.71$, $\eta p^2=.001$). A significant main effect of group was found ($F(2,108)=11.3$, $p<.001$, $\eta p^2=.17$). **No significant interaction** between group and regularity was found ($F(2,108)=0.9$, $p=.41$, $\eta p^2=.02$)(Fig.4.3.A). The main effect of group on accuracy was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that OC produced significantly less errors than both

AwPKU (p<.001) and YC (p=.005), but no difference was found between YC and AwPKU (p=.90), nor any interaction between individual group pairings.

RT analyses found a significant main effect of regularity (F(1,108=23.1, p<.001, ηp^2=.18) with irregular words producing significantly slower RTs than regular words in all groups (Fig.4.3.B). A main effect of group was also found (F(2,108)=6.9, p=.001, ηp^2=.11). There was **no significant interaction between regularity and group** (F(2,108)=1.4, p=.25, ηp^2=.03) demonstrating that all groups were equally impacted by increasing difficulty in terms of word regularity. The main effect of group on RT was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found significant differences between YC and both AwPKU (p=.03) and OC (p=.001), with both groups responding significantly slower than YC.

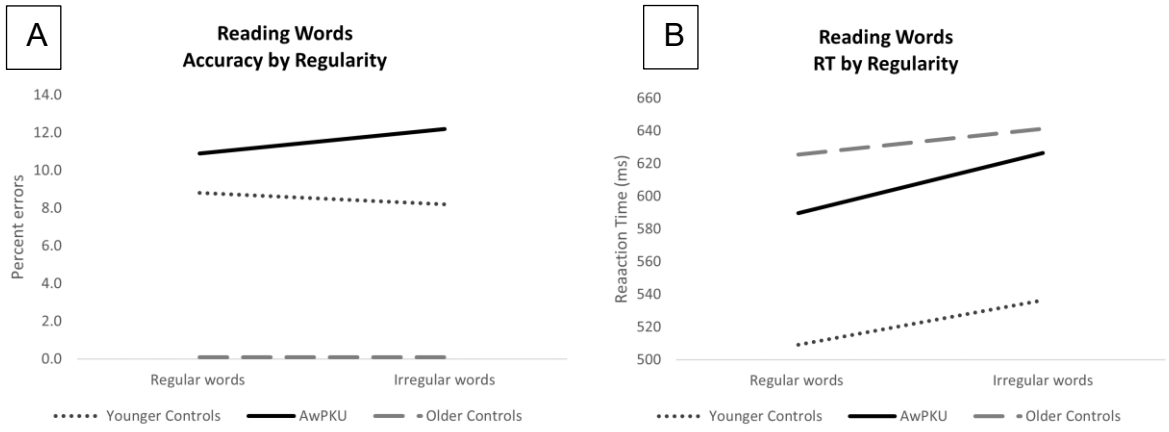


Figure 4.3. – Error rates by word regularity for YC, OC, and AwPKU (A). RTs by regularity for YC, OC, and AwPKU (B) in word reading task.

Frequency. ANOVA analyses of accuracy data between groups found a trend main effect of frequency on error rates (F(1,108)=3.8, p=.054, ηp^2=.03), however the effect was in the opposite direction than expected, with more errors made on high frequency than low frequency words in YC and AwPKU (Fig.4.4.A). A significant effect of group was also found (F(2,108)=9.9, p<.001, ηp^2=.16) but **no significant interaction** between group and word frequency (F(2,108)=1.1, p=.33, ηp^2=.02). The main effect of group on accuracy was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that OC produced significantly fewer errors than both YC (p<.004) and AwPKU (p<.001). No differences in error rates were found between AwPKU and YC, **nor any interactions between individual group pairings and word frequency effects.**

RT analyses found a significant main effect of frequency ($F(1,107=11.8$, $p=.001$, $\eta p^2=.10$) and group ($F(2,107=11.1$, $p<.001$, $\eta p^2=.17$), but **no significant interaction** between group and frequency ($F(2,98)=0.12$, $p=.89$, $\eta p^2=.002$). As expected, higher frequency words elicited faster responses than lower frequency words, demonstrating a possible speed-accuracy trade-off in performance from all groups (Fig.4.4.B). The main effect of group on RT was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that YC were significantly faster that AwPKU ($p=.047$) and OC ($p<.001$). No significant difference was found between AwPKU and OC ($p=.10$). **No interactions between frequency effects and individual group pairings** were found, indicating that all groups were equally impacted by increasing difficulty in terms of word frequency.

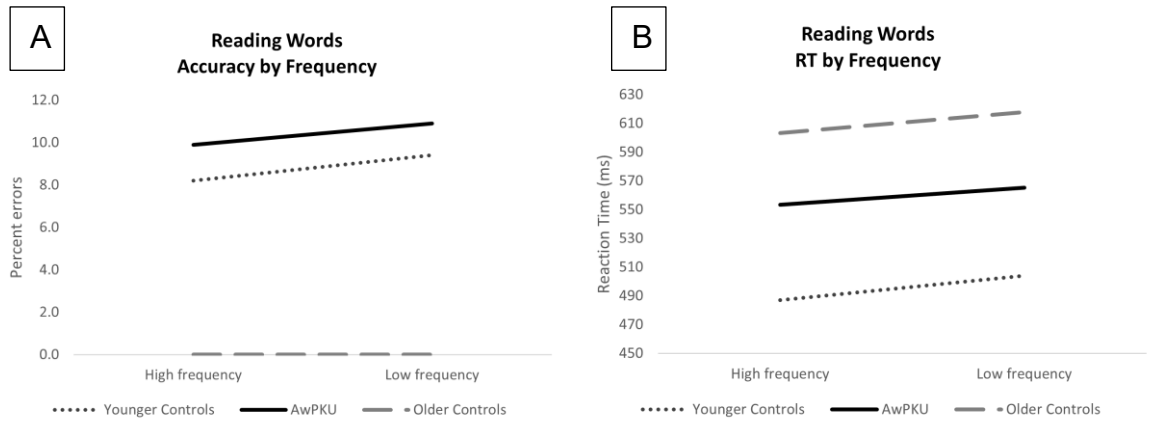


Figure 4.4. – Error rates by word frequency for YC, OC, and AwPKU (A). RTs by frequency for YC, OC, and AwPKU (B) in word reading task.

Word Length. The effects of word length on accuracy and RT of word reading were assessed through ANOVA analyses, with word length (very short: 4 letters, short: 5 letters, medium: 9 letters, long: 7-9 letters) as the within-subject variable, and group as the between-subjects variable. Accuracy analyses found no significant main effect of word length ($F(3,321)=1.8$, $p=.15$, $\eta p^2=.02$) on error rates. A main effect of group was found ($F(2,107)=11.4$, $p<.001$, $\eta p^2=.18$), but **no significant interaction** between these variables ($F(6,321)=1.6$, $p=.15$, $\eta p^2=.03$) (Fig.4.5.A). The main effect of group on accuracy was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that OC performed significantly more accurately than both AwPKU ($p<.001$) and YC ($p=.02$). **No interactions** were found between individual group pairings and word length effects.

RT analyses found significant main effects of word length ($F(3,324=9.5$, $p<.001$, $\eta p^2=.08$), and group ($F(2,108)=11.9$, $p<.001$, $\eta p^2=.18$) on performance, but **no interaction**

between these variables (F(6,324)=1.3, p=.28, ηp^2=.02)(Fig.4.5.B). The main effect of group on RT was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found significant differences in RTs between YC and both AwPKU (p=.04) and OC (p<.001), but no difference between AwPKU and OC (p=.07). **No interactions** were found between individual group pairings and word length effects.

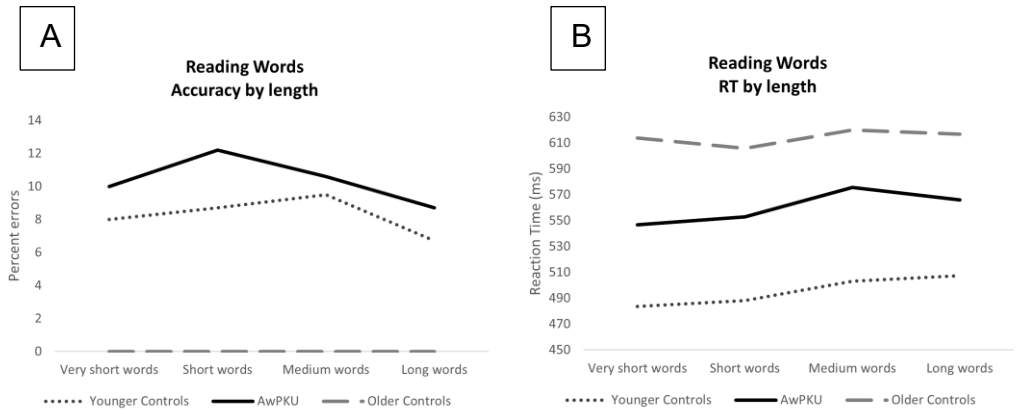


Figure 4.5. – Error rates by word length for YC, OC, and AwPKU (A). RTs by word length for YC, OC, and AwPKU (B) in word reading task.

Non-Word Length. ANOVA analyses investigating the effects of non-word length on accuracy between groups found a significant main effect of non-word length on error rates (F(3,312)=5, p=.002, ηp^2=.05) with a general positive trend of increasing errors with increasing non-word length (Fig.4.6.A). A significant main effect of participant group on accuracy was also found (F(2,104)=6.8, p=.002, ηp^2=.12). **No significant interaction** between the two variables was found (F(6,312)=1.4, p=.21, ηp^2=.03) indicating that accuracy was affected by increasing non-word length equally for all groups, although graphical representation of accuracy by non-word length suggests that AwPKU were notably more impacted by the increase in task difficulty between medium and long words than YC (Fig.4.6.A). The main effect of group on accuracy was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that OC performed more accurately than both YC (p=.007) and AwPKU (p=.008), but **no significant interactions** between individual group pairings and non-word length effects were found.

Analyses of RTs in non-word reading tasks found significant main effects of non-word length (F(3,312)=73.8, p<.001, ηp^2=.42) with RT increasing with increasing word length. A significant main effect of group was also found (F(2,104)=9.7, p<.001, ηp^2=.16), but no significant interaction between non-word length and group (F(6,312)=1.3, p=.25, ηp^2=.03). **A significant interaction** between non-word length and group was, however, found when YC

and AwPKU were compared individually ($F(3,180)=2.8$, $p=.04$, $\eta p^2=.04$) with AwPKU significantly more impacted by increasing non-word length than YC. The main effect of group on RT was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that YC performed significantly faster than both AwPKU ($p=.003$) and OC ($p<.001$). **No significant interactions** were found for non-word length effects between YC and OC ($F(3,219)=1.3$, $p=.28$, $\eta p^2=.02$) or between YC and AwPKU ($F(3,225)=0.5$, $p=.70$, $\eta p^2=.01$) (Fig.4.6.B).

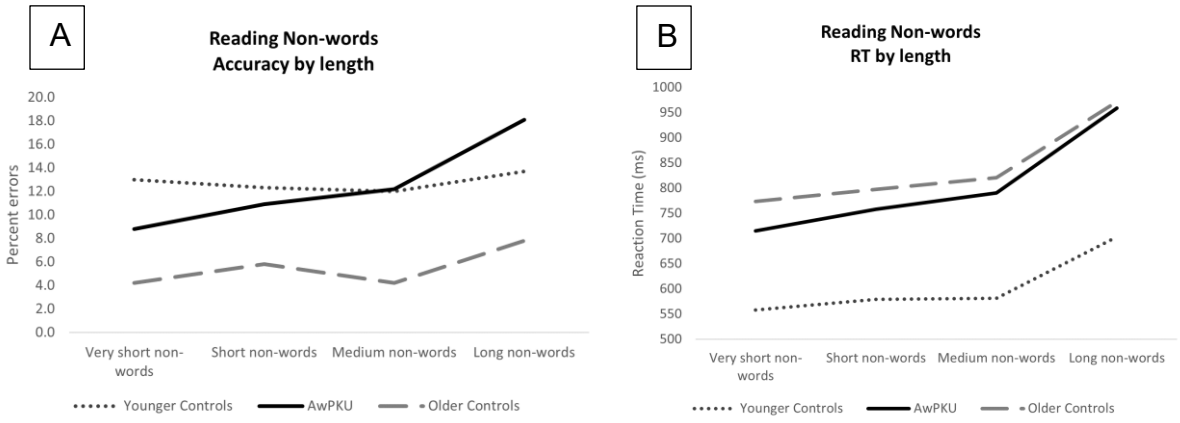


Figure 4.6. – Error rates by non-word length for YC, OC, and AwPKU (A). RTs by non-word length for YC, OC, and AwPKU (B) in word reading task.

Visuo-spatial tasks

To investigate hypothesis 2, that AwPKU and OC will both demonstrate an exponential impact of task difficulty on RTs in visual tasks, error rates and RTs across difficulty conditions in these tasks were compared. Accuracy and RT results across visual search tasks are presented in Tables 4.4. and 4.5. Significant differences in RT from YC were demonstrated across all conditions by both OC and AwPKU. No significant differences in accuracy were found between AwPKU and YC, whilst only two differences were apparent between OC and YC, one of which was in the condition requiring arguably the highest impact of concomitant executive function abilities (conjunction search with absent target). No significant differences between OC and AwPKU were found across conditions in terms accuracy. Significant differences in RT between these two groups, however, were found across all conditions.

Table 4.4. - *Comparison of z-scores across RTs in visual search tasks from OC and AwPKU. Z-scores have been reversed in some instances so that higher z-scores indicate worse performance across tasks. Z-scores of 0.6 or above in either direction have been highlighted in bold. Differences in z-scores are calculated by subtracting AwPKU from OC z-scores.*

Visual search RTs	**Younger Control (YC)**		**Older Control (OC)**		**YC vs. OC**		**Adults with PKU (AwPKU)**			**YC vs. AwPKU**	**OC vs. AwPKU**	
	Mean	**SD**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Difference in z-scores**	**p-value**
Feature search	498.4	71.7	815.1	246.8	**4.4**	<.001**	606.0	167.6	**1.5**	.002**	**2.9**	<.001**
Feature target present	506.6	74.5	808.8	246.2	**4.1**	<.001**	589.1	143.6	**1.1**	.01*	**3.0**	<.001**
Feature target absent	491.3	77.1	823.3	280.6	**4.3**	<.001**	622.6	200.4	**1.7**	.002**	**2.6**	.001**
Conjunction search	840.5	126.3	1307.6	380.1	**3.7**	<.001**	1007.9	239.7	**1.3**	.001**	**2.4**	<.001**
Conjunction target present	754.2	110.5	1126.7	319.6	**3.4**	<.001**	878.2	191.8	**1.1**	.003**	**2.3**	<.001**
Conjunction target absent	926.2	148.8	1468.3	465.5	**3.6**	<.001**	1136.1	297.0	**1.4**	.001**	**2.2**	<.001**
Feature search 4 distractors	495	82.7	788.5	231.3	**3.5**	<.001**	581.7	131.8	**1.0**	<.001**	**2.5**	<.001**
Feature search 8 distractors	496.6	68.5	823.3	262.5	**4.8**	<.001**	613.9	187.9	**1.7**	<.001**	**3.1**	<.001**
Feature search 12 distractors	504	76	835.3	277.3	**4.4**	<.001**	622.3	193.1	**1.6**	<.001**	**2.8**	<.001**
Conjunction search 4 distractors	717.4	112.7	1063	315.5	**3.1**	<.001**	838.9	195.4	**1.1**	<.001**	**2.0**	<.001**
Conjunction search 8 distractors	848.8	130.4	1306.4	365.1	**3.5**	<.001**	1009.9	240.1	**1.2**	<.001**	**2.3**	<.001**
Conjunction search 12 distractors	690.3	159.2	1513.6	496	**3.5**	<.001**	1178.2	305	**1.4**	<.001**	**2.1**	<.001**

* T-test is significant at the 0.05 level (2-tailed).

** T-test is significant at the 0.01 level (2-tailed).

Table 4.5. - *Comparison of z-scores across accuracy scores in visual search tasks from OC and AwPKU. Z-scores have been reversed in some instances so that higher z-scores indicate worse performance across tasks. Z-scores of 0.6 or above in either direction have been highlighted in bold. Differences in z-scores are calculated by subtracting AwPKU from OC z-scores.*

Visual search error rates	**Younger Control (YC)**		**Older Control (OC)**			**YC vs. OC**	**Adults with PKU (AwPKU)**			**YC vs. AwPKU**	**OC vs. AwPKU**	
	Mean	**SD**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Mean**	**SD**	**Mean z-score**	**p-value**	**Difference in z-scores**	**p-value**
Feature search	2.0	2.3	0.8	1.4	**-0.6**	.01*	1.6	2.6	-0.2	.50	-0.4	.10
Feature target present	2.9	11.1	1.8	5.3	-0.1	.60	1.6	2.9	-0.1	.53	3.0	.87
Feature target absent	1.3	2.4	1.3	5.2	0.0	.97	1.6	3.3	0.1	.67	-0.1	.72
Conjunction search	3.1	4.4	4.1	6.4	0.2	.45	2.4	2.7	-0.2	.44	0.4	.12
Conjunction target present	5.2	7.8	4.4	8.4	-0.1	.69	3.0	3.8	-0.3	.14	0.2	.34
Conjunction target absent	0.9	4.1	3.8	7.1	**0.7**	.03*	1.8	3.6	0.2	.39	0.5	.12
Feature search 4 distractors	1.9	3.6	1.5	5.5	-0.1	.69	0.5	2.1	-0.4	.07	0.3	.29
Feature search 8 distractors	2.2	4.3	1.3	5.4	-0.2	.43	2.7	4.5	0.1	.68	-0.3	.24
Feature search 12 distractors	1.9	3.6	1.7	5.6	-0.1	.82	1.6	4.5	-0.1	.75	0.0	.93
Conjunction search 4 distractors	1.4	4.4	2.7	6.2	0.3	.31	2.2	4.3	0.2	.50	0.1	.68
Conjunction search 8 distractors	4.2	6.8	4.5	7.5	0.1	.82	2.7	5.0	-0.2	.27	0.3	.08
Conjunction search 12 distractors	3.9	5.7	5.9	0.4	0.0	.85	5.9	0.3	0.0	.97	0.0	.87

* T-test is significant at the 0.5 level (2-tailed).
** T-test is significant at the 0.1 level (2-tailed).

Visual search

The impact of task difficulty in visual search tasks was investigated by assessing the effects of feature vs. conjunction search on accuracy and response times. An ANOVA was carried out using search type as the within-subject factor, group as the between-subject factor, and RT or error rate as the dependent variable.

ANOVA analyses investigating the differential impact of feature search vs. conjunction search tasks on error rates found a significant main effect of search type ($F(1,1254)=666.8$, $p<.001$, $\eta p^2 = .01$), with error rates increasing for conjunction search tasks, but no significant main effect of group ($F(2,1254)=66.5$, $p=.27$, $\eta p^2 = .002$) **nor any interaction** between group and search type ($F(2,1254)=50.3$, $p=.11$, $\eta p^2 = .004$).

Similar analyses for RTs found significant main effects of search type ($F(1,1254)=408.7$, $p<.001$, $\eta p^2 = .25$) and group ($F(2,1254)=216$, $p<.001$, $\eta p^2 = .26$) as well as a **significant interaction** between search type and group ($F(2,1254)=17.8$, $p<.001$, $\eta p^2 = .03$). Post-hoc ANOVAs with individual group pairings found significant differences between all group pairs, with significant interactions found between YC and AwPKU ($F(1,728)=24.4$, $p<.001$, $\eta p^2 = .03$) and between OC and AwPKU ($F(1,884)=31.4$, $p<.001$, $\eta p^2 = .03$) but not between YC and OC ($F(1,896)=2.8$, $p=.09$, $\eta p^2 = .003$). These analyses indicate that all participants performed slower in conjunction than feature search tasks but that this difference was particularly evident in AwPKU.

Feature search

The impact of task difficulty in feature search tasks was investigated by assessing the effects of number of distractors and target presence on accuracy and response times. ANOVAs were carried out using target presence as a within-subject factor, group as the between-subject factor, and RT or error rate as the dependent variable. MML analyses, meanwhile, were used to assess interactions between display size as a continuous variable, and group as a categorical variable on RT and error rate in displays where the target was present, and in those where it was absent.

ANOVA analyses found a significant main effect of target presence on error rates ($F(1,102)=4.33$, $p=.04$, $\eta p^2 = .04$) with participants unexpectedly showing higher error rates on average for displays where the target was present, possibly reflecting a speed-accuracy trade off in these displays. No significant effect of group was found ($F(2,102)=0.26$, $p=.77$,

ηp^2 = .005) **nor any interaction** between the two variables (F(2,102)=1.6,p=.21, ηp^2 = .03) with all groups demonstrating similar accuracy rates in both conditions.

When error rates were analysed for displays where the **target was present** only, MML analyses found no significant effect of group (F(2,157)=0.3, p=.74) or number of distractors (F(1,206)=1.5, p=.22) on performance accuracy, **nor any interaction between the two** (F(2,206, =0.5, p=.61). In displays where the **target was absent**, there was also no significant effect of group (F(2,161)=0.98, p=.38) or number of distractors (F(1,205)=0.01, p=.92) on performance accuracy, **nor any interaction between the two** (F(2,205, =1.3, p=.28) (Figs.4.7.A and 4.8.A).

ANOVA analyses of RTs found no significant effect of target presence (F(1,628)=2.1, p=.15, ηp^2=.09) on speed of response, however a significant effect of group was found (F(2,102)=27.2, p<.001, ηp^2=.35) but **no interaction between the two** (F(2,102)=1.1, p=.35, ηp^2=.02). The main effect of group on RT was further investigated through post-hoc ANOVA analyses between individual group-pairings. These comparisons found that OC performed significantly slower overall than both YC and AwPKU (p<.001 for both). **No interactions between target presence effects and individual group pairings** were found, however, demonstrating that response times in all groups were equally impacted by the target's presence in the display.

MML analyses assessing the impact of number of distractors on RTs in displays where the **target was present** found a significant main effect of group (F(2,158)=8.7, p<.001) but no effect of number of distractors on response time (F(1,207)=2.8, p=.096) **nor any interaction between the two** (F(2,207)=0.5, p=.63). The main effect of group on RT was further investigated through post-hoc MML analyses between individual group pairings. These comparisons found no significant difference in RTs between YC and AwPKU (F(1,90)=1.8, p=.19) but a significant difference between YC and OC (F(1,112)=12.6, p=.001) and between AwPKU and OC (F(1,113)=6.9, p=.01), with OC performing significantly slower than both other groups (Fig.4.7.B). **No significant interactions between group and number of distractors** were found for any individual group pairing in target present displays.

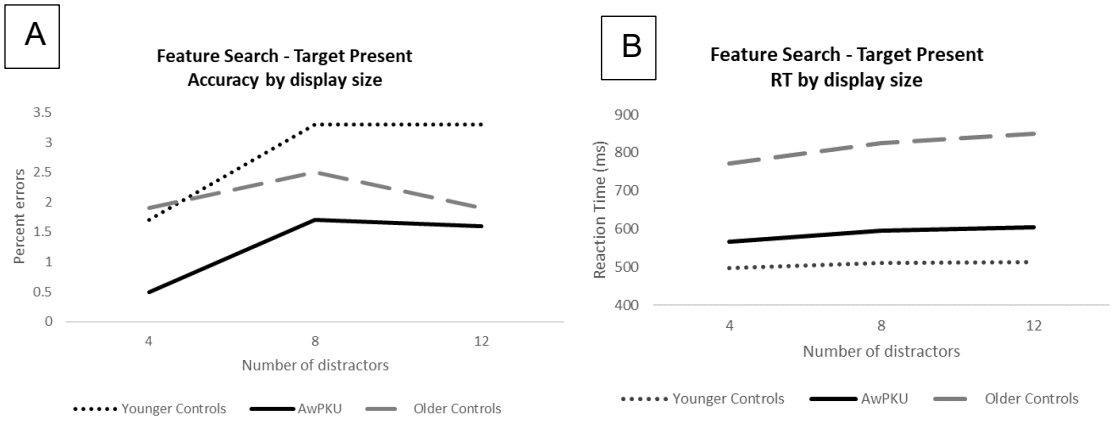


Figure 4.7. – Error rates by number of distractors when target was present for YC, OC, and AwPKU (A). RTs by number of distractors when target was present for YC, OC, and AwPKU (B) in feature search tasks.

Similarly, MML analyses for displays where the **target was absent**, found a significant main effect of group ($F(2,161)=9.4$, $p<.001$) but no effect of number of distractors on response time ($F(1,210)=0.6$, $p=.45$) **nor any interaction between the two** $(2,210)=0.1$, $p=.85$). The main effect of group on RT was further investigated through post-hoc MML analyses between individual group pairings. These comparisons found no significant difference in RTs between YC and AwPKU ($F(1,93)=2.4$, $p=.13$) but a significant difference between YC and OC ($F(1,115)=14.96$, $p<.001$) and between AwPKU and OC ($F(1,115)=6.7$, $p=.01$), with OC performing significantly slower than both other groups (Fig.4.8.B). **No significant interactions between group and number of distractors** were found for any individual group pairing in target absent displays.

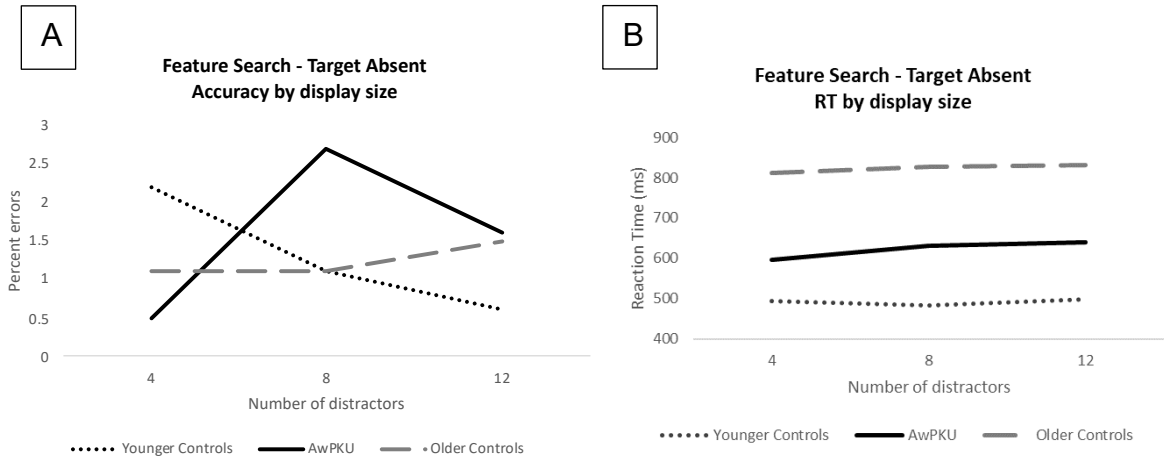


Figure 4.8. – Error rates by number of distractors when target was absent for YC, OC, and AwPKU (A). RTs by number of distractors when target was absent for YC, OC, and AwPKU (B) in feature search tasks.

Conjunction search

Similarly to feature search analyses, the impact of task difficulty in conjunction search tasks was investigated by assessing the effects of number of distractors and target presence on accuracy and response times. ANOVAs were carried out using target presence as a within-subject factor, group as the between-subject factor, and RT or error rate as the dependent variable. MML analyses, meanwhile, were used to assess interactions between display size as a continuous variable, and group as a categorical variable on RT and error rate in displays where the target was present, and in those where it was absent.

ANOVA analyses found a significant main effect of target presence ($F(1,651)=9.6$, $p=.002$, $\eta p^2 = .02$) with participants making more errors in displays where the target was absent. No significant effects of group ($F(2,102)=1.1$, $p=.34$) were found, nor any interaction ($F(2,102)=2$, $p=.14$), with all participants demonstrating similar error rates in both conditions.

In terms of RT ANOVA analyses found a significant effect of target presence ($F(1,102)=241.8$, $p<.001$, $\eta p^2=.70$) and group ($F(2,102)=23.7$, $p<.001$, $\eta p^2=.32$), as well as **an interaction between the two** ($F(2,102)=9.3$, $p<.001$, $\eta p^2=.15$). The main effect of group on RT, as well as the interaction effect, were further investigated through post-hoc ANOVA analyses between individual group pairings. These comparisons found that OC and AwPKU produced slower RTs in the target absent condition whilst YC demonstrated a slightly positive effect in the opposite direction.

MML analyses assessing the impact of number of distractors on accuracy in displays where the **target was present** found no significant effect of group ($F(2,160)=0.51$, $p=.60$), but a significant effect of number of distractors ($F(1,212)=9.8$, $p=.002$), with more errors in displays with higher numbers of distractors. **No interaction between group and number of distractors was found** ($F(2,212, =1.3$, $p=.28$). In displays where the **target was absent**, there was no significant effect of group ($F(2,151)=1.1$, $p=.35$) or number of distractors ($F(1,204)=0.02$, $p=.89$), **nor any interaction between the two** ($F(2,204, =0.02$, $p=.98$) (Figs.4.9.A and 4.10.A).

MML analyses assessing the impact of number of distractors on RT in displays where the **target was present** found a significant main effect of group ($F(2,162)=3.7$, $p=.03$) and a significant effect of number of distractors ($F(1,217)=41.1$, $p<.001$), but **no interaction between the two** ($F(2,217)=1.8$, $p=.17$), with all groups showing an equal negative impact of

increasing number of distractors. The main effect of group on RT was further investigated through post-hoc MML analyses between individual group pairings. These comparisons found no significant difference in RTs between YC and AwPKU ($F(1,94)=2$, $p=.16$) nor between AwPKU and OC ($F(1,116)=1.9$, $p=.17$) however a significant difference was present between YC and OC ($F(1,154)=28.8$, $p=.01$), with OC performing slower than YC (Fig.4.9.B). **No significant interaction between number of distractors and individual group pairings** was found in target-present displays.

RT Analyses for displays where the **target was absent**, found a trend effect of group ($F(2,146)=3$, $p=.051$) and a significant effect of number of distractors ($F(1,183)=49.9$, $p<.001$) as well as **a significant interaction between the two** ($2,183)=6.9$, $p=.001$), indicating that increasing numbers of distractors had negative effects of differing severities across groups. The main effect of group on RT, as well as the interaction effect, were further investigated through post-hoc MML analyses between individual group pairings. These comparisons found no significant difference in overall RTs between YC and AwPKU ($F(1,90)=0.18$, $p=.67$), however a **significant interaction between number of distractors and group** was found between these two cohorts ($F(1,86)=73.9$, $p=.006$), with number of distractors having a greater negative impact on AwPKU than YC. A significant main effect of group was also found between YC and OC ($F(1,100)=4.9$, $p=.03$), as well as an interaction between group and number of distractors ($F(1,122)=10.4$, $p=.002$), with OC performing significantly slower, and showing a greater negative impact of increasing numbers of distractors than YC. No significant main effect of group was found between OC and AwPKU ($F(1,116)=2.5$, $p=.12$), nor any interaction between group and number of distractors (Fig.4.10.B).

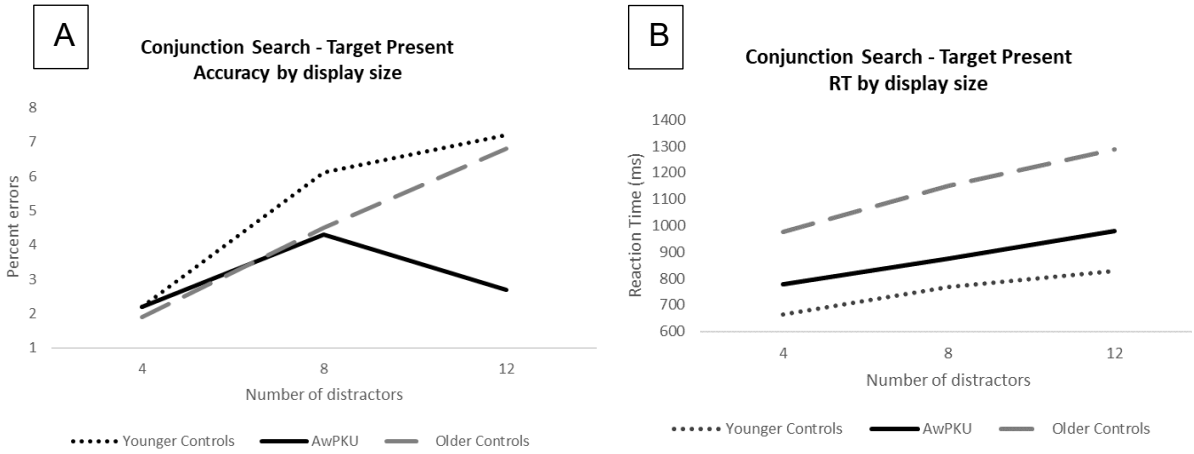


Figure 4.9. – Error rates by number of distractors when target was present for YC, OC, and AwPKU (A). RTs in by number of distractors when target was present for YC, OC, and AwPKU (B) in conjunction search task.

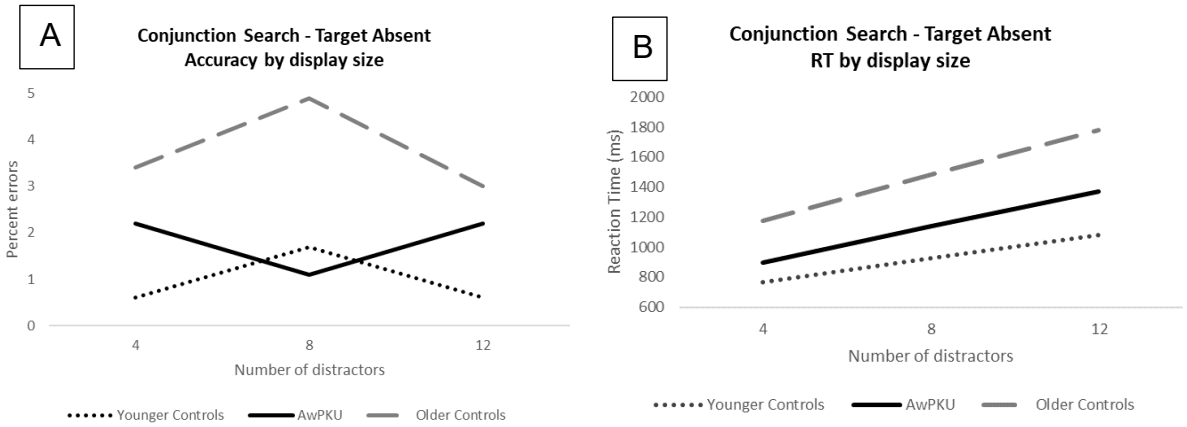


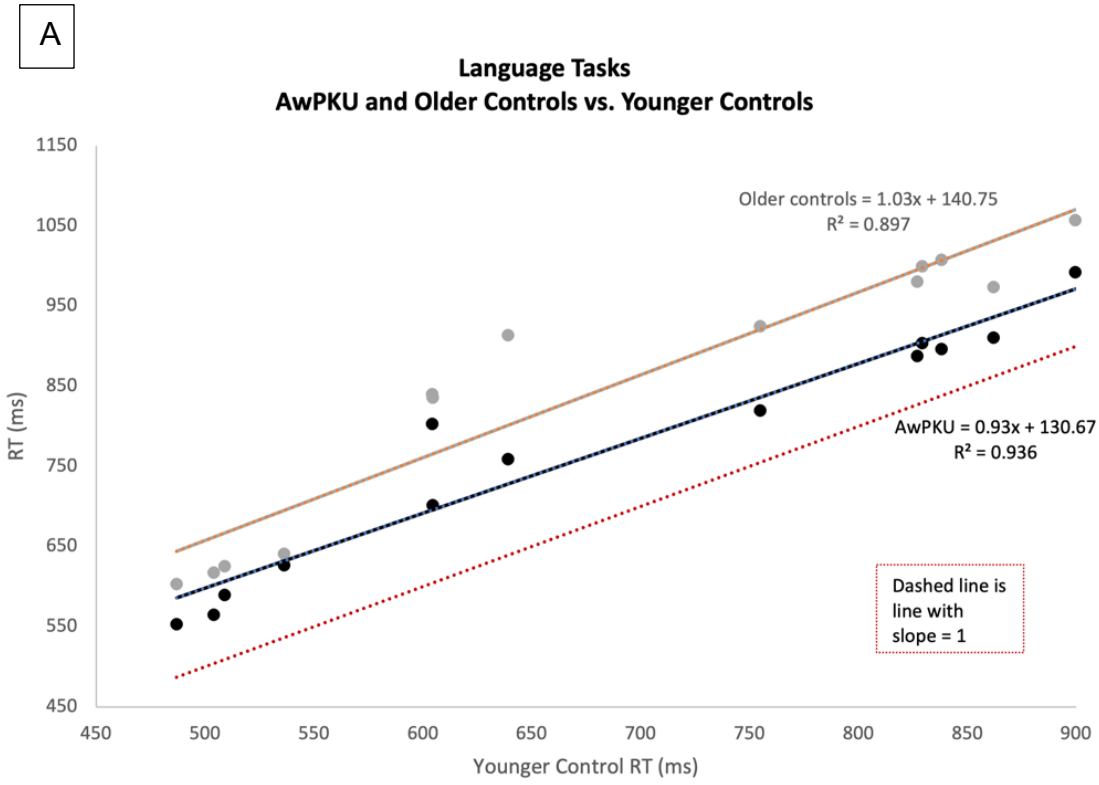
Figure 4.10. – Error rates in by number of distractors when target was absent for YC, OC, and AwPKU (A). RTs by number of distractors when target was absent for YC, OC, and AwPKU (B) in conjunction search task.

Similarly to picture naming analyses, Pearson correlational analyses were carried out between accuracy scores and RTs across groups for visual search tasks. These analyses found a significant negative correlation between errors and RTs in feature search tasks ($r(103)=-.25$, $p=.009$) suggesting that some speed-accuracy trade off was present in this condition. For conjunction search, however, there was a positive, but non-significant correlation indicating that, as reaction times increased, error rates were also increased slightly ($r(103)=.18$, $p=.07$). Increased RTs from AwPKU and OC in conjunction search tasks, therefore, are likely to be the results of slowed visuo-spatial processing in these populations, rather than solely being due to increased caution in more difficult conditions. When correlations were run for individual participant groups, correlations became insignificant for all groups, with YC and AwPKU demonstrating a non-significant negative association between RT and error rates in both feature (YC: $r(29)=-.34$, $p=.07$; AwPKU: $r(28)=-.34$, $p=.07$) and conjunction (YC: $r(29)=-.05$, $p=.78$; AwPKU: $r(28)=-.11$, $p=.56$) searches, whilst OC demonstrated a slight positive association between RT and error rates in both conditions (feature search: $r(42)=.06$, $p=.70$; conjunction search $r(42)=.23$, $p=.14$). These individual group analyses, therefore, suggest that slower responses in visual search tasks from OC, in particular, cannot reasonably be attributed to a speed-accuracy trade off.

Brinley Plots

Hypotheses 1 and 2 were further investigated through conducting Brinley plots wherein RTs of AwPKU and OC were plotted against those of YC to graphically compare the impact of increasing difficulty on RT slowing in AwPKU and OC in language and visual search tasks. The Brinley plots in Figure 4.11. show that both language and visuo-spatial RTs were linear across difficulty conditions, demonstrating a common cognitive dimension

across conditions in these tasks. However, the difference between YC RTs and both AwPKU and OC RTs increases as conditions become more difficult in visuo-spatial tasks, suggesting that the impact of task difficulty on performance in this cognitive domain is disproportionate compared to YC in both groups. Slopes plotting visuo-spatial RTs of AwPKU and OC against those demonstrated by YC accounted for 99% and 98% of the variance displayed in this domain by these groups respectively. They also indicate that RT increases with increased difficulty were 27% higher in AwPKU (y=1.27; R^2=0.99), and 65% higher in OC (y=1.65; R^2=0.98) than they were in YC, suggesting some interaction between visuo-spatial attention and task difficulty in these groups. In contrast, when we look at the plots for language tasks, RT slopes for all three populations run parallel to one another, demonstrating that, whilst there is a difference in RTs between groups across difficulty levels, this difference remains constant despite the increasing difficulty across conditions. These plots, again, explain 94% and 90% of the variance in RTs demonstrated by AwPKU and OC in this domain respectively. OC show a 3% higher increase in RT with increasing difficulty than YC across language tasks. AwPKU, however, demonstrate less of an impact of task difficulty on RT than YC, with the distance between slopes getting smaller as conditions become more difficult (AwPKU: y=0.93; R^2=0.94, OC: y=1.03; R^2=0.90). These plots suggest a fixed delay in OC and AwPKU across language tasks, regardless of task difficulty.



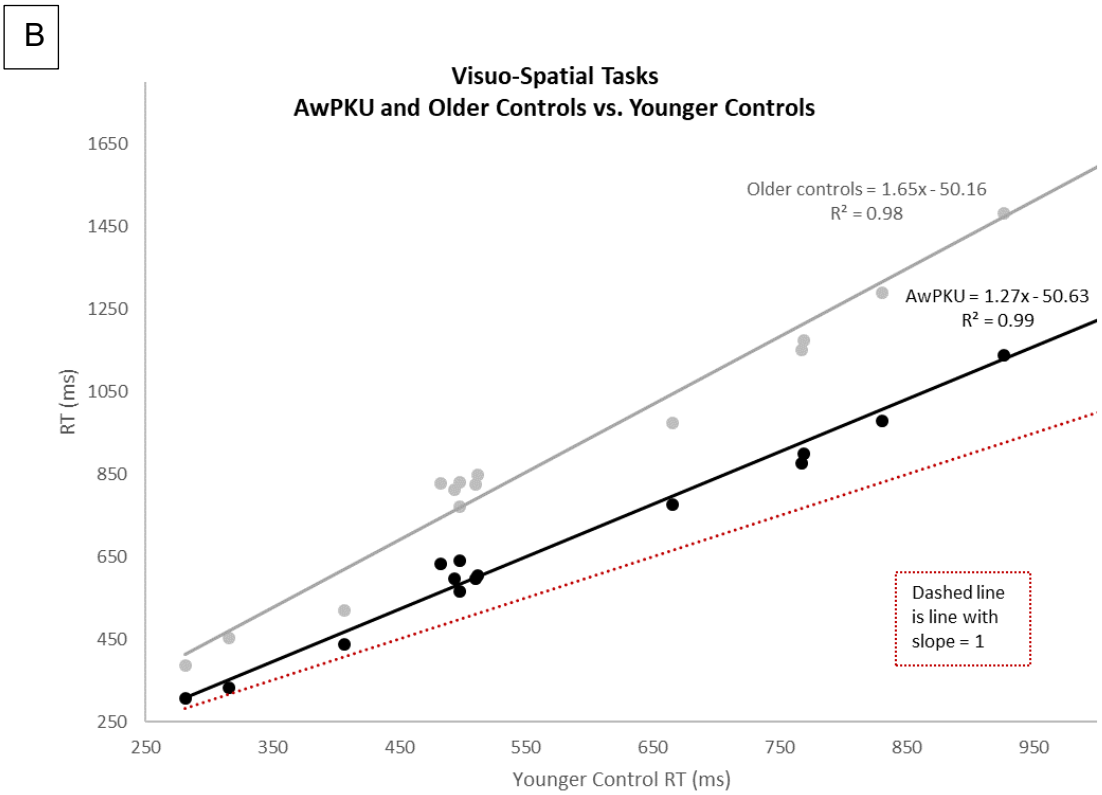


Figure 4.11. - Brinley plots for language (A) and visuo-spatial (B) tasks. RTs of AwPKU and OC plotted against the YC RTs for different conditions. Each point on the graph refers to a different condition. In both graphs, the red dotted line represents the equality line. If the RTs slope for a particular population was the same as that of controls, it would fall along the dotted line. If a population is slower dots will fall above the dotted line.

Correlation analyses

Finally, analyses were carried out to investigate hypothesis 3, that performance in executive function tasks and RTs in visual search and language tasks of increasing difficulty will be significantly correlated, reflecting a fixed impact of executive function deficits (rather than a generalised speed of processing deficit) on reaction times across domains (Table 4.6.). Performance differences between target present and target absent trials, and between small (4 distractors) and large (12 distractors) displays in visual search tasks, as well as between high and low regularity and frequency word reading tasks, long and very short word and non-word reading tasks, and high and low semantic interference picture naming conditions were correlated with performance differences known to be due to executive function demands including inhibition (incongruent vs. congruent Stroop RTs), task switching (WCST errors) and planning (Tower of Hanoi errors).

Analyses found no correlations between performance in executive function tasks and the impact of task difficulty in feature search tasks. Significant correlations were found, however, between WCST performance and the effects of task difficulty in conjunction tasks

in terms of both display size (r=.33, p=.001) and target presence (r=38, p<.001). The effects of increasing display size on RTs in conjunction search tasks were also found to correlate significantly with the effects of congruency on Stroop RTs (r=.29, p=.005), as were the effects of word and non-word length (r=.23, p=.02, and r=.46, p<.001 respectively), and word frequency, although this was in the opposite to expected direction (r=-.24, p=.04). Finally, significant correlations were also apparent between the difference between Trail making conditions A and B and the effects of word length (r=.24, p=01) and word frequency, although frequency effects were, again, in the opposite to expected direction (r=-.23, p=.02)

These analyses therefore indicate that fixed delays across domains may be due, at least in part, to executive deficits leading to a fixed increase in time to respond to tasks of increasing difficulty.

Table 4.6. – *Correlations between differences between high and low difficulty conditions in visual search task RTs and performance in tasks measuring executive function.*

	Trail Making B - A	**Stroop Incongruent - congruent**	**WCST % errors**	**Tower of Hanoi % errors**
Feature Search Target Absent - Present	r=-.09, p=.39	r=-.03, p=.76	r=-.15, p=.14	r=.02, p=.83
Feature Search Large Display - Small	r=-.07, p=.49	r=.02, p=.83	r=.04, p=.69	r=.12, p=.26
Conjunction Search Target Absent - Present	r=.02, p=.81	r=.17, p=.10	**r=.33, p=.001****	r==.19, p=.07
Conjunction Search Large Display - Small	r=.02, p=.86	**r=.29, p=.005****	**r=.38, p<.001****	r=-.14, p=.19
Reading Words High – Low Regularity	r=.16, p=.11	r=.09, p=.40	r=-.09, p=.37	r=.17, p=.12
Reading Words High – Low Frequency	**r=-.23, p=.02***	**r=-.21, p=.04***	r=-.07, p=.50	r=-.16, p=.12
Reading Words Long – Very Short	**r=.24, p=.01****	**r=.23, p=.02***	r=.07, p=.48	r=.12, p=.27
Reading Non-Words Long – Very Short	r=.12, p=.24	**r=.46, p<.001****	r=.11, p=.27	r=.08, p=.46
Picture Naming Position 5 – Position 1	r=-.06, p=.56	r=.00, p=.97	r=-.14, p=.18	r=-.05, p=.67

Discussion

This chapter built on the results of the previous chapter to further investigate similarities in cognitive impairments between AwPKU and OC, in particular with regards to the nature of potential speed impairments in these populations. Analyses of interactions between RTs and difficulty conditions in the domains of visuo-spatial processing were carried out to investigate the following three hypotheses:

1. Both AwPKU and OC will demonstrate an exponential impact of task difficulty on RTs in visual search tasks, with differences in RTs from YC getting progressively larger as task difficulty increases, demonstrating a domain-specific speed of processing deficit in this domain.

2. Both AwPKU and OC will demonstrate a fixed delay in language tasks, with the difference in RTs from YC remaining stable across conditions, demonstrating no domain-specific speed impairment in this domain.

4. We expect to find correlations between performance in executive function tasks and RTs in tasks of increasing difficulty in both visual and language tasks, reflecting the fixed impact of executive function impairments on RTs across domains.

Task difficulty in visual search

Analyses RTs in visual search tasks supported our first hypothesis. AwPKU were more impacted by the increased difficulty of feature vs. conjunction searches than YC, and that this effect was also evident, to a more severe degree, OC. Enhanced effects of target presence on performance in conjunction search tasks was apparent in both AwPKU and OC, as well as a significantly stronger negative impact of increasing number of distractors in both of these groups (although this was only apparent in displays where the target as absent). These findings are consistent with Romani et al.'s (2018) hypothesis that impairments in visual search in AwPKU are due to a reduced efficiency or capacity of visuo-attentional spotlight mechanisms, with attentional allocation moving more slowly that in healthy controls. A similar impaired mechanism appears to be apparent in OC, to a more severe extent that in AwPKU, as this group exhibited both slower overall RTs, and a stronger exponential effect of increasing task difficulty on search times.

Task difficulty in language tasks

Analyses RTs in language tasks supported our second hypothesis. Slower overall RTs were apparent in both AwPKU, and OC compared YC, with this effect more apparent in OC. No differential effects of task difficulty on response times across YC, AwPKU, or OC were apparent in picture naming and word reading tasks, indicating that the effects of semantic interference, word frequency, word regularity, and word length are the same across these cohorts. An exponential increase in RTs of OC and AwPKU compared to YC, however, was apparent when reading non-words of increasing length. This effect is likely indicative of an impairment in allocating attention to orthographic units, supporting the previous findings of Allen et al. (2011) who reported no significant interactions between age and either frequency or regularity in terms of response times between older and younger adults, but a significant effect of frequency of spatial pattern on response times from older participants. This is supported by previous findings of the role of visuo-spatial attention in non-word reading (e.g., Romani et al., 2014).

Correlations between RTs and executive function

Evidence of differential effects of task difficulty on RTs across domains suggests that the slow response times exhibited by AwPKU and older controls are not likely to result from an impairment in a generalised speed of processing mechanism. Slowed responses across domains, therefore, may be reflective of a deficit in inhibitory control, and/or increased caution in executive mechanisms, wherein individuals in these cohorts require more evidence and/or rechecking of answers than healthy, younger controls before they feel confident enough to return a response to a task. An impairment of this kind would be consistent with the unimpaired accuracy exhibited by both cohorts (with OC performing significantly more accurately than YC in a number of language tasks), as well as by the particularly strong, negative impact of target absence on RTs in these cohorts, as responses in this condition are likely to be the most uncertain, and therefore the most in need of re-checking. Correlations between the impact of increasing display size on conjunction search RTs, and increasing word and non-word length on reading RTs, with differences between congruent and incongruent Stroop RTs indicate that impaired inhibitory mechanisms may elicit some impact on performance in conditions of increased cognitive load. This is further supported by findings of significant correlations between RTs for words of increasing length and differences between response times on the Trail making A and B tasks. Findings of a negative correlation between frequency effects on RTs and inhibition effects in Trail making and Stroop tasks, however, suggest that impaired inhibitiory mechansims are not solely

responsible for the slow response times found across domains. Inhibitory mechanisms may, therefore, interact with other impaired executive mechanisms, such as an increased caution in responding, to impact speed of responses in both of these cohorts. This is supported by findings that conjunction search differences between groups were significantly correlated with increased errors on the WCST task, indicating that slowed performance in tasks of increasing difficulty could result from interactions with impairments in updating, planning, and working memory mechanisms in AwPKU and OC.

Implications

The lack of evidence supporting a generalised speed deficit in these two cohorts of participants suggests that PKU and ageing both affect speed of processing in a domain-specific manner, with some cognitive domains significantly more impacted than others. These findings support Romani et al.'s (2018) previous assertion that speed of processing is an intrinsic process, linked to the efficiency of specific cognitive domains, rather than being a general resource that is tapped equally across domains. The slower responses observed across domains in cohorts affected by PKU or ageing, therefore, are unlikely to result from the impairment of a single 'speed of processing component' or 'executive mechanism', but may instead reflect an interaction between mildly impaired inhibitory mechanims and a preference of these cohorts to respond cautiously to tasks across these domains, sacrificing response speed in an effort to avoid incorrect responses.

Findings of similar patterns of impairment in visuo-spatial and lexical processing systems in AwPKU and OC relative to YC also lend further support to the findings reported in Chapter 3, indicating that the mechanisms underlying impaired performance in both cohorts are similar and therefore may interact with one another as AwPKU get older. These findings, then, further emphasise the need for longitudinal research with AwPKU to allow monitoring of cognitive decline with ageing in this cohort, and to better prepare clinicians and AwPKU to address any potential interactions that may occur between pathologies associated with high Phe levels, and those associated with increasing age.

Chapter 5: Cognitive and Well-Being Outcomes in Middle-Aged AwPKU

Introduction

As has been discussed in previous chapters, the introduction of a low-Phe diet to treat PKU has resulted in early-treated adults with PKU (AwPKU) no longer experiencing significant intellectual disability or limited life expectancy as a result of their disorder. As the diet was only established as standard treatment for neonates with PKU in the 1950s, however, limited evidence exists regarding how early-treatment of PKU impacts adults' cognitive and well-being outcomes in later life. As the first cohort of early-treated AwPKU begin to reach middle-age, it is important that we continue to investigate the impact that PKU and ongoing metabolic control have on cognitive and well-being outcomes in later life. A few studies have begun to consider the impact of increasing age on cognition and well-being in AwPKU, although the literature in this area remains sparse, with few direct comparisons between older and younger AwPKU being carried out.

Cognition in AwPKU

Current evidence suggests that cognitive outcomes in early-treated AwPKU are below those of healthy controls (for reviews see: Burlina et al., 2019; Hofman et al., 2018). Studies investigating cognition in early-treated AwPKU have highlighted impairments in the domains of speed of processing, executive function (namely planning, switching, and monitoring), working memory, and visuospatial attention in this cohort, with some studies also finding evidence of some language deficit in terms of verbal fluency and verbal language and memory (see Aitkenhead et al., 2021; Christ et al., 2010; Hofman et al., 2018; Moyle, Fox, Arthur, et al., 2007 for reviews; also see Chapter 1 for a more detailed review of the literature). Metabolic control in adulthood has been consistently linked to performance in all of these domains, as well as being suggested to play a role in inhibition abilities in AwPKU, although previous research suggests that visuospatial processing and verbal memory and learning may be more significantly impacted by childhood metabolic control than by concurrent Phe levels in adulthood (Aitkenhead et al., 2021; Bartus et al., 2018; Bik-Multanowski et al., 2011; Brumm et al., 2004; Channon et al., 2004; Channon et al., 2007; Channon et al., 2005; Moyle, Fox, Bynevelt, et al., 2007; Pietz, 1998; Ris et al., 1994; Romani et al., 2017; Schmidt et al., 1994; Smith et al., 1996; ten Hoedt et al., 2011).

Well-being in AwPKU

In addition to findings of impaired cognition in early-treated AwPKU, a number of studies have also reported reduced quality of life (QoL) in this cohort, with measures of well-being often found to correlate significantly with metabolic control. Cazzorla et al. (2014) found significant correlations between scores on the WHOQOL and both average Phe levels over the past 12 months, and Phe on the day of testing, in 17 AwPKU aged 17-35. They further found that AwPKU had a significantly lower QoL than students with PKU, whilst long-treated participants demonstrated a better QoL than those who had stopped treatment (either dietary or pharmacological), although normal levels of depression and anxiety were apparent in all participants. Similarly, Bosch et al. (2015) investigated health-related QoL in 104 AwPKU (mean age 25.8) using the SF-36 Health and PKU QoL scales. They found lower scores in mental domains for the SF-36 compared to norms, as well as high levels of anxiety and guilt related to Phe levels and supplement intake from AwPKU. Notably, AwPKU demonstrated higher physical health scores in the SF-36 than US norms, however it is noteworthy that the mean age of adults included in the study was significantly lower than that used to calculate adult physical health norms (25.8 years vs. 50.7 years), possibly leading to this difference. In contrast, Bosch et al. (2009), found no differences between 32 AwPKU (mean age 24.6) and healthy controls in terms of course of life or health-related QoL as measured using the Course of Life Questionnaire.

With regards to the relationship between metabolic control and mood scores in AwPKU, findings have been mixed. ten Hoedt et al. (2011) reported increased depression and fatigue, and decreased vigour in nine AwPKU during the high-Phe phase of their study, along with lower overall scores on the Profile of Mood States questionnaire. Similarly, Brumm et al. (2004) found higher mean scores on the BDI and BAI in AwPKU with Phe levels >1000μmol/L compared to those with levels <1000μmol/L, although this difference was not statistically significant. Correlation analyses in this study further found significant correlations between BDI scores and age of diet initiation, and between BAI scores and median Phe aged 5.5 to 6 years, but no correlations with adulthood Phe levels. Channon et al. (2007) found no differences between on-diet AwPKU, off-diet AwPKU, or healthy controls, on the short form SF-36v2 quality of life questionnaire, the Beck Anxiety Inventory (BAI), or the Beck Depression Inventory (BDI). Aitkenhead et al. (2021) reported similar findings with 149 AwPKU, who demonstrated no difference from controls on the HADS anxiety and depression measures, although they did demonstrate a more anxious relationship style than controls. This study further found that partially-adherent AwPKU demonstrated significantly

lower mental and physical QoL scores than both on-diet and off-diet groups, suggesting that lower well-being scores in AwPKU may be more linked to lifestyle and stability of metabolic control, than to blood-Phe levels specifically.

Changes in impairments with age

Despite some variability in cohort size and testing methods across studies investigating cognitive and well-being outcomes in AwPKU, a common element amongst many of them is the young age of the AwPKU assessed, with the majority of studies in this area testing early-treated AwPKU with a mean age of mid-20s to early-30s. It is well established that cognitive abilities and neuropathological integrity decrease with increasing age in a healthy population (for cognition see Craik, 1994; Der & Deary, 2006; Salthouse, 2004 for reviews; for neuropathology e.g., Enzinger et al., 2005; Galluzzi et al., 2008; Grueter & Schulz, 2012; Hedman et al., 2012; Sigurdsson et al., 2012), with Der and Deary (2006) suggesting that slowing in simple RT tasks is evident from the age of around 50 (for a detailed review of the impact of increasing age on cognition and neurophysiology in healthy adults, see Chapter 3, pgs. 44-57). Middle- and older age, therefore, may prove to be a particularly vulnerable time for AwPKU, as age-related neural deterioration begins to further impact cognitive functions already impaired in PKU, as well as potentially affecting compensational mechanisms previously used to mitigate the impact of Phe accumulation in the brain.

The impact of increasing age on neuropathological integrity in PKU has been demonstrated by studies assessing neural changes in CwPKU as they move into adolescence and early adulthood. Wesonga et al. (2016) identified an age-related decrease in mean diffusivity in 31 CwPKU, aged 7-18. They reported a significant negative correlation between age and mean diffusivity in CwPKU that was not apparent in healthy controls, with medium to large effect sizes apparent in all regions of interest (including the prefrontal cortex, centrum semiovale, posterior parietal-occipital cortex, optic radiation, putamen, thalamus, hippocampus, and corpus callosum) therefore suggesting a developmental deterioration in brain health. From a cognitive perspective, White et al. (2002) also found a difference in the developmental trajectory of prefrontal function in PKU. This study reported significant differences in working memory for letters, abstract objects, and spatial location in 20 CwPKU aged 6 to 17 years, compared to healthy controls. Significant differences from controls, however, emerged only in older children, further suggesting some developmental deterioration or delay in cognitive abilities in CwPKU compared to controls.

Studies investigating neuropathological abnormalities in children/adolescents and adults with PKU have also reported differences between these two age groups. Mastrangelo et al. (2015) investigated associations between age and WM abnormalities in 47 participants with PKU, aged 12 to 37 years. This study reported an increased frequency and severity of WM alterations from the second decade of life onwards, with significant associations found between WM abnormalities and both age and persistent exposure to high Phe. There was, however, a significant amount of interindividual variability in the progression of WM alterations, possibly related to differences in consistency of metabolic control during childhood. No significant associations between WM alterations and IQ were reported in this study. Similar findings were reported in a longitudinal study by Nardecchia et al. (2015), which assessed IQ, neuropsychological function (including executive function, spatial planning, sustained attention, visual memory), and WM alterations, in 14 participants with PKU as they moved from childhood to adulthood across a period of 14 years (mean age at first assessment: 7.8-13.5, mean age at second assessment 22.2-27.7). This study reported the emergence of WM alterations, ranging from mild to severe, at follow up, which had not been reported during childhood. However, similar to Mastrangelo et al. (2015), neuropsychological assessments identified no changes in IQ performance with increasing age. Additionally, although impaired planning, shifting, sustained attention, and visuo-spatial abilities were still apparent at follow-up assessment, overall neuropsychological functioning was found to improve during adolescence and early adulthood, despite increased Phe levels and WM alterations.

This finding of improved cognition with age was further supported by Jahja, Van Spronsen et al. (2017) who reported an improvement in motor function in a longitudinal study investigating cognition in 21 participants with PKU during childhood (mean age 10.4) and adulthood (mean age 25.8). This study also reported stable levels of impairment in inhibition and cognitive flexibility, despite increased Phe levels with age. These findings suggest that increasing age from adolescence to adulthood has a positive impact on cognition in PKU, despite increased WM alterations. Improvements or stability in cognitive abilities, therefore, may reflect the development of compensatory cognitive strategies in AwPKU.

Due to the recent introduction of treatment for PKU from birth, few studies have investigated the impact of ageing in adulthood on cognition and brain health in early-treated AwPKU. Weglage et al., however, carried out a study in 2013 assessing IQ, attention, information processing, and WM integrity over 5 years in AwPKU age 19 to 41 years. Despite mean blood-Phe levels increasing with age, this study found that cognitive performance remained stable as participants aged, with MRI findings also indicating no

additional WM deterioration with age. Older AwPKU (i.e., those over the age of 32) were, however, found to perform slower in the Trail-making test than younger AwPKU. A similar longitudinal study by Feldmann et al. (2019) assessed IQ, attention, and information processing (measured using the WASI-IV, Tests d-2, and Trail-making tests respectively) in 35 early-treated AwPKU aged 29 to 51. Assessments were conducted at a 5-year, and then a 10-year, follow up. Although performance on all three of these measures was, again, found to remain consistent, comparisons between older AwPKU (aged >42 years), and younger AwPKU (aged <42 years) found that the older cohort demonstrated significantly slower information processing at both assessment times.

With regards to the impact of age on neurological health in AwPKU, Pilotto et al. (2019) assessed neuropathological integrity and neurotransmitter levels in 10 early-treated AwPKU aged 31 to 45 years. This study reported evidence of reduced serotonin and dopamine metabolites in AwPKU compared to controls which correlated significantly with grey matter atrophy in both frontal areas and in the precuneus. The precuneus, in particular, is a region known to be vulnerable to the negative effects of ageing, with atrophy in this area associated with mild cognitive impairment and the early phases of Alzheimer's disease in healthy adults (Frisoni et al., 2017; Nathan et al., 2017; Pievani et al., 2011; Weston et al., 2016). These impairments were also found to be associated with Phe levels in AwPKU, indicating that poor metabolic control in older age may lead to accelerated neurotransmitter depletion and neural atrophy linked with increasing age in this population.

Finally, when we consider the impact of ageing on well-being in PKU, Douglas et al. (2013) found that PKU QoL scores were inversely correlated with age in a cohort of 37 PKU participants aged 10 to 49. Comparisons between groups, however, were only split into participants <=19 or >20 years, with differences between younger and older adults remaining unexplored. Meanwhile, Bilder et al. (2013) assessed well-being in 64 early-treated AwPKU aged 17 to 48, using the Brief Symptom Inventory (BSI). Regression analyses found a significant association between age and BSI scores, with increasing age associated with worse scores, in particular on the Global Severity Index, Positive Symptom Distress Index, obsessive-compulsion, depression, and psychoticism sub-domains. Direct comparisons between older and younger AwPKU, however, were not carried out. These studies both indicate that well-being may be another area particularly vulnerable to the negative effects of ageing in AwPKU, although more research is required investigating differences between older and younger AwPKU.

In this study, we continue to investigate the impact that PKU and ongoing metabolic control have on cognitive and well-being outcomes in later life by comparing levels of impairment on a range of cognitive assessments between young and middle-aged AwPKU. We also explore associations between blood-Phe concentrations and cognitive and well-being scores in middle-aged AwPKU.

Remote testing

Although this study was originally designed to be carried out in person (using the materials presented in Chapters 3 and 4), the ongoing COVID-19 pandemic required the re-design of study materials and delivery protocols to allow cognitive assessments of participants to be carried out in remote conditions. Prior to the pandemic, remote delivery of psychological of assessments has been growing in popularity due to the opportunities and advantages that it presents, including cost-effectiveness, and reduced time expenditure for both researchers and participants (Lefever et al., 2007; Lochner et al., 2016; Tuten et al., 2002). Remote assessments also allow for more flexibility in terms of who is able to participate in studies, such as those with busy schedules, or living in different parts of the country, which can be of particular value when recruiting specialised samples (Kraut et al., 2004; Tuten et al., 2002). Previous research investigating the efficacy of remote administration of cognitive assessments has reported good correlations between results gathered remotely and in face-to-face conditions using both between-participant and within-participant designs (Abdolahi et al., 2016; Brearly et al., 2017; Chapman et al., 2021; Dauphinot et al., 2020; Lindauer et al., 2017; Randolph et al., 2014; Stillerova et al., 2016; Zeghari et al., 2021).

Studies comparing telephone delivery of cognitive assessments to older participants with mild to severe cognitive impairments have reported no significant differences in scores achieved in assessments delivered over the telephone and those delivered in face-to-face settings (Dauphinot et al., 2020; Randolph et al., 2014). Similarly, a feasibility study comparing performance on neuropsychological tests delivered face-to-face or via video call demonstrated comparable performance across the two modalities, with no significant difference in overall scores reported (Zeghari et al., 2021). In particular, significant agreement across scores between modalities were reported for the Mini Mental State Examination, letter fluency, and forwards digit span assessments. Some differences, however, were apparent between scores for semantic fluency and backwards digit span tasks. Notably, this feasibility study only included eight participants, therefore further

research is needed before clear conclusions about the efficacy of remote delivery of these neuropsychological assessments can be drawn.

A systematic review and meta-analysis of adult neurocognitive assessments delivered via video call reported consistent performance in video call and face-to-face assessment scores across 12 studies and 497 participants (Brearly et al., 2017). These studies included both healthy adults and those with psychiatric or neurocognitive disorders, with mean participant ages ranging from 34 to 88 years. In particular, verbally-mediated tasks, similar to those included in our cognitive assessment battery (i.e. digit span, verbal fluency, and list learning), were found to be unaffected by mode of delivery, however Boston Naming Scores did fall by approximately 1/10th of a standard deviation below face-to-face scores. Although no significant differences were found between scores on assessments delivered via video call or face-to-face in this review, overall video call scores were approximately 1/33rd of a standard deviation below face-to-face scores, and there was significant heterogeneity between included studies.

Although the majority of studies found good consistency between assessment delivered face-to-face and remotely, some differences have been reported in either direction. One between-participant study with children with mild to moderate intellectual disability due to William's Syndrome reported that, whilst scores on Raven's coloured Progressive Matrices were equivalent when delivered remotely vs. face-to-face, performance on the British Vocabulary Scale was better from those participants who were tested remotely (Ashworth et al., 2021). Meanwhile a study investigating online vs. face-to-face delivery of cognitive assessments to adults with multiple sclerosis found a consistent trend for higher scores during in-person administration of the automated neuropsychological assessment metrics, as well as significantly better performance in face-to-face conditions of the symbol digit modalities test (Settle et al., 2015).

Our study provided an opportunity to further assess the efficacy of remotely delivered cognitive assessments across a large range of tasks, in particular in terms of their sensitivity to mild impairments (such as those usually exhibited by AwPKU) in comparison to their face-to-face counterparts.

Aim and objectives

We assessed performance on a remotely-delivered battery of cognitive assessments, and gathered well-being and QoL information, from a cohort of middle-aged, early treated

AwPKU. Cognitive and well-being outcomes were then compared to age-matched healthy controls to assess levels of impairment. Emerging impairments were then compared and contrasted with impairments previously demonstrated by young AwPKU, with the aim of better understanding the impact of PKU in later adulthood.
This was an exploratory study, investigating the research question, “How does increasing age impact cognition and well-being in AwPKU?”

Method

Design

This study employed a between-participants, independent groups design, with two participant groups (middle-aged AwPKU, and middle-aged controls). Additional analyses were then conducted with previously collected and published data from another participant group (young AwPKU; Palermo et al., 2017; Romani et al., 2017, 2019).

Participants

Young PKU Participants

Data from 37 young adults with classical PKU reported in Romani et al. (2017, 2019) and Palermo et al. (2017) was also included in comparisons. These participants were recruited from the Department of Inherited Metabolic Disorders at the University Hospitals Birmingham. All participants were diagnosed through new-born screening 5-7 days after birth and were continuously treated with a low-Phe diet from diagnosis. Data on historical Phe levels was gathered from the Clinical Chemistry Department at Birmingham Children’s Hospital.

At time of testing, seven participants were on an unrestricted diet and 30 were following a low-Phe diet. All early-treated PKU individuals attending the clinic were invited to participate, as well as a number of individuals who were not currently attending clinic follow-up appointments but were still contactable. All individuals who responded to the invitation were tested.

Participants for this group were recruited and assessed prior to the undertaking of this PhD. Data from this cohort has previously been published in Palermo et al. (2017) and

Romani et al. (2017, 2019). It has also been previously analysed in Chapters 3 and 4 of this thesis.

Middle-aged PKU Participants

19 early-treated AwPKU, aged 40 to 56 years, with classical PKU were recruited from the Charles Dent Metabolic Unit at University College Hospitals London (UCLH; N=13) and the Mark Holland Metabolic Unit at Salford Royal NHS Foundation (SR; N=6). All participants were diagnosed through new-born screening conducted 5–7 days after birth and were continuously treated with a low-Phe diet from diagnosis. Participants from both clinics were approached by dieticians or clinicians during their scheduled clinic visits (either in person or via telephone). Additional recruitment was conducted by Salford Royal through sending letters and participant information sheets to eligible patients via post. Interested participants then contacted the research team at Aston University for more information about the study, and to provide consent to participate. Data on adulthood Phe levels was gathered from the UCLH and SR patient databases.

At time of testing, six participants were on an unrestricted diet and 13 were following a restricted diet. All early-treated PKU individuals attending the clinic were invited to participate, as well as a number of individuals who were still contactable. Forty-one individuals responded to the invitation and were contacted by the research team. Twenty-three patients who responded to the original study invitation did not go on to take part in the study due to a lack of time (N=2), a lack of access to an appropriate device to run the remote assessments (N=1), ill health (N=1), changing their mind about wanting to participate (N=4), or not responding to communications (N=15). Twenty participants were consented and tested. One participant was later excluded as they did not receive treatment until 3 months after birth.

Participants for this group were recruited and assessed for the purpose of this study. Data from this cohort has not been analysed or presented in any previous publication or chapter of this thesis.

Middle-aged Control Participants

Controls consisted of a group of 31 healthy adults, matched with PKU participants for age, gender, and education (Table 5.1.). Healthy volunteers were recruited through the Aston University volunteering website, and through adverts posted on social media and in public spaces (e.g., village halls, Aston University campus). Forty control participants in total

were recruited and assessed. Data from nine participants was excluded due to concerns regarding data integrity stemming from doubts about the participants' truthfulness in answering background questions and their commitment to provide quality data (these participants formed a group who contacted researchers at about the same time).

Participants for this group were recruited and assessed for the purpose of this study. Data from this cohort has not been analysed or presented in any previous publication or chapter of this thesis.

Table 5.1. – *Demographic information for each participant group*

		Sex		Age (yrs.)			Education (yrs.)	
Participant Type	**N**	**% Male**	**% Female**	**Mean**	**SD**	**Range**	**Mean**	**SD**
Middle-aged AwPKU	19	57.9	42.1	45.8	4.4	40-56	14.5	2.7
Middle-aged control	31	60	40	46.5	4.3	40-55	15.8	2.7
Young AwPKU	37	35	65	27.5	7.3	18-41	14.4	2.0
Younger control	30	33	67	27.6	7.4	18-41	15.2	1.7

This research was approved by the NHS and Aston University ethics committees. All participants gave voluntary informed consent to take part.

Cognitive Measures

Remote assessments were designed to reflect the face-to-face assessments previously used by Palermo et al. (2017) and Romani et al. (2017, 2019) as closely as possible. However, due to the limitations of technologies available, some differences were unavoidable. Brief descriptions of tasks used to measure performance in each domain are laid out below. For a detailed descriptions of task design, delivery, and scoring, as well as differences between face-to-face and remote versions of tasks, see the methodology chapter of this thesis (Chapter 2 – remote delivery).

IQ

WASI – The vocabulary and similarities subtests from the WASI were administered. Verbal IQ was computed.

Executive function

WCST (64-card version) - Participants were asked to correctly match 64 cards with symbols on them with corresponding place-holding cards. Sorting rules (colour, form, or number) were not disclosed to participants, and were changed following 10 correct responses. Number of correct responses was scored.

Verbal fluency – Participants were asked to name as many words as possible, either starting with a particular letter or belonging to a specific category, within 1 minute. The task included two letter fluency conditions (with letters ‘c’ and ‘p’) and two semantic fluency conditions (with categories ‘clothing’ and ‘animals’). Number of correct responses was scored.

Picture naming – Inhibitory control was measured through comparing response times to images presented later in a semantic category to those presented earlier in the same category to assess the effect of semantic interference on speed of response in the picture naming task (see language tasks).

Stroop task – Inhibitory control was measured through differences in response times to congruent and incongruent stimuli in the Stroop task (see language tasks).

Memory and learning

Digit span – Participants were asked to repeat back sequences of digits read to them. Lists started at a length of four digits. If the participant recalled more than half of the 10 sequences presented for a given length, they were presented with a longer sequence (up to a maximum of eight digits). Number of correct responses was scored.

Corsi span – Participants were presented with nine static, white squares on their screen. A sequence of squares flashed red, immediately after which participants were asked to click on squares in the same order that they flashed. Three trials were presented for each sequence length (from 2 to 9). If the participant responded correctly to at least one trial of a given length, they were presented with the next length, otherwise the task was stopped. Number of correct responses was scored.

Non-word repetition – Participants were asked to repeat sequences of non-words. Ten trials were presented for each sequence-length (2, 3, and 4 non-words). The task was stopped if more than half of the sequences for a given length were incorrect. Number of correct responses was scored.

Rey Auditory Verbal Learning Task – Participants were read a list of 15 nouns and asked to recall as many words as they could, in any order. The list was repeated five times and participants were asked to repeat what they could remember after each presentation. After a 20-minute delay, participants were asked again to recall as many items as they could remember.

Verbal paired associate learning – Participants were presented with nine pictures, each associated with a written non-word. They were then presented with the images again and asked to write down the associated non-word. The task ended when all nine words were recalled correctly, or after five presentations of the whole list. Number of correct responses was scored.

Visual attention

Choice reaction time – Participants were presented with an arrow pointing to either the left or the right of the computer screen. They were asked to respond by pressing the corresponding arrow on their keyboard as quickly as possible. Reaction time and number of correct responses was scored.

Detection with distractors – Participants were asked to respond as quickly as possible when a target (ladybird) was presented on a screen, with or without a distractor (green bug), and not to respond if only distractors (green bugs) were presented. Upon completion of 64 trials, the task was run again but required responses were reversed (i.e. participants were asked to respond for the green bug and not respond for the ladybird). Reaction time and number of correct responses was scored.

Visual search – Participants were asked to search for a target (the red ladybird) among a number of distractors (4, 8, or 12 other bugs) on their screen. They were asked to press the ‘m’ key on their keyboard if the target was present, or the ‘z’ key if it was absent (responses were reversed for left-handed participants). In a feature search condition, distractors were green bugs only. In a conjunction search condition, distractors were both green and red bugs. Reaction time and number of correct responses was scored.

Sustained attention

Rapid Visual Information Processing (RVP) – Participants were asked to detect target sequences of three digits (3-5-7, 2-4-6, or 4-6-8) in a rapidly presented string of digits between ‘1’ and ‘9’. Reaction time and number of correct responses was scored.

Language

Picture naming – Participants were asked to name 120 images, split into 24 semantic categories, presented on the screen, one at a time. They were asked to respond as quickly as possible. Reaction time and number of correct responses was scored.

Word and non-word spelling – Participants were given a list of words to spell to dictation. Thirty regular and irregular words were presented, as well as 40 non-words. Number of correct responses was scored.

Phoneme deletion – Participants were read a word (e.g. table), and then one sound from within that word (e.g. /t/). They were then asked to repeat back the word without the given sound. The task included 20 trials which resulted in real words (e.g. powder, /d/ = power) and 20 which resulted in non-words (e.g. cabbage, /k/ = abbage). Number of correct responses was scored.

Spoonerisms – Participants were read two words. They were asked to exchange the first two sounds and repeat the words back (e.g. bad, sin = sad, bin). The task included 12 trials which resulted in two real words, and 12 trials which resulted in two nonsense words. Number of correct responses was scored.

Sentence completion – Participants were read a sentence that was missing its last word. In a logical condition, they were asked to finish the sentence with the correct word as quickly as possible. In an illogical condition, they were asked to finish the sentence with an unconnected word. Reaction time and number of correct responses was scored.

Stroop task – Participants were asked to name the colour of stimuli presented on a screen. In a neutral, condition they were presented with a string of X’s. In a congruent condition, they were presented with a word in the same colour text as the word (e.g. the word ‘Green’ in green text). In an incongruent condition, they were presented with a word in a difference colour text than the word (e.g. the word ‘Blue’ in red text).

Well-being Measures

After completing the cognitive assessment, all participants were asked to complete depression and anxiety questionnaires delivered to them via a Qualtrics form. To better understand how the COVID-19 pandemic may have influenced responses to these measures, participants were also asked about how the UK-wide lockdown had affected their mood and (in the case of participants with PKU) how it had impacted their ability to manage their PKU. In addition to these questionnaires, participants with PKU were also asked to complete online questionnaires assessing their health-related quality of life. Copies of all questionnaires administered can be found in Appendix A.

Emotional well-being (delivered to all participants)

Beck Depression Inventory (BDI-II; Beck et al., 1996) – Participants were presented with a 21-item self-reporting questionnaire. All items were presented on a 4-point Likert scale ranging from 0 (symptom absent) to 3 (severe symptoms). The questionnaire assesses affective, cognitive, somatic, and vegetative symptoms over the last 2 weeks, in line with the DSM-IV criteria for major depression. Participants were presented with the question "*Please read each group of statements carefully, then fill in the circle next to the statement in each group that best describes the way you have been feeling over the past two weeks, including today*", followed by four possible statements (e.g. "0 – I do not feel sad, 1- I feel sad, 2 – I am sad all the time and I can't snap out of it, 3 – I am so sad and unhappy that I can't stand it"). Responses were scored by summing participants' responses (0-3) to all items, the minimum score being 0 and the maximum being 63. Higher scores reflect more severe symptoms, with scores above 20 indicating depression in non-clinical populations.

Beck Anxiety Inventory (BAI; Beck et al., 1988) – Participants were presented with 21-item self-reporting questionnaire. All items were presented on a 4-point Likert scale ranging from 0 (symptom absent) to 3 (severe symptoms). The questionnaire assesses anxiety symptoms over the last month. Participants were presented with the question "*Below is a list of common symptoms of anxiety. Please carefully read each item in the list. Indicate how much you have been bothered by that symptom during the past month, including today, by selecting the button in the corresponding space in the column next to each symptom.*", followed by a list of symptoms (e.g. "numbness or tingling") and a Likert scale of possible responses ("Not at all; Mildly, but it didn't bother me much; Moderately – it wasn't pleasant at times; Severely – it bothered me a lot"). Responses were scored by summing participants' responses (0-3) to all items, the minimum score being 0 and the maximum being 63. Higher

scores reflect more severe symptoms. A score of 0-21 indicates low anxiety, 22-35 indicates moderation anxiety, and 36+ indicates potentially concerning levels of anxiety.

COVID-19 questionnaire – a non-validated questionnaire was presented to participants asking them to rate how the UK-wide lockdown due to the ongoing pandemic had affected their mood. There were 7 questions asking about stress, isolation, anxiety, exhaustion, sadness, anger, and concern about their health. Questions were answered on a 7-point Likert scale with 0 indicating no change, scores +1, +2, +3 indicating progressive increases in a given emotion, and scores -1, -2 and -3 indicating progressive decreases in a given emotion (e.g., *"Please rate how you feel since the UK-wide lockdown was declared on 23rd March 2020. 1) Stressed"* – Significantly more, Moderately more, Slightly more, No more or less, Slightly less, Moderately less, Significantly less). Thus an overall score of -1 to -21 indicated better emotional health since the lockdown, a score of 0 indicated no impact of the lockdown, and a score from 1 to 21 indicated a negative impact of the lockdown. Participants with PKU were asked to answer two further questions regarding their PKU management, and access to supplements, during the lockdown (rated on a scale from significantly more difficult – significantly less difficult). Responses were scored similarly to mood scores, with a score of -1 to -3 indicating that they had found managing their PKU easier during the lockdown, 0 indicating that there had been no change, and 1 to 3 indicating that PKU management had become more difficult.

Health-related quality of life (delivered to AwPKU only)

Two health-related quality of life questionnaires were delivered to AwPKU only. The purpose of these questionnaires was to investigate how blood-Phe levels, and following a PKU diet, affected the areas of quality of life known to be impacted by living with a chronic condition. For this reason, control participants did not complete these questionnaires.

PKU Quality of Life Questionnaire (PKU-QoL; Regnault et al., 2015) – Participants were presented with a 65-item questionnaire split into four modules: PKU symptoms (e.g., *"In the past 7 days, I had headaches"* – Never, A little of the time, Sometimes, Often, Very often, *"If you had this, do you think it was related to PKU?"* – Yes, No, I don't know), PKU in general (e.g., *"In the past 7 days, it was hard to do everything I needed to do for my PKU"* – Never, A little of the time, Sometimes, Most of the time, Always), Administration of Phe-free protein supplements (e.g., *"In the past 7 days, I missed taking some supplements"* – Never, 1 or 2 times, 3 to 5 times, 6 or 7 times, More than 7 times, I don't take a supplements), and Dietary protein restriction (e.g., *"In the past 7 days, I followed my PKU diet"* – Never, A little

of the time, Sometimes, Most of the time, Always, I don't follow a PKU diet). Responses to each item of the questionnaire were scored from 0-4. Domain scores were then calculated by summing the response scores and applying a linear transformation to the sum so that all domain scores ranged from 1 to 100. Higher scores were associated with more frequent symptoms, poorer adherence, or a greater impact of PKU on quality of life. Domain scores were only calculated if 70% of the questions in a domain were answered, otherwise the domain score was set as missing.

36-Item Short Form Health Survey (SF-36; Ware Jr & Sherbourne, 1992) – Participants were presented with a 36-item questionnaire split into eight domains: physical functioning (e.g., *"During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health? Cut down the amount of times you spend on work or other activities"* – Yes, No), bodily pain (e.g., *"How much bodily pain have you had during the past 4 weeks?"* – None, Very Mild, Mild, Moderate, Severe, Very Severe), role limitation due to physical health problems (e.g., *"Does your health limit you in these activities? If so, how much? Lifting or carrying groceries"* – Yes, Limited a Lot; Yes, Limited a Little; No, Not Limited at all), role limitation due to personal or emotional problems (e.g., *"Emotional problems interfered with your normal social activities with family, friends, neighbours, or groups?"* – Not at all, Slightly, Moderately, Severe, Very Severe), emotional well-being (e.g., *"[during the last 4 weeks] Have you been a very nervous person?"* – All of the time, Most of the time, A good bit of the time, Some of the time, A little bit of the time, None of the time), social functioning (e.g., *"During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities?"* – All of the time, Most of the time, some of the time, A little bit of the time, None of the time), energy/fatigue (e.g., *"[during the last 4 weeks] Did you have a lot of energy?"* – All of the time, Most of the time, A good bit of the time, Some of the time, A little bit of the time, None of the time), and general health perceptions (e.g., *"I am as a healthy as anybody I know"* – Definitely true, Mostly true, Don't know, Mostly false, Definitely false). The questionnaire also includes a single question which provides an indication of perceived change in health (*"Compared to one year ago, how would you rate your health in general now?"* – Much better now than one year ago, somewhat better now that one year ago, About the same, Somewhat worse now than one year ago, Much worse than one year ago). Reponses to the questionnaire were scored following the RAND 36-Item Health Survey (version 1.0) guidelines (Hays et al., 1993). Numeric values of responses were recoded so that all responses were scored on a range of 0 to 100 with a higher score representing a more favourable health state. Values for all individual responses were then averaged together to create a single score representing each of the eight domains. Scores were then

reversed so that higher scores reflected a worse health state (for accurate comparison with other well-being measures).

Procedure

Remote assessment sessions were completed over two or three sessions, all of which were conducted over video call (either through Zoom or Microsoft Teams). During the first session, eight online tasks were completed via links sent through the chat function of the video call platform. The researcher explained each task to the participant in advance of sending across the relevant link. The researcher was then available to answer any queries that participants had whilst completing these tasks. Online materials were created using the PsychoPy3 Experiment Builder and hosted online by Pavlovia. Online tasks included: Simple Reaction Time, Choice Reaction Time, Detection with Distractors, Visual Search, Wisconsin Card Sorting, Rapid Visual Processing (RVP), and Corsi Span tasks.

Once all online tasks had been completed, the second session was booked for a date convenient to the participant. Where possible second (and third) sessions were scheduled for within 1 week of completing the first assessment, although timeframes of up to 3 weeks were permitted. During the second session, assessments were conducted via video link. Tasks completed during this session were: WASI Vocabulary and Similarity subtests, Digit Span, Non-word Repetition, Picture Naming, Word and Non-word Spelling, Phoneme Deletion, Spoonerisms, Sentence Completion, Stroop, Rey Auditory Verbal Learning, Verbal Non-word Paired Associate Learning, and Verbal Fluency. The second session took approximately 2 hours to complete, including a break for participants in the middle of the session. Participants were also given the option to split the second session into two, shorter, sessions (to be completed within a week of one another) if this was more convenient for their schedule. Where possible, tasks were administered in a fixed order to all participants, however some flexibility on task order was required dependent on participants' schedules and engagement levels (e.g., if time restraints meant that the next task in the assessment order would take too long to complete in the current session, this may be swapped out for a shorter task, and the longer task then completed in the second session).

Where possible, all participants were assessed on all measures. Technical difficulties when first moving tasks online for remote delivery resulted in RT data being unavailable for 12 controls and seven AwPKU for the Stroop and picture naming tasks. Four control participants only attended the first assessment session and did not return for their second

assessment session leading to further missing data. Administration errors and technical difficulties also resulted in some missing data across tasks (see Table 5.2.).

Data handling and scoring

Outcomes on all measures were scored and input into one database. Z-scores were calculated for each group ([individual score] – [average of control group]/[control group SD]), allowing direct comparisons to be made of levels of impairment across different tasks in different cognitive domains and across different testing modalities. Due to this study's interest in group, rather than individual, z-scores, as well as the particular focus on differences in z-scores between groups, a relatively low threshold for identifying mild impairments was applied. As such, for this study, z>-0.5 = no or minimal impairments, z=-0.5 to -2 = mild to moderate cognitive impairments, and z≤-2 = severe cognitive impairment. Similar z-score impairment categories have previously been applied in studies investigating mild cognitive impairments in clinical cohorts (Bohnen et al., 2015; Rowe et al., 2013). Where necessary, z-scores were reversed to ensure that lower z-scores always reflected poorer performance. Individual participant reaction times (RTs) were cleaned to exclude error responses and outliers (responses <100ms or ±3SD from participant mean). Due to the low participant numbers, and the nature of the varied impact of PKU on cognitive performance, outliers within participant groups were generally not excluded. One control participant, however, was removed from detection with distractors attentional switch RT analyses as an outlier, as their RT was >10SD away from the mean. Both composite and individual scores demonstrated high levels of skewness and kurtosis, therefore Kolmogorov-Smirnov tests of normalcy were run for all measures, and those which violated the assumption of normalcy were assessed using both parametric (t-test) and non-parametric (Mann-Whitney U) comparisons between groups. As Mann-Whitney U tests did not impact significance in any of these instances compared to t-tests, all reported results are for t-test comparisons between groups.

Composite scores for speed of response, accuracy, executive function, and well-being were calculated for comparison between participants groups. Composite speed of processing scores were calculated by averaging z-scores across the simple detection, choice RT, detection with distractors, feature search, conjunction search, RVP, Stroop, and picture naming tasks. Composite accuracy scores were calculated by averaging mean z-scores for error rates across choice RT, detection with distractors, feature search, conjunction search, RVP, Stroop, and picture naming tasks. Composite executive function

scores were calculated by averaging mean z-scores for WCST error rate; sentence completion differences in error rates and RTs between illogical minus logical conditions; Stroop differences in error rates and RTs between incongruent minus congruent conditions; detection with distractors differences in error rates and RTs between the condition where the green bug was the target and the conditions where the ladybird was switched as the target; picture naming differences in error rates and RTs between items presented 4th and 5th within a semantic category and items presented1st and 2nd. Composite well-being scores were calculated by averaging z-scores for BAI and BDI responses.

Statistical analysis

Exploratory analyses were carried out with t-test comparisons between middle-aged AwPKU and both middle-aged controls and young AwPKU for individual tasks, however it must be noted that the large number of individual assessments puts these analyses at risk of a Type I error, where seemingly significant differences may emerge due to change. We have reported significance with a standard p-value, without applying corrections. Cohen's D effect sizes have also been computed, to contextualise the strength of results. Significant differences in individual tests may warrant follow up in further research. However, it is important to stress that individual significant results are of limited value. Our primary interest lies in whether patterns of impairment can be identified, where impairments in one task are corroborated by impairments in similar tasks. For this reason, composite z-scores have been computed so that scores across heterogenous tasks in different cognitive domains can be directly compared using a single value representing 'level of impairment'. T-test comparisons were therefore conducted for composite scores between middle-aged adults and both middle-aged controls and young AwPKU. For measures where Levene's Test for Equality of Variance was found to be significant, t-test comparisons were carried out under conditions where equal variances were not assumed.

Table 5.2. – *Number of and reason for missing data points for each participant cohort*

Reason for Missing Data	Control	AwPKU
Participant attrition		
Did not return for second session	4	0
Did not complete all questionnaires	4	4
Data collection error		
Stroop	1	0
Nonword repetition	1	0
Technological issues		
Detection with distractors	4	0
Visual search	1	0
Stroop RT	11	5
Stroop accuracy	3	3
RVP	3	0
Picture naming RT	11	5
Picture naming accuracy	2	3
Did not understand/complete task		
Phoneme deletion	0	1
Spoonerisms	1	1
Word spelling	0	1
RVP	0	1
Picture non-word paired associates	0	1
Visual search	1	0

Results

Middle-Aged AwPKU vs. Controls

Middle-aged AwPKU produced an overall cognition z-score of -0.04 relative to age-matched controls, indicating no overall cognitive impairment in this cohort. When comparisons between middle-aged AwPKU and age-matched controls were made across individual cognitive tasks, AwPKU demonstrated mildly impaired performance (defined as z-scores <-0.5) in measures of speed of processing, orthographic language, and short-term memory. See Table 5.4.

Mildly impaired z-scores for RVP, visual search, and picture naming RTs demonstrate some slowed response times across domains in middle-aged AwPKU compared to age-matched controls, although t-test comparisons were only significant for picture naming RTs for items presented fifth (last) within a semantic category ($t(27)=-2.2$, $p=.04$) and conjunction search RTs ($t(46)=-2.7$, $p=.01$; see Table 5.4). Further analyses into individual task conditions found that, similar to patterns previously observed in young AwPKU, differences in RT were compounded by the effects of additional cognitive load. A measure of semantic interference (difference in RT between the final two items and the first two items presented within a semantic category) demonstrated a mildly negative z-score ($z=-0.5$), whilst display size impacted RTs in conjunction search tasks, with significant

differences found between RTs for displays with four distractors (t(46)=-1.97, p=.005), eight distractors (t(46)=-2.52, p=.015), and 12 distractors (t(46)=-2.43, p=.019), although the difference between groups decreased slightly as display sizes increased (z=-0.9, z=-0.8, z=-0.7 respectively).

In addition to differences in RT, negative z-scores in AwPKU were apparent in word spelling errors and Corsi span accuracy (Table 5.4.). Further investigation of spelling differences found that differences in spelling ability were non-significant and were only apparent for irregular words (z=-0.8) with error analyses finding a higher number of insertion and substitution errors, in particular, from AwPKU compared to controls. A significant proportion (27.5%) of all word spelling errors demonstrated by AwPKU, however, were made by a single participant (LS120). When this participant was removed from analyses, z-scores reduced significantly to -0.3 for both overall word spelling errors, and irregular word spelling errors. Assessment of Corsi span scores found that AwPKU showed impaired spatial short-term memory in comparison to controls, with a significantly lower average Corsi span of 5.3 compared to 5.9 in controls (t(48)=2.1, p=.045).

Composite scores

Comparisons of composite scores between middle-aged AwPKU and controls are shown in Table 5.3. Between-subjects t-tests were carried out between z-scores for each composite measure. Middle-aged AwPKU were significantly mildly to moderately impaired in overall measures of RT (z=-0.5, t(48)=-3.1, p=.003) and well-being (z=-1.7, t(19)=-2.8, p=.01) but not in accuracy and executive function.

Table 5.3. - *Composite z-scores and SDs for middle-aged AwPKU compared to control participants. Z-scores have been normalised so that lower z-scores indicate worse performance across tasks. Z-scores of 0.5 or above in either direction have been highlighted. T-test analyses were carried out between z-scores for each composite score.*

Composite Scores	**Middle-Aged AwPKU**			**p value of t-test**
	N	**z-score**	**SD**	
Speed of processing	19	-0.5	0.8	**.03***
Accuracy	19	0.0	0.4	.95
Executive function	19	0.1	0.3	.34
Well-being	16	-1.7	2.2	**.002****

* T-test is significant at the 0.05 level (2-tailed)
** T-test is significant at the 0.01 level (2-tailed).

Table 5.4. - *Comparison of scores on all tasks across cognitive domains. Z-scores reflect differences from age-matched controls. They have been reversed for some tasks so that lower z-scores always indicate worse performance. Z-scores of 0.5 or less have been highlighted in green. Z-scores of -0.5 or more have been highlighted in red. Effect sizes of 0.8 or above have been highlighted in bold text. AwPKU = Adults with PKU. Err = errors*

Cognitive task	Middle- Aged AwPKU				Middle-Aged Controls			Middle-Aged AwPKU vs. Controls t-test		Young AwPKU (z-score relative to young controls)				Young Controls			Young AwPKU vs. Controls t-test	
	N	Mean	SD	z-score	N	Mean	SD	p-value	Cohen's D	N	Mean	SD	z-score	N	Mean	SD	p-value	Cohen's D
IQ																		
Verbal IQ	19	111.4	11.5	0.1	27	110.1	11.7	.73	0.2	37	102.3	12.9	-1.0	30	112.2	10.2	**<.001****	**1.2**
Vocabulary subtest	19	67.1	8.1	0.3	27	65.1	7.1	.37	0.4	37	58.4	8.7	-0.8	30	64.2	7.1	**.004****	**1.0**
Similarities subtest	19	39.5	4.2	0.2	27	38.3	5.3	.44	0.4	37	36	5.8	-0.9	30	39.2	3.6	**.008****	**0.9**
Spelling accuracy																		
Word spelling % err.	18	7.4	9	-0.7	27	3.7	6.4	.11	0.7	31	4	4.6	0.2	30	4.8	5.7	.52	0.2
Non-word spelling % err.	19	16.2	12.2	0.5	27	22.4	13	.11	0.7	36	13.5	8.2	-0.2	30	11.9	7.6	.43	0.3
Language/lexical access																		
Verbal fluency (letters)	19	16.2	4.9	0.5	27	14.3	4.1	.16	0.6	37	36	11.4	-0.4	30	41.8	12.9	.06	0.7
Stroop congruent % err.	16	0.1	0.3	-0.1	24	0	0.2	.77	0.6	30	0	0	0.2	28	0.0	0.2	.33	0.0
Stroop incongruent % err.	16	0.3	0.5	0.2	24	0.6	1.5	.45	0.4	30	0.7	1.1	0.1	28	0.8	1.0	.76	0.1
Stroop neutral % err.	16	0.1	0.3	-0.1	24	0.1	0.3	.68	0.0	30	0.1	0.4	-0.1	28	0.1	0.3	.76	0.0
Logical Sentence completion % err	19	0.5	2.3	0.3	27	3.3	8.3	.16	0.7	-	-	-	-	-	-	-	-	-
Picture naming % err.	16	6.9	3.2	0.2	25	8.2	6.8	.47	0.3	30	7.4	3.8	-0.2	30	6.5	4.0	.45	0.3
Phoneme deletion % err.	18	8.5	8.1	0.3	27	11	8.8	.33	0.4	36	14.4	13	-0.4	30	11.2	8.7	.24	0.4
Spoonerism % err.	18	6.3	6	0.1	26	6.8	6.7	.78	0.1	31	10.8	14.5	-0.6	30	6.8	6.6	.18	0.5
Visual search accuracy																		
Choice RT % err.	19	2.1	2.2	-0.3	31	1.5	2.1	.37	0.4	37	0.5	0.8	-0.1	30	0.4	0.8	.68	0.2
Detection with distractors (Ladybird) % err.	19	4.9	5.8	-0.2	28	3.7	6.4	.50	0.3	31	1.4	2.1	-0.3	30	1.0	0.9	.45	0.4
Attentional switch (Bug) % err.	19	4.9	8.8	0.1	28	5.7	11.8	.78	0.1	31	0.3	0.6	-0.1	30	0.2	0.5	.76	0.3

Detection with distractors overall % err.	19	9.8	14	0.0	27	9.3	16.7	.91	0.0	31	0.8	1.1	-0.3	30	0.6	0.6	.44	0.3
Feature search overall % err.	19	3.1	2.9	0.0	29	3.3	5.1	.89	0.1	31	1.6	2.6	0.2	30	2.0	2.3	.5	0.2
Conjunction search overall % err.	19	3.5	3.3	0.2	29	6.5	13.2	.34	0.4	31	2.4	2.7	0.2	30	3.1	4.4	.44	0.3
Short Term Memory																		
Digit span	19	6.3	1.5	0.0	27	6.3	1.3	.93	0.0	37	6.1	1	-0.5	30	6.5	0.9	.05	0.6
Non-word repetition % err.	19	49.5	12.7	0.1	25	51.3	21.2	.74	0.1	31	48.2	12.5	-0.9	30	39.3	10.4	**<.001****	**1.1**
Corsi span	19	5.3	1.2	-0.7	31	5.9	1	**.045***	**0.8**	37	5.3	0.9	-0.4	30	5.7	0.9	.1	0.6
Inhibition																		
Stroop incongruent-congruent % err.	16	0.3	0.9	0.3	24	0.6	1.3	.38	0.4	30	0.7	1.1	0.0	30	0.7	1.0	.87	0.0
Stroop incongruent-congruent RT	12	167.8	106.2	-0.2	15	146	96.3	.58	0.3	30	117.8	80.1	-0.5	30	92.4	51.6	.15	0.5
Illogical Sentence completion % err.	19	17.9	18.1	0.2	27	23	21.1	.40	0.4	-	-	-	-	-	-	-	-	-
Illogical-logical sentence completion err.	19	17.4	18.8	0.1	27	19.6	22.3	.72	0.2	-	-	-	-	-	-	-	-	-
Illogical-logical sentence completion RT	19	1600	1800	0.0	27	1600	1200	.96	0.0	-	-	-	-	-	-	-	-	-
Detection with distractors bug-ladybird % err.	19	-0.1	5.2	0.3	27	2.7	9.2	.25	0.5	31	-1.1	2.2	-0.3	30	-0.8	0.9	.52	0.3
Detection with distractors bug-ladybird RT	19	-23	44.6	0.3	26	11.2	122.8	.22	0.5	31	13.8	33	-0.2	30	4.2	38.4	.3	0.4
Executive Function																		
WCST Total err.	19	16.4	7.6	0.2	31	18.4	12.8	.56	0.3	37	14	8.2	-0.7	30	10.8	4.9	.05	0.7
WCST Score	19	47.6	7.6	-0.1	31	46.3	13.2	.70	0.2	37	50	8.2	-0.7	30	53.2	4.9	.05	0.7
Verbal fluency (semantic)	19	24.2	4.3	0.1	37	23.7	7.2	.80	0.1	37	21.4	6.2	-0.7	30	25.0	5.0	**.01***	**0.9**
Sustained attention																		
RVP % err.	18	17.6	17.1	-0.4	28	11.8	14.9	.23	0.5	37	18.9	11.3	-0.6	30	13.2	8.8	**.02***	**0.8**
RVP RT	18	534.4	248.8	-0.7	27	445.6	121.4	.11	0.6	37	431.4	108.5	-0.5	30	379.1	105.5	.05	0.7
LTM & Learning																		
Picture non-word paired associates % err.	18	38.5	24.3	-0.1	27	35.4	23.9	.67	0.2	31	43.5	22.9	0.1	30	47.0	24.9	.58	0.2

Picture non-word paired associates delayed recall % err.	18	17.9	28	0.2	27	23.5	22.5	.47	0.3	31	20.4	23.7	0.4	30	29.3	24.5	.16	0.5
Rey A learning % err.	19	27.6	15.1	-0.4	27	23.4	11.1	.28	0.4	31	24.1	11.2	-0.4	30	20.5	8.4	.17	0.5
Rey A delayed recall % err.	19	11.6	2.5	0.0	27	11.7	2.7	.96	0.1	31	16.3	15.9	-0.1	30	15.1	15.1	.76	0.1
Language RTs																		
Stroop neutral RT	13	681.7	89.7	-0.3	15	643.6	112.3	.34	0.5	30	701.6	129.6	-1.2	28	604.5	92.0	**<.001****	**1.2**
Stroop congruent RT	13	704.4	82.7	0	15	708.8	126.6	.92	0.1	30	758.9	151.7	-1.1	28	639.2	111.0	**<.001****	**1.3**
Stroop incongruent RT	12	867.7	103.7	-0.1	15	854.8	146.3	.80	0.1	30	876.7	206.1	-1.1	28	731.6	134.9	**<.001****	**1.2**
Logical Sentence completion RT	19	1500	300	0.0	27	1500	400	.94	0.0	-	-	-	-	-	-	-	-	-
Illogical Sentence completion RT	19	3000	1900	0.0	27	3100	1300	.98	0.1	-	-	-	-	-	-	-	-	-
Picture naming RT	14	1005.6	119.8	-0.5	16	920.8	184.5	.15	**0.8**	30	890	130.8	-0.5	29	825.7	125.3	**.02***	0.7
Visual search RT																		
Simple detection RT	19	376.2	69.6	-0.2	31	366.1	66.5	.61	0.2	31	331.7	52.9	-0.3	30	315.9	57.0	.27	0.4
Choice reaction RT	19	371	51.5	-0.1	31	367.5	58.5	.83	0.1	37	306.9	42	-0.8	30	281.3	31.3	**.01***	**1.0**
Detection with distractors (Ladybird) RT	19	512.2	81.1	-0.2	28	499.2	72.7	.57	0.2	31	431.2	68.4	-0.3	30	405.1	83.6	.19	0.5
Attentional switch (Bug) RT	19	489.3	98.1	0.1	27	503.2	131.1	.37	0.2	31	445	76.3	-0.6	30	409.3	64.6	.05	0.7
Detection with distractors overall RT	19	500.7	87.2	0.0	26	501.7	88.3	.49	0.0	31	438.1	70.6	-0.4	30	407.2	72.2	.10	0.6
Feature search RT	19	699.3	159.6	-0.6	29	629	109.1	.08	0.7	31	606	167.6	-1.5	30	498.4	71.7	**<.001****	**1.2**
Conjunction search RT	19	1000.4	162.9	-0.8	29	869.4	168.5	**.01***	**1.1**	31	1007.9	239.7	-1.3	30	840.5	126.3	**<.001****	**1.2**
Well-being																		
BAI	16	12.2	10.4	-1.4	25	4.8	5.4	**.002***	**1.3**	-	-	-	-	-	-	-	-	-
BDI	16	10.4	8.4	-2.0	25	4.4	3	**.005***	**1.3**	-	-	-	-	-	-	-	-	-
COVID-19 mood	16	6.8	5.1	-0.9	24	1.1	6.1	**.003****	**1.4**	-	-	-	-	-	-	-	-	-
Overall																		
Overall Cognition	-	-	-	-0.04	-	-	-	-		-	-	-	-0.4	-	-	-	-	-
Overall Well-being	-	-	-	-1.4	-	-	-	-		-	-	-	-	-	-	-	-	-

* T-test is significant at the 0.05 level (2-tailed), ** T-test is significant at the 0.01 level (2-tailed).

Well-being

When scores on measures of well-being between middle-aged AwPKU and controls are considered, a clear difference is apparent, with AwPKU demonstrating significantly poorer well-being scores on all measures (Table 5.4.). AwPKU exhibited increased anxiety and depression, as measured by the BAI (z=-1.4; t(39), p=.002) and BDI (z=-2.0; t(39), p=.005), as well as an increased negative impact of the COVID-19 lockdown on their general mood (z=-0.9; t(39), p=.003). AwPKU reported a negative impact of the lockdown on all mood dimensions, whilst controls indicated a slight improvement in anger and health worries (Table 5.5.). Differences between the groups were significant across the dimensions of anxiety, exhaustion, stress, and worries about their health. AwPKU further reported a slight negative impact of the lockdown on their general PKU management, and on their access to their PKU supplements.

Table 5.5. - *Comparison of responses to COVID-19 questionnaire between middle-aged AwPKU and control participants. Negative scores represent an improvement since the UK lockdown, positive scores represent a decline. An average score of zero represents no impact of the lockdown.*

COVID-19 questionnaire	**Middle-Aged AwPKU**			**Middle-Aged Controls**			**Middle-Aged AwPKU vs. Controls t-test**
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**	
Anger	0.6	1.1	-2 – 3	-0.1	1.2	-3 – 2	.06
Anxiety	1.1	1.2	-1 – 3	0.2	1.0	-2 – 2	**.02***
Exhaustion	1.1	1.1	-1 – 3	0.0	1.4	-3 – 3	**.01****
Isolation	0.9	1.0	0 – 3	0.6	1.1	-2 – 3	.43
Sadness	0.9	1.0	-1 – 3	0.3	1.1	-3 – 2	.053
Stress	1.3	1.2	0 – 3	0.2	1.5	-3 – 3	**.02****
Worried for health	0.9	1.1	-1 – 3	-0.1	1.1	-3 – 2	**.008****
PKU - general management	0.1	1.2	-3 – 3	-	-	-	-
PKU - access to supplements	0.3	0.6	0 – 2	-	-	-	-

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Middle-Aged AwPKU vs. Young AwPKU

When z-scores were compared across individual tasks, middle-aged AwPKU demonstrated notably fewer cognitive impairments than young AwPKU (Table 5.4.). Middle-aged AwPKU produced an overall cognition z-score of -0.04 which was significantly higher than the overall score of -0.4 produced by YAwPKU (t(94)=5.2, p<.001), indicating that the middle-aged cohort were significantly less impaired overall than their younger counterparts. When overall pattern of impairment between these two cohorts was considered, a small correlation was observed, however this was not significant (r=.24, p=.11).

Young AwPKU were mildly impaired across executive function tasks whereas middle-aged AwPKU were not. Although fewer language tasks were carried out in remote assessments than in previous face-to-face batteries, RTs on tasks that were assessed in

both cohorts were notably impaired in young AwPKU but were comparable to controls in middle-aged AwPKU (aside from a mildly impaired z-score of -0.5 in picture naming RTs). Furthermore, whilst young AwPKU exhibited normal performance in letter fluency and non-word spelling, middle-aged AwPKU notably outperformed matched controls in these tasks. Conversely, young AwPKU exhibited slightly better performance than matched controls in word spelling accuracy whilst middle-aged AwPKU exhibited a mild impairment on this task. As discussed earlier, however, impaired spelling performance in this cohort was largely attributable to a single participant's poor performance on this task.

Composite scores

Magnitude of impairment in middle-aged AwPKU was further compared to that of young AwPKU by considering the difference in composite scores between these two cohorts. Statistical comparisons of true scores between these two cohorts were not conducted as these may reflect differences in testing modalities (young AwPKU: face-to-face, middle-aged AwPKU: online) and normal age-related differences, rather than specific impairments due to interactions between PKU and increasing age. As z-scores represent levels of impairment compared to age-matched controls (tested using the same materials as clinical cohorts), t-test comparisons used these values (Table 5.6.). Young AwPKU demonstrated equal or larger impairments than middle-aged AwPKU across all composite scores, although a significant difference was only apparent for executive function where young AwPKU showed a slight impairment whilst middle-aged AwPKU demonstrated no impairment relative to controls (middle-aged AwPKU: z=0.1, young AwPKU: z=-0.3; t(52)=2.7. p=.008).

Table 5.6. - *Comparison of composite z-scores between middle-aged AwPKU and young AwPKU. Z-scores have been normalised so that lower z-scores indicate worse performance across tasks. Z-scores of 0.5 or above in either direction have been highlighted. T-test analyses were carried out between z-scores for each measure.*

Composite Scores	Middle-Aged AwPKU			Young AwPKU			Diff in z-scores	p value of t-test
	N	z-score	SD	N	z-score	SD		
Speed of processing	19	-0.5	0.8	32	-0.9	1.1	0.4	.18
Accuracy	19	0.0	0.4	32	-0.3	0.7	-0.3	.26
Executive function	19	0.1	0.3	30	-0.3	0.8	0.4	**.008***

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Middle-Aged AwPKU; High vs. Low Phe.

To assess the relationship between metabolic control and performance, AwPKU were split into high and low-Phe groups where 'high-Phe' was defined as blood-Phe levels

>1000µmol/L and 'low-Phe' was defined as blood-Phe levels <900µmol/L. AwPKU were split into groups based on concurrent Phe levels (high-concurrent Phe: N=7, low-concurrent Phe: N=9), and on mean Phe levels over the past 20 years (labelled as mean adulthood Phe; high-adulthood Phe: N=9, low-adulthood: Phe N=8).

Composite z-scores for speed of processing, accuracy, executive function, and emotional well-being were compared between middle-aged AwPKU with high- vs. low-concurrent Phe (Table 5.7.) and with high- vs. low-adulthood Phe (Table 5.8.). All comparisons between high- and low- concurrent and adulthood Phe groups were insignificant. Correlations with concurrent Phe and adulthood Phe were also non-significant across all composite scores (Table 5.9.).

No significant differences between high- and low-Phe groups were found in any dimension of the PKU QoL or SF-36 questionnaires. A significant difference between high- and low-adulthood Phe differences was apparent, however, on reports of the impact of the COVID-19 lockdown on feelings of exhaustion ($t(13)=2.3$, $p=.04$), with the low-adulthood Phe group reporting a larger increase in exhaustion since the start of the lockdown than the high-adulthood Phe group. No other significant differences were found across responses regarding the impact of the lockdown on mood or PKU management.

Table 5.7. - *Comparison of composite z-scores between middle-aged AwPKU with low- vs. high-concurrent Phe. T-test analyses were carried out between z-scores for each measure.*

Composite Scores	Low Concurrent Phe (<900umol/L)			High Concurrent Phe (>100umol/L)			Diff in z-scores	p value of t-test
	N	z-score	SD	N	z-score	SD		
Speed of processing	9	-0.5	0.8	7	-0.3	0.8	-0.2	.71
Accuracy	9	0.2	0.3	7	-0.3	0.5	0.5	.18
Executive function	9	0.1	0.3	7	0.2	0.4	-0.1	.42
Well-being	9	-1.8	2.9	5	-1.9	1.3	0.1	.90

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Table 5.8. - *Comparison of composite z-scores between middle-aged AwPKU with low- vs. high-adulthood Phe. T-test analyses were carried out between z-scores for each measure.*

Composite Scores	Low Adulthood Phe (<900umol/L)			High Adulthood (>100umol/L)			Diff in z-scores	p value of t-test
	N	z-score	SD	N	z-score	SD		
Speed of processing	8	-0.4	0.9	9	-0.3	0.7	0.1	.76
Accuracy	8	0.2	0.3	9	-0.2	0.5	0.4	.06
Executive function	8	0.2	0.3	9	0.2	0.4	0.0	.90
Well-being	8	-1.8	3.1	6	-1.6	1.3	-0.2	.86

* T-test is significant at the 0.05 level (2-tailed).
** T-test is significant at the 0.01 level (2-tailed).

Table 5.9. - *Correlations between composite z-scores and concurrent and adulthood Phe levels.*

	Speed of processing	Accuracy	Executive function	Well-being
Concurrent Phe				
Pearson's r	-0.21	0.27	-0.12	0.07
p-value	0.39	0.26	0.64	0.8
Adulthood Phe				
Pearson's r	-0.22	0.23	-0.06	0.04
p-value	0.36	0.34	0.81	0.89

Discussion

In this chapter, we carried out exploratory analyses with young and middle-aged AwPKU, along with age matched controls, to address the research question, “How does increasing age impact cognition and well-being in AwPKU?”

Cognitive outcomes

Overall, middle-aged AwPKU demonstrated significantly less impairment across cognitive domains than young AwPKU. These findings suggests that there is no evidence of accelerated ageing impacting cognitive health in AwPKU. In fact, our results suggest the opposite effect of ageing on cognition in this group, with severity of impairment appearing to decrease with age. These findings could be due to a multitude of impacting factors. For example, previous research has indicated that cognitive abilities in PKU improve from adolescence to adulthood, despite increased brain abnormalities (Jahja, Van Spronsen et al., 2017; Mastrangelo et al., 2015; Nardecchia et al., 2015). This may be due to the development of compensatory strategies in this cohort in early adulthood. These compensatory strategies, then, may provide a cognitive advantage to AwPKU as they start to age, compared to healthy controls who need to develop these strategies in later life as the ageing process begins to deteriorate the cognitive abilities that they have previously relied upon.

Additionally, the difference in testing modalities between young and middle-aged cohorts may be an influencing factor in the differing impairment levels observed between these two age groups. Whilst previous research has found good consistency between cognitive assessments delivered face-to-face and remotely (Abdolahi et al., 2016; Brearly et al., 2017; Chapman et al., 2021; Dauphinot et al., 2020; Lindauer et al., 2017; Randolph et

al., 2014; Stillerova et al., 2016; Zeghari et al., 2021), there has been some evidence to suggest that certain tasks may elicit better performance from mild to moderately intellectually disabled participants when delivered online (Ashworth et al., 2021). It may be, then, that the clinical cohort in this study performed better when tasks were delivered remotely, than when they were delivered in a face-to-face setting.

Additionally, previous research has found impairments in executive functions, such as planning, switching, and reasoning, in participants with PKU (for reviews see Burlina et al., 2019; Canton et al., 2019; Hofman et al., 2018). However, no such impairment was identified in middle-aged AwPKU in our study, with younger AwPKU demonstrating a significantly larger impairment than middle-aged AwPKU in this domain. This large difference in findings between age groups could be reflective of better executive function in middle-aged AwPKU, or it may be due, at least in part, to the differing testing modalities used between age groups. Moving assessments online for this study necessitated the removal of many of the more sensitive executive function tasks included in the original face-to-face testing battery (i.e. Tower of Hanoi, Digit Symbol coding, Trail making). As such, only the WCST was included as a measure of task-switching. Other measures included in the executive function composite score were representative of inhibitory abilities (rather than planning, switching, or reasoning abilities), an area of executive function previously shown to be relatively unaffected in AwPKU (e.g., Palermo et al., 2017; Romani et al., 2017).

A final consideration when exploring differences in impairment levels between young and middle-aged AwPKU is the known neurodiversity present within the PKU population (Romani et al., 2019). Cognitive impairment varies dramatically within populations of AwPKU, and it is likely that the subset of AwPKU who agreed to participate in a remotely delivered study, during a global pandemic, were those who felt relatively self-assured and confident with online processes. Additionally, participation was limited to those individuals who had access to the necessary hardware to complete the tasks and possessed sufficient understanding of technology to successfully run the tasks at home. Individuals fitting these criteria, then, are likely to be those with a higher level of cognitive functioning. Furthermore, smaller differences in performance abilities between middle-aged AwPKU and controls could be reflective of a difference in motivation between these two groups. Whilst control participants were financially motivated to take part in the study, AwPKU were motivated by a desire to improve understanding of their condition, as well as of their own current cognitive ability. Therefore it is likely that participants with PKU would try harder and be more mindful of their environment whilst completing online assessments.

Well-being outcomes

This study further hypothesised that middle-aged AwPKU would demonstrate poorer scores on measures of well-being than controls, therefore supporting previous findings (e.g. Brumm et al., 2004; ten Hoedt et al., 2011). This hypothesis was supported by self-reported emotional well-being measures which indicated that AwPKU were significantly more anxious and depressed than age-matched controls. Questionnaire data further found that AwPKU were more negatively impacted by the ongoing COVID-19 lockdown in terms of feelings of anger, anxiety, exhaustion, isolation, sadness, stress, and overall worries about their general health. Interestingly, controls demonstrated a decrease in anger and worries about their health during the lockdown, possibly reflective of an increased feeling of security due to the extensive application of health-protection protocols put in place during this time, leading to a decreased risk of health issues in those who were not considered to be significantly vulnerable to the Covid-19 virus. AwPKU, however, were understandably more worried about their health during this time. This was likely due to the unknown potential interactions between their existing conditions and the Covid-19 virus. These findings support those of recent studies which have reported that adults with chronic illnesses have been disproportionately impacted by the COVID-19 pandemic in terms of well-being and mental health (Khan et al., 2021; Umucu et al., 2021; Wang et al., 2020).

Impact of metabolic control

Overall, less impact of metabolic control on performance in AwPKU was found than was expected based upon the current literature. Some studies have reported no differences in performance between AwPKU with high and low blood-Phe levels (Aitkenhead et al., 2021; Bartus et al., 2018). Aitkenhead et al. (2021) did report that higher average adulthood (from age 18+) and concurrent Phe levels were significantly correlated with slower performance and worse learning scores in their cohort of AwPKU, although once Phe concentrations at all other life stages were controlled for, speed of processing was only associated with concurrent Phe and Phe in early childhood, and learning was only associated with early childhood Phe. A number of previous studies have reported significant differences in the domains of sustained attention, working memory, speed of processing and executive function (Bik-Multanowski et al., 2011; Moyle, Fox, Bynevelt, et al., 2007; Romani et al., 2017; Schmidt et al., 1994; Smith et al., 1996; ten Hoedt et al., 2011) which was not replicated in the current study. This study found no significant differences when composite scores for speed of processing, accuracy, or executive function were compared between AwPKU with low- vs. high-concurrent Phe or between AwPKU with low- vs. high adulthood

Phe. A trend difference in overall accuracy, however, did emerge between AwPKU and with low- vs. high-adulthood Phe (p=.06), with the low-adulthood Phe group producing less errors overall than the high-adulthood Phe group. Correlation analyses found no associations between cognitive performance and either concurrent or adulthood Phe levels.

With regards to the impact of metabolic control on well-being in middle-aged AwPKU, no significant differences were found between high- and low-Phe groups on standardised measures of emotional well-being or QoL, supporting Aitkenhead et al.'s (2021) assertion that metabolic control does not impact anxiety or depression in adulthood. Correlation analyses further found no relationship between concurrent or adulthood Phe levels and measures of emotional well-being or QoL.

Limitations

A number of limitations must be considered when interpreting the findings of this study, in particular those caused by the ongoing COVID-19 pandemic and subsequent lockdowns in place throughout the data collection period. Power-calculations provided an initial recruitment target of 25 AwPKU and 50 controls. However NHS recruitment holds, alteration of assessment modalities, and the increased stressors associated with the pandemic (e.g., home-schooling) meant that many participants could not spare the time to participate, resulting in recruitment numbers for this study being much lower than this. This was particularly true for the cohort of clinical participants. As such, the participants recruited from clinics may not have been representative of the clinical population as a whole.

Results of exploratory analyses comparing the performance of different cohorts on individual tasks are also limited, due to the increased risk of a Type I error when t-test analyses are conducted across a large number of individual measures. Whilst this limitation was recognised, the potential to identify clinically significant differences within a cohort of participants with a rare clinical disorder was considered sufficiently important to conduct analyses across the entire assessment battery with a more liberal p-value than may otherwise have been applied to such a large number of analyses. The potential impact of Type 1 error on the interpretation of results, however, was mitigated by focussing our results on comparisons of overall pattern of performance, and differences in composite scores across domains.

In addition to recruitment issues stemming from the pandemic, the adaptation of assessments to allow remote testing further limits the reliability of data gathered for this

study. The requirement for all participants to have access to, and knowledge of how to use, a computer with internet and video call capabilities will have likely biased recruitment to those from a higher socioeconomic background and those with fewer concerns about their cognition. This bias, therefore, may be responsible for the apparent smaller impairments observed in participants tested through this modality compared to those assessed using face-to-face assessments.

Finally, the impact of the pandemic on participants' reported well-being must be considered when interpreting responses to questionnaires. It may be that, whilst PKU does have some detrimental impact on well-being in middle-aged AwPKU, this effect was inflated by the increased negative impact of the pandemic and lockdown on well-being in this population, as demonstrated by increased negative responses of AwPKU compared to controls when asked about the impact of the COVID-19 lockdown on their mood.

Directions for future research

This study highlights the need for future research in two areas. First, more research is required to better understand the impact of early-treated PKU on neurological health in middle-aged AwPKU. Assessments with larger cohorts of clinical participants are needed, as well as longitudinal studies that investigate the ongoing impact of PKU on cognition, neuropathology, and neurophysiology throughout the lifespan, ideally utilising a combination of cognitive assessments and scanning technologies. Furthermore, early-treated AwPKU who are currently in their 40s and 50s should continue to be monitored, using uniform assessment batteries, to allow assessment of impairments as these participants move into older age and start to experience additional cognitive decline as a result of normal ageing. Interactions between cognitive decline and metabolic control during this period are also important, to allow clinicians to give the best possible advice to AwPKU who are reaching old age.

In addition to research assessing the impact of PKU on adults of increasing age, this study has highlighted the need for new research investigating the efficacy of remote assessment of cognitive abilities. The COVID-19 pandemic has led to the widespread adoption of telemedicine and has caused many researchers to move their studies online (Iyer et al., 2021). This has particularly been the case for studies with clinical cohorts (Elbaz et al., 2021; Rohr et al., 2021; Sathian et al., 2020). Due to the unique circumstance of the last few years, this shift towards remote assessments has happened rapidly, resulting in

many unanswered questions remaining about the most effective modalities for this type of testing, and their comparability to face-to-face equivalents. Whilst studies are now starting to compare the efficacy of cognitive assessments delivered remotely via video call vs. telephone call (see Elbaz et al., 2021 for a review; Goodman-Casanova et al., 2020; Lai et al., 2020), research directly comparing cognition data gathered in face-to-face vs. remote environments is vital to ensure that data quality does not suffer as a result of the ongoing trend towards conducting assessments remotely rather than in laboratory settings. Initial research comparing online teleconsultations to in-person assessments has found no significant differences in scores (Arighi et al., 2021; Zeghari et al., 2021), however, additional research comparing findings across a range of different assessment batteries, and a range of cohorts with varying needs, is vital moving forward.

Chapter 6: The Impact of Metabolic Control on Cognition, Neurophysiology, and Well-Being in PKU: A Systematic Review and Meta-Analysis of the Within-Participant Literature

This chapter has been published by Molecular Genetics as a self-contained manuscript (DOI: https://doi.org/10.1016/j.ymgme.2022.106969). As such, there may be a degree of redundancy and repetition, especially with respect to the introductory section.

Phenylketonuria (PKU) is a metabolic disease where Phenylalanine (Phe) rises much above normal levels. Cross-sectional and correlational studies provide valuable information on the importance of maintaining low blood-Phe to achieve good outcomes, but they may be confounded, at least partially, by differences in participant demographics. Moreover, the effect of Phe at older ages is difficult to ascertain because of strong associations between Phe levels across ages. Within-participant studies avoid confounding issues. We have reviewed these studies.

We followed PRISMA guidelines to search the literature for studies reporting the impact of Phe changes within participants. Phe was either increased or decreased through diet relaxation/resumption or through pharmacological interventions. Forty-six separate articles reported, singly or in combination, results on cognition (N=37), well-being (N=22) and neurophysiological health (N=14). For all studies, we established, in a binary way, whether a benefit of lower Phe was or not demonstrated and compared numbers showing benefit versus a null or negative outcome. We then analyzed whether critical parameters (e.g., length of the study/condition for the change, size of Phe change achieved) influenced presence or absence of benefit. For a subset of studies that reported quantitative cognitive outcomes, we carried out a meta-analysis to estimate the size a change in cognitive performance associated with a change in Phe and its significance.

There were significantly more studies with benefits than no benefits, both for cognitive and well-being outcomes, and a trend in this direction for neuro-physiological outcomes. The meta-analysis showed a highly significant effect size both overall (.55) and when studies with adults/adolescents were considered separately (.57). There was some indication that benefits were easier to demonstrate when differences in Phe were larger and achieved across a longer period, but these effects were not always consistent.

These results reinforce results from the literature by demonstrating the importance of lower Phe in children as well as in adolescents and adults, even when confounding factors in group composition are eliminated. The field would benefit from further studies where Phe levels are contrasted within-participants to ascertain how much Phe needs to be changed and for how long to see a difference and which measures demonstrate a difference (e.g., which cognitive tasks).

Introduction

Phenylketonuria (PKU) is a metabolic disorder affecting between 1 in 10,000 to 1 in 12,000 births (Christ, 2003; Scriver & Kaufman, 2001). Individuals with PKU have mutations in the genes which code for the enzyme phenylalanine hydroxylase (PAH) which is essential for the hydroxylation of phenylalanine (Phe) into tyrosine in the liver. Since tyrosine is a precursor of dopamine, this reduces the availability of this neurotransmitter. Moreover, large amounts of Phe in the blood stream reduce the chances of other essential amino acids such as tryptophan (a precursor of serotonin) being able to pass through the blood-brain barrier as they must compete with the Phe, leading to neurotransmitter imbalances. This, in addition to the accumulation of Phe in the brain, which is toxic for white matter, impairs brain functioning (Blau et al., 2010; Surtees & Blau, 2000). If left untreated, PKU causes significant neurological impairment including intellectual disability, microcephaly, seizures, and behavioural difficulties (Albrecht et al., 2009; Blau et al., 2010). Fortunately, early and continuous treatment (e.g., via reducing Phe intake) prevents much of the neurocognitive damage associated with untreated PKU.

The importance of maintaining low blood-Phe levels from infancy is well documented, with several reviews reporting findings of cognitive impairments being particularly evident in individuals with poor metabolic control (see Canton et al., 2019; Christ et al., 2010; Hofman et al., 2018 for reviews). Beyond cognition, early-treated patients with PKU have been found to exhibit a number of emotional and behavioural deficits, including high levels of depression, anxiety, hyperactivity, inattention, poor self-image, withdrawal, and a lack of autonomy and drive (see Bilder et al., 2016; van Spronsen et al., 2011; Smith & Knowles, 2000 for reviews). Metabolic control has also been linked to these cognitive and emotional impairments (e.g., Brumm et al., 2004; Bik-Multanowski et al., 2011; ten Hoedt et al., 2011; Romani et al., 2017). Finally, the high prevalence of cerebral white matter abnormalities in PKU has been recognised by several scoping reviews, with consistent findings of an association between increased blood-Phe levels and more severe abnormalities (Anderson & Leuzzi, 2010; Ferreira et al., 2020; van Spronsen et al., 2011).

Whilst the importance of metabolic control for maintaining cognition, mental health, and brain health in children with PKU (CwPKU) has been widely accepted, the importance of maintaining low blood-Phe levels during and after adolescence is less clear. European guidelines recommend maintaining Phe levels between 120 and 360 µmol/L up to 12 years of age and <600 µmol/L after that (van Spronsen et al., 2017; van Wegberg et al., 2017). US guidelines are even more strict, recommending maintaining Phe levels <360 µmol/L throughout life (American College of Medical Genetics and Genomics, ACMG, Vockley et al., 2014). Many individuals with PKU, however, choose to relax their treatment diet in adolescence or early adulthood so that Phe levels often exceed what is recommended. For example, Phe levels were found to remain steady in Dutch CwPKU until around 13 years of age, after which they started to increase (Crone et al., 2005) and 34% of 13-18-year-olds in the US were found to no longer attend PKU clinics (Berry et a., 2013). Walter and White (2002) reported that 30% of Phe samples were above 700 µmol/L (the target recommended at the time by the United Kingdom's National Society for Phenylketonuria or NSPKU) in children younger than 10 years, but this figure rose to 80% in people aged 15 and older, with levels reaching a mean of 800 µmol/L by the age of 15 and increasing after that (Walter & White, 2004). A clearer understanding of the potential impact of reduced metabolic control throughout the lifespan, therefore, is needed. This will allow clinicians to provide their patients with the necessary guidance to make an informed decision about continuing, relaxing or abandoning their treatment diet post-childhood. Also, the parameters of blood-Phe required to maintain optimal cognitive, emotional, and neurophysiological health in children and adults with PKU, are unclear, as well as the extent to which damage caused by excessively high levels can be reversed.

Most of the existing evidence on the importance of metabolic control comes from studies which have used either a correlational methodology or a contrast between groups of participants with high vs. low metabolic control. The problem with these studies is that variability in cognition, emotional health and brain health will be determined, not only by metabolic control, but also by individual differences in genetic endowment, educational opportunities, socio-economic status, etc., which are difficult to control. Within-participant study designs offer a gold-standard way to overcome these limitations, but results from the literature are limited. The aim of this paper is to carry out a systematic review of studies which have used a within-participant design to see whether results can be accrued in a meaningful way. Before detailing our aims and methodology, however, we will briefly review existing evidence from studies using between-participant designs. We will concentrate, in particular, on evidence regarding levels of Phe that can be considered safe. We will review evidence in the cognitive, well-being and neurophysiological domains.

Studies involving both children and adults with PKU have found worse performance in groups of individuals with higher Phe levels, but interpretations are limited by the fact that different criteria have been used for group selection. Studies considering IQ in CwPKU have reported impaired intelligence in groups with average Phe levels between 800 µmol/L and 1300 µmol/L (Azen et al., 1991; Holtzman et al., 1986; Smith et al., 1990; Waisbren et al., 2007). In other cognitive domains, performance has been found to be impaired at lower Phe levels. For example, Schmidt et al. (1996) found that, in CwPKU, 1-9 year-olds with mean blood-Phe levels **>620 µmol**/L were more impaired in sustained attention and calculation than children with mean Phe **<=240 µmol**/L. Similarly, Huijbregts et al. (2002) found that 7-14 year-olds with blood-Phe levels **>360 µmol/**L made more errors in focussed attention tasks and were slower at feature identification than children with mean Phe **<=360 µmol**/L. Waisbren et al. (2007) looked at correlations between IQ and childhood blood-Phe levels across 40 different studies. Meta-analytic results showed significant correlations between IQ and blood-Phe levels within the first 12 years of life, with each increase in blood Phe of 100 µmol/L during this age range predicting a decrease in IQ of 1.9 to 4.1 points. Concurrent blood-Phe levels equally correlated significantly with IQ, with each increase of 100 µmol/L predicting a drop in IQ of 0.5 to 1.4 points, (over a range of Phe from 394-666 µmol/L in the studies included in the meta-analyses).

Between-subject studies with adults and adolescents with PKU (AwPKU) suggest that they may be able to maintain higher blood-Phe levels than children without compromising cognitive abilities. Romani et al. (2019) assessed cognitive impairment in 56 early-treated AwPKU, and found that maintaining Phe levels **<600 µmol/L** post-childhood reduced the risk of cognitive impairment in this cohort by approximately 30%, with AwPKU with current levels below this threshold demonstrating little to no cognitive impairment. Analyses of the impact of childhood Phe levels on adult outcomes, meanwhile, found a lower threshold of 360 µmol/L, after which cognitive impairments became apparent. Romani et al. (2017), compared performance in AwPKU with mean adult blood-Phe levels of **>950 µmol/**L, to those with mean levels of **<650 µmol**/L, and reported differences across a range of tasks including differences in executive function (excluding inhibition), and in measures of visuo-spatial attention, short-term/working memory, sustained attention, written and spoken language, and verbal memory and learning. When compared to controls, those with Phe <650 µmol/L showed a modest, but still significant, impairment. Furthermore, this study found that both lifetime and concurrent adult Phe levels were correlated significantly with performance on most tasks (excluding inhibition, short-term memory, and language tasks).

However, Jahja, et al. (2017) reported differences in performance between groups of AwPKU contrasting at lower Phe levels. AwPKU with lifetime Phe **>=360 µmol**/L performed worse than both controls and AwPKU with Phe <360 µmol/L in working memory and sustained attention. Moreover, concurrent blood-Phe levels correlated with slower performance in several tasks, including a feature integration task (all conditions), a memory search task (high working memory condition only), and a sustained attention dots task. Lifetime Phe also correlated negatively with accuracy in the high working memory load condition of the feature integration task. Finally, Aitkenhead et al. (2021) assessed correlations between metabolic measures and performance in several cognitive tasks in a sample of 154 early-treated AwPKU with Phe levels ranging from 200-800 µmol/L. They found that performance on digit symbol coding was best predicted by concurrent Phe, with a decrease of 1.05 SD for every increase of 1000 µmol/L. This test requires participants to write down symbols under corresponding digits following a key, and it assesses a variety of skills impaired in PKU including speed of processing, visuo-motor coordination, and working memory. Importantly, correlations with concurrent Phe were maintained even when Phe levels at previous ages were taken into consideration. Similarly, a task involving learning of lists of words was best predicted by concurrent Phe, although this contribution did not remain significant when Phe at previous ages was taken into consideration. These findings suggest that, even in adulthood, maintaining Phe at lower levels is beneficial for certain domains.

Metabolic control has also been found to influence emotional well-being. For example, Smith et al. (1988) measured behaviour in 544 early-treated, 8-year-old CwPKU by asking school teachers to assess the frequency of abnormal behaviours using the Rutter Behaviour Questionnaire. Those with good metabolic control (average levels **<600 µmol/L**) had fewer problems than those with worse control (>600 µmol/L) although they were still impaired compared to healthy controls (1.6 vs. 2.2. times more likely to show deviant behaviours). Relaxing dietary control in childhood has also been reported to impact well-being later in life. Koch et al. (2002) reported more well-being difficulties (including phobias and depression) in AwPKU who had discontinued their diet aged 10, compared to those who were still following the diet, with 8/58 participants who were off-diet suffering from hyperactivity and 11/57 from lethargy. No such symptoms were reported by the nine participants who remained on diet. These findings support those of Holtzman et al. (1986), who found a significant correlation between the age at which participants lost dietary control (defined as levels reaching **908 µmol/L),** and higher behaviour problem scores.

With regards to well-being in adulthood, Jahja, et al. (2017) found that increased depressive and avoidant personality traits in AwPKU compared to controls were only apparent when lifetime blood-Phe level was **>=360 μmol/L**, with no significant correlation found with concurrent blood-Phe. Burton et al. (2013) reported higher levels of psychiatric distress in children, adolescents, and adults with PKU with higher median Phe and concurrent Phe, (582 and 683 μmol/L respectively) compared in those with lower levels (354 and 442 μmol/L). These findings suggest that negative effects on well-being may become apparent as soon as levels rise above 350-400μmol/L. However, other studies have reported no effect, or a negative effect, of metabolic control on emotional well-being (e.g., Aitkenhead et al., 2021; Brumm et al. 2004, Palermo et al., 2020; Pietz et al., 1997). These contrasting findings may reflect the fact that following a PKU diet may have mixed effects because, on the one hand, reduced Phe will improve brain health with positive effects on well-being but, on the other hand, following a strict diet may be stressful and unsociable, causing emotional difficulties.

Finally, outcomes in PKU can be measured in terms of brain abnormalities and neurological symptoms beyond cognitive impairments. White matter pathology has been commonly observed in people with PKU, with significant associations with elevated Phe levels reported across all ages (for studies with adults only see Cleary et al., 1994, Hellewell et al., 2021, Nardecchia et al., 2015; for studies with mixed cohorts see Hawks et al., 2019; Mastrangelo et al., 2014; Thompson et al., 1993). For example, Thompson et al. (1993) assessed white matter changes in early-treated (N=25) and late-treated (N=9) adults and children with PKU aged 8 to 33 years who had been treated until a minimum age of 7 years and found that, when other factors in the model were controlled for, the likelihood of more severe MRI abnormalities increased 1.5 times for every 100 μmol/L rise in concurrent blood-Phe levels, and 1.3 times for every additional year off a low-Phe diet (current Phe range: 350-2000 μmol/L). Mastrangelo et al. (2015) reported that, in AwPKU with concurrent Phe levels ranging from 148μmol/L to 1900 μmol/L, each increase in blood-Phe of 100 μmol/L was associated with a 0.46 increase in participants' white matter severity score[1]. Use of modern neuroimaging techniques, such as diffusion kurtosis imaging (DKI) have further expanded our understanding of the impact of PKU on white matter integrity. With this technique, Hellewell et al. (2021) reported findings of subclinical white matter abnormalities in AwPKU, as well as a significant relationship with metabolic control, with AwPKU with

[1] WMSS - see Leuzzi et al. 1993; Severity scores 0-4 points for increasing T2 signal alterations, plus 1 point for each location, plus 1 point for presence of patchy lesions, plus 1 point for decreased T1 signal.

lower adult Phe levels demonstrating more preserved white matter microstructure (mean lifetime Phe 461 vs. 609 µmol/L; mean adulthood Phe 487 vs. 1645 µmol/L). Importantly, there were suggestions that white matter damage mediates cognitive impairments. In a 3-year longitudinal study, Hawks et al. (2019) using DTI and MRI found different trajectories of white matter development in 35 participants with PKU aged 7-21 years compared to controls, as well as whole-brain and regional white and grey matter abnormalities. Importantly, mediation analyses found that whole-brain diffusivity, as well as regional diffusivity in the corpus callosum and centrum semiovale, mediated the relationship between metabolic control and executive functioning.

Neurological symptoms beyond cognitive impairments involve mainly motor and visual disturbances (e.g., tremors, motor stereotypies, tics, epilepsy blurred vision). These symptoms occur often in untreated people with PKU (see Mainka et al., 2021), but also, occasionally, in early-treated people (e.g., see González et al. 2011, who found respectively in early- vs. late-treated % -- epilepsy: 1.1% vs. 31%; tremor: 12% vs. 93%; clumsiness: 11% vs. 90%). White matter abnormalities and neurological symptoms were present in 97% of late-treated participants, and in 25% of early-treated participants (González et al. 2011). Other studies have reported cortical blindness and vision loss associated with periventricular white matter and cortico-subcortical occipital lesions, and neuromyelitis optica[2] in early-treated AwPKU (Anwar et al., 2013; Rubin et al., 2013). Importantly, when these symptoms appear later in life they have been found to regress or ameliorate after the reintroduction of a PKU diet, or after introduction of a PKU diet in a few cases of participants with missed diagnosis who had never been on diet (see Jaulent et al., 2020).

To summarize, cross-sectional and correlational studies provide valuable information about associations between blood-Phe levels and outcomes in PKU. However, this evidence suffers from several limitations. Firstly, it is difficult to derive conclusions about the Phe levels which are safe at different ages because of the differences in the cut offs used by different studies to compare participants with good vs. poor metabolic control. Secondly, between-group comparisons may be confounded by differences in demographics and other individual variables. For example, AwPKU and parents of CwPKU with better metabolic control may have more education and a higher socio-economic status than those with worse control (for evidence in this sense see Channon et al., 2007; Koch et al., 2002; MacDonald et al., 2008). Therefore, better cognition, and possibly better mental health and brain health, could reflect a more stimulating environment instead of (or in addition to) better metabolic

[2] A demyelinating autoimmune inflammatory process affecting the central nervous system

control (but for evidence of effects of Phe controlling for education see Cazzorla et al., 2014; Romani et al., 2019). As another example, sensitivity to Phe can vary from one individual to the next, and this may increase noise in correlational and between-groups studies (e.g., some individuals with untreated PKU seem to escape the worse effects of Phe; see van Vliet et al., 2018 for a review). Thirdly, comparisons between participants with PKU and healthy controls do not account for differences caused by living with a chronic disease (such as perceived social exclusion or the stress of arranging to eat special foods) and/or the potential nutritional deficiencies caused by a PKU diet which can contribute to poor outcomes beyond high Phe levels. Finally, between-participant studies do not address the question of the extent to which existing deficits are reversed or exacerbated by changes in Phe and provide little insight into additional factors (e.g., length of treatment, magnitude of metabolic change) that contribute to the likelihood of observing changes in outcome.

Within-participant studies offer the ideal conditions for addressing questions related to the effects of changes in Phe, while at the same time controlling for any confounding individual variables. Studies of this type are limited in number and have used participants with different levels of Phe at different times. However, carrying out a systematic review can both accrue results, and capitalise on between-study differences in metabolic levels to provide evidence for the levels at which Phe becomes detrimental, and for the differences in Phe able to produce a difference in outcomes. One can carry out correlations between Phe variables and outcomes but also compare the parameters of studies where differences in Phe did, or did not, produce differences in outcomes. In particular, one can compare baseline Phe and differences in Phe. This will provide us with some indication of the parameters within which Phe manipulations can be effective.

Whilst several scoping and systematic reviews of cross-sectional studies exist, no systematic comparisons of within-participant effects of altering Phe levels have been carried out so far. We reviewed within-subject studies reporting cognitive, well-being, and/or neurophysiological outcomes in early- and late-treated children, adolescents, and adults with PKU, whose blood-Phe levels changed through dietary or pharmaceutical intervention. We wanted to establish, first, whether increasing or decreasing Phe had significant effects, respectively worsening, or improving outcomes and whether this was true both in child and adult samples. Additionally, we wanted to assess how effects were related to metabolic parameters, both by carrying out correlations, and by comparing the parameters of studies which did/did not find significant effects. For well-being and neurophysiological/neurological outcomes we could only carry-out systematic reviews and provide descriptive statistics of results since it was not possible to accrue results in a quantitative way. However, this was

possible for a subset of the studies reporting cognitive outcomes and a meta-analysis was carried out on this subset.

Method

Where applicable, our review was conducted in adherence with the updated guidelines of the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Statement (Page et al., 2020).

Literature search

A literature search was conducted using the Web of Science, PubMed, Cochrane Library and PsychArticles databases (from inception to August 2022). Abstracts, titles, and author keywords were searched with the combination of terms:

PKU OR Phenylketonuria
AND Phe OR Phenylalanine OR metaboli*
AND manipulat* OR load* OR diet* OR pharmacolog* OR interven* OR treat* OR restrict* OR resum*

'*' functions as a wildcard symbol to broaden the search by finding words that start with the same letters.

The inclusion criteria for a study were that it:

1) Included human participants with phenylketonuria (children, adolescents, or adults; classical or mild PKU);
2) Changes in participants' blood-Phe concentrations either through termination, relaxation, resumption, or restriction of a low-PKU diet, or through pharmacological interventions (e.g., treatment with BH4, Pegvaliase, Phe loading); we included both longitudinal studies and 'proper' intervention studies;
3) Included post-change cognitive, well-being, or neurophysiological results;
4) Included baseline and/or post-change blood-Phe levels;
5) Was written or available in English.

The exclusion criteria for a study were that it:

1) Was a review, meta-analysis, or conference abstract;

2) Phe change observed over an average period of more than 4 years to limit the impact of more general long-term influences.

Titles and abstracts of returned papers were initially screened by the first author to exclude duplicate and clearly ineligible studies. After the initial screening, all retrieved papers were read in full. Further references were obtained from the reference lists of reviews and relevant papers.

From the selected papers, the following data were extracted for the PKU groups:

1) Type of intervention: dietary or pharmacological;
2) Number and age of participants;
3) Age at initiation and termination of treatment;
4) Length of study-related Phe change – this is the time span over which the change in Phe was achieved;
5) Blood Phe concentrations pre- and post-change, and average difference between these two measures;
6) Outcome measures: cognitive, well-being, and neurophysiological (including neurological symptoms other than cognitive), at baseline and post Phe change.

All studies were checked for inclusion criteria and data extraction by the third author (C.Romani)[3]. Where papers did not include all data of interest, we included all papers that met a minimum information requirement of number of participants, age, Phe pre and/or post change, and outcome(s) in at least one cognitive, neurophysiological, or well-being domain. Where articles compared outcomes in two or more different Phe conditions (e.g., increasing and decreasing Phe) or outcomes from different domains (cognitive, well-being, or neurophysiological), these were considered separately as individual 'studies'. When only results in graph format were available, we used digitization software (PlotDigitizer) to calculate actual scores. Where blood Phe concentrations were given in mg/L or % mg, these were converted into μmol/L. Where the range of blood-Phe concentrations was given instead of the standard deviation, standard deviations were approximated using sd = (maximum-minimum)/4.

[3] The study by Ficicioglu et al. (2013) was not included in our review because, although it aimed to reduce Phe level through the introduction of sapropterin, it was not successful in this endeavour as the five participants did not show any decrease in Phe level before/ after sapropterin.

Data analyses

Most studies reported well-being or neurophysiological health outcomes only in qualitative terms or in varied formats which did not allow quantitative aggregation. The same was true for a number of studies reporting cognitive outcomes. In these instances, we computed descriptive statistics in terms of number of studies showing a benefit of lower Phe levels vs. those showing no benefit or an effect in the opposite direction. Studies were defined as showing a benefit if changes in outcomes were all in the expected direction or, when changes were in both directions, if significant changes were only in the expected direction or, when statistical analyses were not reported, if positive changes were reported for at least 50% of participants and changes in the opposite direction for less than 20% of participants.

We ran different types of analyses. We compared numbers of studies showing benefit vs. no benefit for each outcome type (cognitive, well-being, neurophysiological) using χ^2 tests. We assessed differences between the parameters for studies showing or not showing benefits using univariate analysis of variance; and we assessed whether size of Phe-difference and/or length of study/condition predicted outcomes using binomial logistic regression. It is unclear whether results are better reflected by considering arithmetic averages or averages weighted by number of participants. Weighted averages favour large studies, which are more trustworthy, but underestimate the information provided by smaller studies which give converging sources of information from different paradigms/groups of participants. Arithmetic means have the opposite advantages and disadvantages: they properly consider results of diverse studies but fail to give more weight to studies with more participants. Therefore, we have analysed results both using arithmetic average and weighted averages, but taking care, in this last case, not to inflate results. This was managed by ensuring degrees of freedom reflected number of studies and not number of participants. If convergence is achieved with different analyses this will increase our confidence.

Finally, for the subset of cognitive studies which reported quantitative data, we computed effect sizes and ran a meta-analysis of those.

Meta-analysis. We standardised differences in performance before and after changes in Phe using Hedge's g effect sizes (ESs). A Hedges' g ES is computed by dividing the difference in measures taken before and after a change in Phe by the pooled standard deviation of the two conditions and then applying a correction for small sample sizes:

Hedges' g $= \left(\frac{\bar{y}_{low} - \bar{y}_{high}}{s_p}\right) \times \left(1 - \left(\frac{3}{4(df)-1}\right)\right)$

Where '$\bar{y}_{low}$' is cognitive performance when Phe is low; '$\bar{y}_{high}$' is performance when Phe is high; 's_p' is the pooled standard deviation calculated as $\sqrt{\frac{1}{2}\left(s_{low}^2 + s_{high}^2\right)}$; 's' is the SD of each condition; and n is the number of participants. *Df* is an estimate of the degrees of freedom from Pustejovsky (2016):

$$\frac{2(n-1)}{1 + r^2}$$

To make sure that positive effect sizes always reflected a benefit of lower Phe, we subtracted outcomes in high Phe conditions from outcomes in low Phe conditions and then multiplied the effect size by -1 when higher values indicated worse performance (e.g., error or RT measures). Thus, positive ESs always reflected better performance in the low-Phe condition.

The meta-analysis derived a weighted, cumulative ES for change in cognitive performance after a change in Phe and assessed its significance. Individual ESs were entered into a mixed-effect meta-analysis. The model included random effects for 'study' and 'measure' nested within study. This accounted for studies where the same cohort of participants was tested with different tasks (or measures). A random effects model was appropriate because we assumed that the cumulative effect size reflected a *population* of effects which differed depending on the characteristics of the PKU group tested and the type of cognitive measure used. In contrast, a fixed effects model would assume that variation is solely due to random error (see Hunter & Schmidt, 2000 for a detailed account of fixed vs. random effects models in meta-analyses). Finally, since results were within-participants, the Phe conditions are, by definition, not independent of each other. Because the correlation between pre and post Phe-change measures could influence the estimate of the variance of effect sizes, we included a term for the correlation between measures (see Pustejovsky, 2016, citing Borenstein, 2009). Pre/post correlations were typically not reported in the studies we reviewed, so we ran a sensitivity analysis with estimated correlations of 0.2, 0.5, and 0.8 (see Higgins et al., 2019). If our estimate of the ES was relatively insensitive to different levels of correlation, this reassured us of the stability of our estimate.

With all meta-analyses there is a risk of a publication bias inflating the estimated ES (Begg, 1994). The 'file drawer problem' is a label for the tendency not to publish non-significant findings, or findings in the direction opposite from those expected. This is more likely for studies with small samples. We assessed the potential influence of a publication bias through a funnel plot and the influence of small studies through a sensitivity analysis. The funnel plot plotted ESs on the x-axis and a measure of precision (in our case, sample sizes) on the y-axis, with high precision at the top. This should produce a plot with a funnel shape because ESs should cluster tightly around the estimated mean at the top, where the sample size is large and precision is high, and distribute more widely at the bottom, where measures have less precision. If there is no publication bias, ESs should be distributed *symmetrically* around the mean, even when sample sizes are small. In the case of publication bias, instead, there may be more measures from small studies on the positive side of the mean, with missing studies on the other side. Funnel plots often use standard error as a measure of precision, but, in our case, the number of participants is more appropriate since standard errors can be affected by variability in metabolic control which could have a larger range in larger studies. Thus, a smaller standard error may not automatically signal a more precise estimate of population values.

A sensitivity analysis (or cumulative meta-analysis) assessed whether our cumulative estimate of ES was influenced by studies with small samples (see Borenstein, Hedges, Higgins & Rothstein, 2009; Chapter 30, p. 288). In this analysis, we ordered our ESs according to the number of participants and subdivided them into bins. The first bin included the largest studies and the following bins added progressively smaller studies. There were six bins, with each bin adding 4 studies. If effects were unduly biased by small sample studies, the cumulative effect size should increase as smaller studies are added. Instead, if the effect size stays the same or gets smaller, we can be confident that the estimate is not inflated by small samples or publication bias (see logic outlined in Borenstein, et al., 2009, Chapter 30).

Results

The initial keyword search yielded 4,620 results. 1,571 were excluded as duplicates and a further 2,914 were excluded through inspection of titles and abstracts. The remaining 162 articles were subject to full text review and an additional 27 studies were added through examination of reference lists or reviews. 116 records were excluded for a number of reasons (see Figure 6.1; flow chart). Forty-six articles were finally included for review. As described above, articles were split into separate 'studies' if there were multiple Phe-change

conditions (e.g. increasing vs. decreasing Phe), or multiple participant groups were assessed. Overall, we considered 73 different studies, 37 of which reported cognitive outcomes, 22 well-being outcomes, and 14 neurophysiological outcomes and effects on neurological symptoms. We were particularly interested in assessing results in adolescents/adult cohorts. We used a cut-off age of 15 years to split participants into child vs. adolescent/adult cohorts. This is based on the suggestion that diets start to be relaxed and Phe levels begin to rise in mid-late adolescence (Berry et al., 2013; Crone et al., 2005; Walter et al., 2002; Walter & White, 2004). Studies including children (<15 years old), both alone and in combination with adults/adolescents, were grouped together under the heading 'mixed-age'.

Cognitive outcomes: Results and discussion

We reviewed 37 studies reporting cognitive outcomes. Twenty-three were conducted with mixed-age cohorts (15 were children only) and 14 with adults/adolescents. Tables 6.A1 and 6.A2 (in Appendix H) show parameters and outcomes in qualitative terms for individual studies. Both arithmetic and weighted averages across studies are reported.

Table 6.A1 shows results for studies where Phe was decreased. This was mainly through diet resumption (N=11) and, in a few cases, through Sapropterin or Pegvaliase treatment (N=2 and 2). Most studies (11/15) involved early-treated participants only. Study-length varied widely, between 1 and 209 weeks (mean=112, SD=68). Panel A shows results for mixed-age studies (N=9), and panel B for adult/adolescent studies (N=6)**.** Weighted Average Phe at baseline was **931** µmol/L for the mixed-age studies and **1246** µmol/L for the adult studies. After decrease, average Phe was **488** µmol/L for the mixed-age studies, and **562** µmol/L for the adult studies, with respective reductions in Phe of **387** µmol/L and **688** µmol/L. (arithmetic averages: 717 and 885). Overall, significantly more studies (13/15; 86.7%) showed a positive effect of decreasing Phe levels than showed no benefit ($\chi^2(1) = 8.1$, $p=.005$).

Table 6.A2 shows results for studies where Phe was increased either through diet discontinuation (N=15), or Phe loading (N=7). Panel a) shows results for mixed-age studies (N=14), panel b) for adult/adolescent studies (N=8). Again, most studies (17/22) involved early-treated participants only. Study-length varied from 0.01 to 209 weeks (Mean=101, SD=89). Weighted average Phe at baseline was **543** µmol/L for the mixed-age studies and **763** µmol/L for the adult studies. Average Phe after increase was **1327** µmol/L for the mixed-age studies and **1535** µmol/L for the adult/adolescent studies with respective increases in

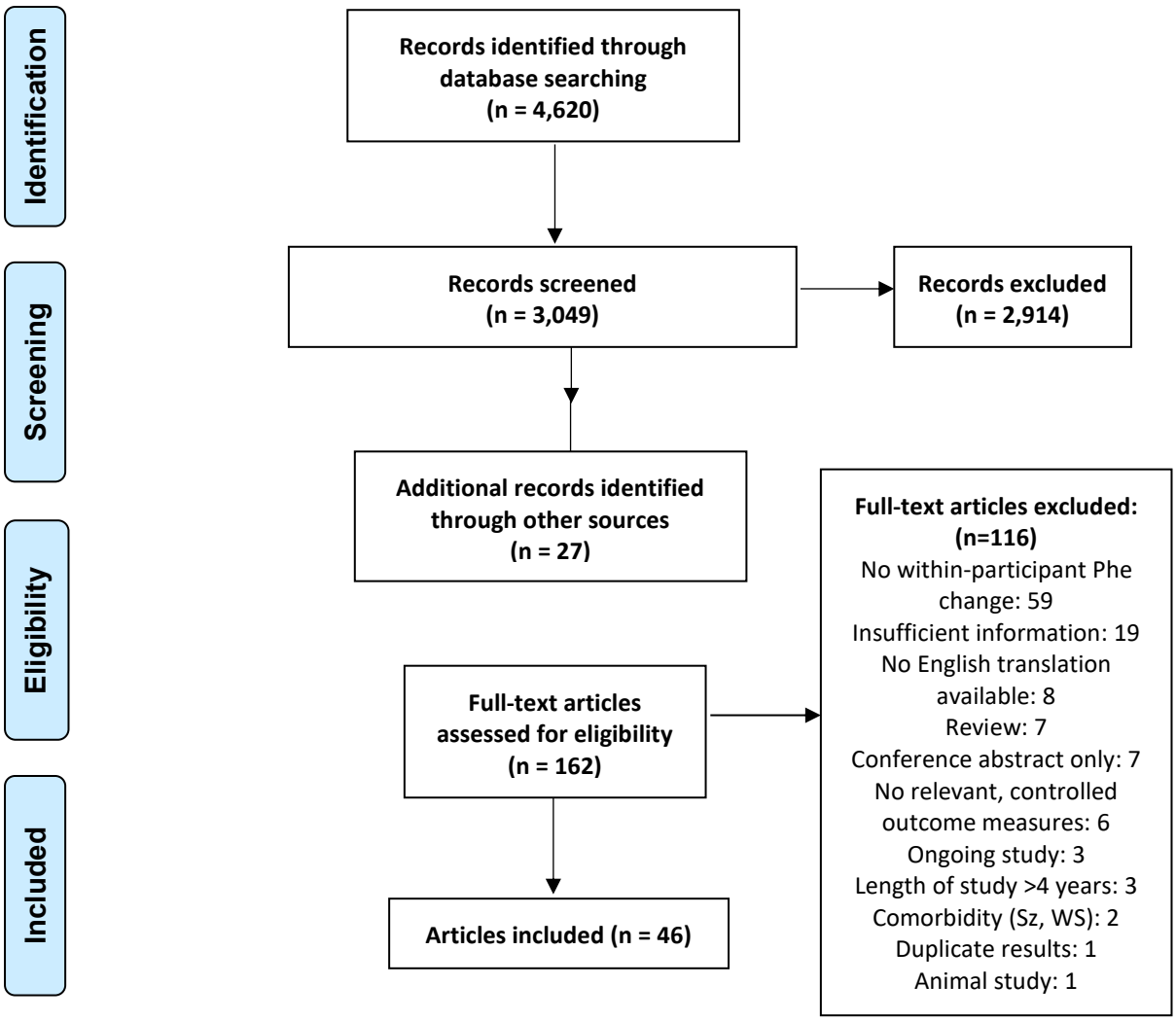


Figure 6.1. – Flowchart for systematic review of articles reporting cognitive, well-being, and/or neurophysiological outcomes of within-participant Phe manipulation; Sz=Schizophrenia, WS=West Syndrome.

Phe of **783** µmol/L and **772** µmol/L (arithmetic averages 918 and 710 µmol/L). Overall, most studies (14/22; 63.6%) showed a negative effect of increasing Phe levels (thus a positive effect of lower Phe levels), although this difference was not statistically significant ($\chi^2(1) = 1.6$, p=.20).

Considering all studies with cognitive outcomes, one can note that the higher Phe condition was well above levels recommended by both European and US guidelines, both in studies where Phe decreased and in those where it increased (but more so in the second case). Instead, lower Phe conditions for adult studies approached recommended levels at least according to European guidelines (**562** µmol/L and **763** µmol/L for studies decreasing and increasing Phe, when European guidelines recommend target levels of 600 µmol/L).

With few exceptions, differences in Phe across conditions were >**500** µmol/L. In studies showing a benefit with decreasing Phe, the smallest difference was **–208** µmol/L in a mixed cohort study (Burton et al., 2008) and **-668** µmol/L in an adult study (Thomas et al., 2018). Among studies where Phe was increased, the smallest difference in studies showing worse performance was **635** µmol/L among studies with children (reported by Holtzman et al., 1975) and **571** µmol/L among adult/adolescent studies (reported by ten Hoedt et al., 2011). This gives us a rough indication of the amount of difference in Phe that will result in a difference in cognitive outcomes.

Table 6.1 compares the number of studies showing vs. not showing a cognitive benefit when Phe is lower (using χ^2) and compares the respective parameters (Phe values, Phe-difference, length of study). Results are subdivided by age of cohort (here mixed-age studies were removed so that studies where participants were all <15 years old were compared with studies where participants were all >=15 years old) and by direction of Phe change (increase vs. decrease). Statistical comparisons of parameters were carried out only when there were >=5 studies in both the 'benefit' and 'no benefit' groups. However, we also report parameters for these subsets as they are informative on the range of values where cognitive effects were or were not demonstrated. We will do the same in the other tables showing results in the other domains.

When all studies were considered together, significantly more studies (and more participants) showed benefits than a null result (number of studies showing a benefit over total number, 27/37 = 73%, $\chi^2(1) = 7.8$, p=.005). The same was true when children and adult/adolescent studies were considered separately (rate of studies showing a benefit: children: 73%; adults: 79%). When studies where Phe was decreased/increased were considered separately, 87% and 62% of studies, respectively, showed a benefit, with only studies where Phe was decreased reaching significance. Studies showing a benefit numerically lasted longer, although the difference was not significant. This was true for all sub-comparisons considering weighted means. Also, studies showing a benefit had numerically a larger difference in Phe (significant only with unweighted means). This was true numerically for all sub-comparisons except for the comparison involving adult/adolescent studies where studies showing no benefit appear to have a larger difference in Phe. However, there were only three studies in this group (see Table 6.A2 panel B). They all involved an increase in Phe from baseline, but, perhaps notably, in 2/3 of these studies Phe, even at baseline, was well above what is recommended by current guidelines (1033 µmol/L and 1180 µmol/L compared with a recommendation of <600 µmol/L

by European guidelines, van Wegberg et al., 2017). This may indicate that negative effects of Phe do not increase linearly (i.e., the same difference in Phe does not have the same effect when Phe is higher overall). However, this is only speculative in the absence of more studies and corroborating evidence.

There was no difference in the proportion of studies showing a benefit when they involved children vs. adult/adolescents (χ^2= 0.11; p=.74). Equally, there was no difference in the parameters used by children and adult/adolescent studies showing a benefit both using unweighted means (study-length: t-test =-1.6; p=.13: Phe-difference t-test=1.8; p=0.9) and weighted means (study-length: F=1.6; p=.21; Phe-difference: F=2.0; p=.17). These results provide no evidence that the effects of Phe change at different ages, but one has to stress that they are null results, gathered post-hoc, and from otherwise very diverse studies. Binary logistic regression predicting benefit/no benefit using the Phe difference was significant using unweighted means (β=-.003; Wald 3.89; p=.05), but not with weighted means (because two large studies with average Phe differences dominate the results; β=-.005; Wald 0.32; p=.57). When adding study length as a predictor the significance of Phe difference in an unweighted analysis reduces to .07 (β=-.003; Wald 3.31; p=.07). Study length was not significant (β=.001; Wald .08; p=.78).

Meta-analyses. Twenty of the 37 cognitive studies provided quantitative outcomes allowing computation of effect sizes. A forest plot showing included studies, measures, and results is shown in Figure 6.2. Table 6.A3 (Appendix H) outlines the parameters of all included studies. We calculated models using assumed pre-post correlations of 0.2, 0.5, and 0.8. Results were always similar. The results presented here and in Figure 6.2 assume a correlation of 0.5. Overall, there was a positive and significant ES which tells us that cognition was better when Phe was lower. The pooled effect size, which captures the general effect of changing Phe, was 0.55, and clearly different from zero (95% confidence interval = 0.17– 0.94, z=2.8, p = .005).

Table 6.1. - Comparison of studies reporting a benefit vs. no benefit of lower Phe **on cognitive outcomes** in terms of their number (compared with χ^2), and treatment parameters (compared with univariate analysis of variance using either weighting or unweighted means, but keeping degree of freedom the same). 'Phe difference' is the difference in Phe mol/L achieved within studies. Note: values for Phe difference do not exactly match differences calculated from the table values because Phe post manipulation value was missed for one adult study decreasing Phe so difference could not be calculated. *significant at the .05 level, **significant at the .01 level

	Cognitive Studies (N=37)						
	Benefit			No benefit			χ^2/t-test/GLM
Overall	**Value**	**SD**	**Range**	**Value**	**SD**	**Range**	**p value**
N studies	**27**	-	-	**10**	-	-	**.005****
Sum participants	**758**	-	-	**134**	-	-	**<.001****
Mean N participants	**28**	*45*	1-178	**13**	*7*	4-24	.31
N studies with early treated ppts	**18**	-	-	**9**	-	-	.08
Study length (weeks) - weighted	**104**	*78*	0.01-209	**54**	*89*	1-209	.19
Study length (weeks) - arithmetic	**63**	*84*	0.01-209	**63**	*83*	1-209	.99
Low Phe - weighted	**563**	*122*	121-993	**588**	*301*	284-1180	.73
Low Phe - arithmetic	**562**	*197*	121-993	**626**	*288*	284-1180	.46
High Phe - weighted	**1230**	*283*	618-2365	**1210**	*539*	470-2170	.89
High Phe - arithmetic	**1448**	*400*	618-2365	**1206**	*481*	470-2170	.13
Phe Difference - weighted	**691**	*303*	208-1751	**622**	*260*	186-990	.58
Phe Difference - arithmetic	**916**	*431*	208-1751	**580**	*280*	186-990	**.03***
Studies with Children (<15years only)							
N studies	**11**	-	-	**4**	-	-	**.05***
Study length (weeks) - weighted	**133**	*78*	1-209	**113**	99	1-209	
Study length (weeks) - arithmetic	**94**	*96*	1-209	**118**	*107*	1-209	
Low Phe - weighted	**556**	*180*	358-993	**407**	*125*	284-666	
Low Phe - arithmetic	**529**	*196*	358-993	**443**	*161*	284-666	
High Phe - weighted	**1378**	*367*	618-2365	**960**	*404*	470-1544	
High Phe - arithmetic	**1590**	*472*	618-2365	**1012**	*463*	470-1544	
Phe Difference - weighted	**823**	*317*	260-1751	**554**	*295*	186-878	
Phe Difference - arithmetic	**1061**	*470*	260-1751	**570**	*319*	186-878	
Studies with Adults (>=15 years only)							
N studies	**11**	-	-	**3**	-	-	**.046***

Study length (weeks) - weighted	**89**	*62*	0.01-104	**14**	*29*	3-164	
Study length (weeks) - arithmetic	**33**	*54*	0.01-104	**37**	*58*	3-164	
Low Phe - weighted	**576**	*80*	121-863	**823**	*313*	536-1180	
Low Phe - arithmetic	**583**	*224*	121-863	**916**	*337*	536-1180	
High Phe - weighted	**1263**	*71*	1108-1600	**1653**	*408*	1337-2170	
High Phe - arithmetic	**1358**	*169*	1108-1600	**1618**	*478*	1337-2170	
Phe Difference - weighted	**690**	*72*	571-1443	**829**	*186*	315-990	
Phe Difference - arithmetic	**800**	*250*	571-1443	**702**	*348*	315-990	
Studies decreasing Phe							
N studies	**13**	-	-	**2**	-	-	**.005****
Study length (weeks) - weighted	**97**	*74*	1-209	**20**	*8*	9-26	
Study length (weeks) - arithmetic	**47**	*72*	1-209	**18**	*12*	9-26	
Low Phe - weighted	**544**	*84*	121-761	**521**	*151*	409-713	
Low Phe - arithmetic	**498**	*196*	121-761	**561**	*215*	409-713	
High Phe - weighted	**1147**	*248*	618-2365	**901**	*334*	653-1327	
High Phe - arithmetic	**1301**	*447*	618-2365	**990**	*477*	653-1327	
Phe Reduction - weighted	**622**	*232*	208-1751	**380**	*183*	244-614	
Phe Reduction - arithmetic	**846**	*488*	208-1751	**429**	*262*	244-614	
Studies increasing Phe							
N studies	**14**	-	-	**8**	-	-	.20
Study length (weeks) - weighted	**123**	*83*	1-209	**60**	*86*	.01-209	.13
Study length (weeks) - arithmetic	**75**	*92*	1-209	**75**	*90*	.01-209	.99
Low Phe - weighted	**604**	*168*	381-993	**599**	*287*	284-1180	.97
Low Phe - arithmetic	**613**	*190*	381-993	**642**	*314*	284-1180	.82
High Phe - weighted	**1450**	*230*	1200-2192	**1261**	*494*	470-2170	.36
High Phe - arithmetic	**1584**	*306*	1200-2192	**1260**	*498*	470-2170	.13
Phe Increase- weighted	**846**	*239*	571-1729	**662**	*268*	186-990	.15
Phe Increase - arithmetic	**971**	*389*	571-1729	**618**	*288*	186-990	.03*

We also ran separate analyses using age, Phe difference and study length as moderators. The best model had an interaction between *Study-length* and *Phe-difference* and a main effect of *Age*. Studies with larger Phe differences produced larger effects and this difference was clearest in longer studies (interaction: $G^2(1) = 13.4$, p=.0002). The best model was clearly better than a model that did not have any terms with Phe differences and included only *Study Length* + *Age,* showing the importance of Phe differences ($G^2(2) = 13.6$, p=.001). The main effect of *Age* was significant ($G^2(1) = 12.6$, p = .0004). The ES was also significant when adult/adolescent studies were considered on their own (ES= .57, z=2.4, p=.01;).

To check for bias due to studies with fewer participants we plotted a funnel plot (Figure 6.3), that displays ES against sample size. Small N studies did not cluster on the right of the mean, which would be the hallmark of bias introduced by small N studies. An Egger's test, which uses number of participants as a moderator of effect size, was not significant (z=0.52, p=.61), showing that participant numbers did not influence ES estimates. There was no evidence, therefore, that smaller studies with null/opposite effects had been excluded because of a publication bias, nor that the overall estimate of ES was biased by small N studies.

Finally, to check whether the overall effect size changed as studies with smaller N were added to the model, we ran a sensitivity analysis. The four largest studies were used for an initial estimate, and then studies were added in sets of 4. At each stage, a new cumulative estimate of effect size was calculated. Figure 4 shows the results, assuming a correlation between 'pre' and 'post' conditions of 0.5. There was no indication that the estimate of the effect size was distorted by smaller studies and the overall estimate remained stable throughout.

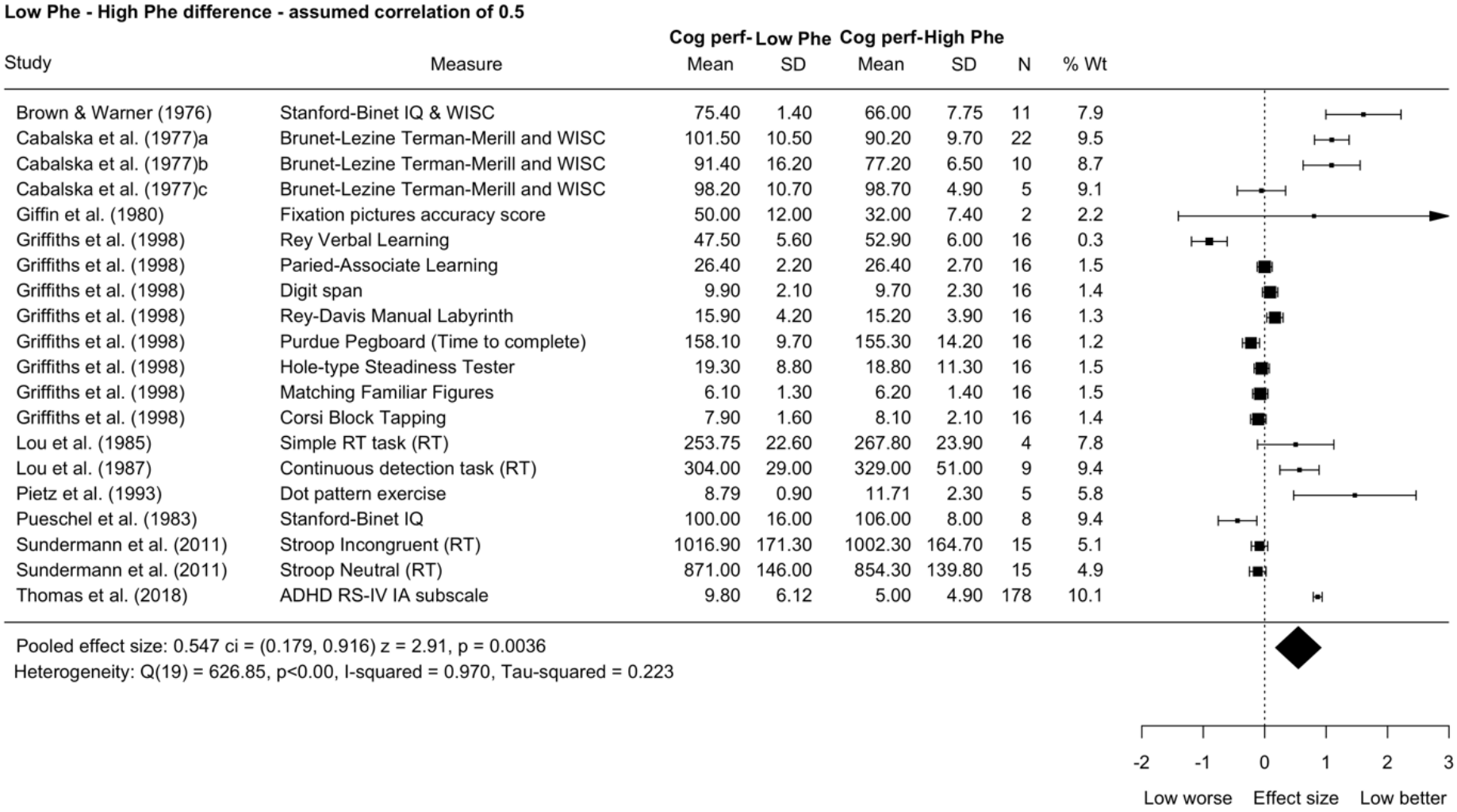


Figure 6.2. – Forest plot of studies included in the meta-analyses with an assumed correlation between studies of 0.5. Mean and SD for results on cognitive measures in low Phe and high Phe conditions are presented, along with number of participants (N) and study weightings (% Wt). Effect sizes for individual studies are indicated by the squares on the horizontal axis. Horizontal bars indicate 95% confidence intervals for each effect size. The black diamond indicates the overall point estimate and confidence interval across all studies.

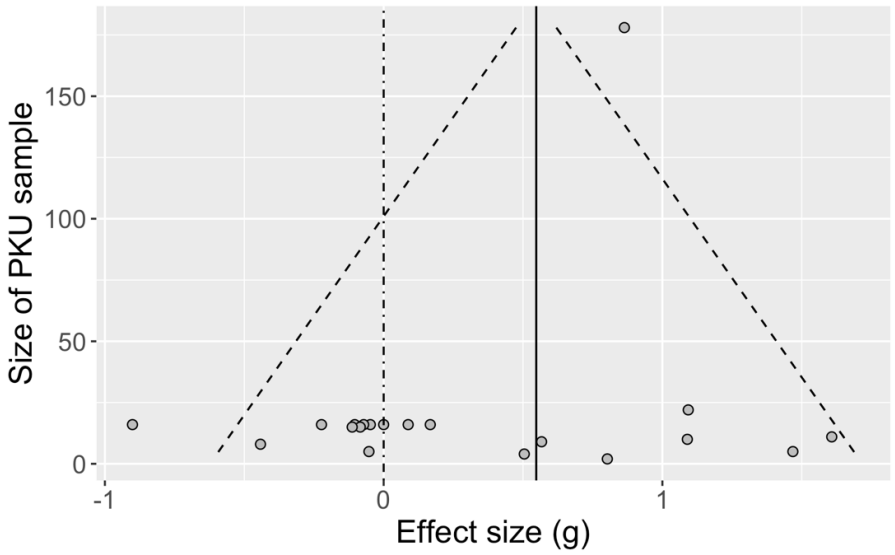


Figure 6.3. – Funnel plot of effect sizes as a function of sample size. The dash-dot vertical line marks the position of no benefit. The solid vertical line shows the overall effect size returned by the meta-analysis. The dashed diagonal lines are approximate 95% confidence limits around the overall effect size. These connect the confidence limit width from the study with the highest N to the confidence limits from the average of the 5 studies with the lowest N.

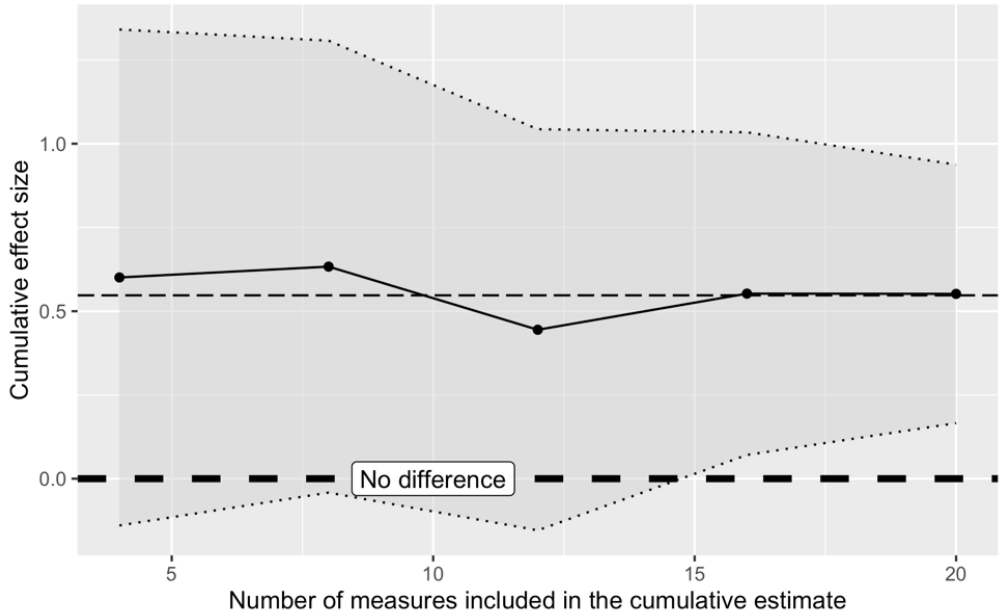


Figure 6.4. – Cumulative effect size estimates when studies with fewer participants are added across five bins.

Well-being outcomes: results and discussion

We reviewed 22 studies reporting well-being outcomes. Eight were conducted with mixed-age cohorts (5 of which were with children only) and 14 with adults/adolescents. Tables 6.A4 and 6.A5 in the Appendix H show results when Phe was respectively decreased and increased. Both weighted and arithmetic means are reported.

Table 6.A4 shows the results for studies where Phe was decreased. This was mainly through diet resumption (N=15) and, in two cases, through Sapropterin or Pegvaliase treatment (N=1 and 1). Most studies (12/17) involved late-treated participants. Study-length was between 8 and 209 weeks (mean=109, SD=58). Panel A shows results for mixed-age studies (N=5), and panel B for adult studies (N=12)**.** Weighted average Phe at baseline was **1011** µmol/L for the mixed-age studies and **1365** µmol/L for the adult studies. After decrease, it was **619** µmol/L for the mixed-age studies and **588 µ**mol/L for the adult studies, with respective reductions in Phe of **406** µmol/L and **783** µmol/L.

Table 6.A5 shows the results for studies where Phe was increased, either through diet discontinuation (N=4) or Pegvaliase discontinuation (N=1). Given the small number of studies, all age cohorts are presented together. 3/5 studies involved early-treated participants only. Study-length varied from 0.9 to 37 weeks (Mean=11, SD=12). Weighted average Phe at baseline was **536** µmol/L; Phe after increase was **1354** µmol/L, with an average increase of **818** µmol/L.

More studies (3/5; 60%) showed a negative effect of increasing Phe levels (thus a positive effect of lower Phe levels) although this is not significant given the small number of studies ($\chi^2(1) = 0.2$, p=.66). Notably, one study reported results in the opposite direction, with reduced emotional problems with diet relaxation. There were large variations in the size of Phe differences.

Considering the difference in Phe achieved by studies showing a benefit, the smallest difference in studies decreasing Phe was **-188** µmol/L in a mixed-age cohort (Douglas et al., 2013) and **–297** µmol/L in adults (Williams 1998). In studies showing decrements with increasing Phe the smallest difference was **1611** µmol/L in children (Leuzzi et al., 1997) and **571** µmol/L in adults (ten-Hoedt et al., 2011).

Table 6.2. - Comparison of studies reporting a benefit vs. no benefit of lower Phe **on well-being**, in terms of their number and treatment parameters. 'Phe difference' is the difference in Phe mol/L achieved within studies. Note: values for Phe difference do not exactly match differences calculated from the table values because Phe post manipulation value was missed for one adult study and one child studies decreasing Phe so difference could not be calculated. *significant at the .05 level, **significant at the .01 level

	Well-Being Studies (N=22)					
	Benefit			No benefit		
Overall	**Value**	**SD**	**Range**	**Value**	**SD**	**Range**
N studies	**19**	-	-	**3**	-	-
Sum participants	**451**	-	-	**36**	-	-
Mean N participants	**103**	*66*	1-178	**18**	*9*	5-24
N studies with early treated ppts	**7**	-	-	**1**	-	-
Length of study (weeks) - weighted	**107**	*59*	0.9-209	**14**	*12*	8-37
Length of study (weeks) - arithmetic	**46**	*50*	0.9-209	**18**	*17*	8-37
Low Phe - weighted	**595**	*67*	121-940	**490**	*67*	381-536
Low Phe - arithmetic	**611**	*183*	121-940	**447**	*80*	381-536
High Phe - weighted	**1290**	*348*	690-5448	**1343**	*390*	793-2143
High Phe - arithmetic	**1753**	*1161*	690-5448	**1424**	*679*	793-2143
Phe Difference - weighted	**758**	*354*	188-4722	**852**	*385*	412-1719
Phe Difference- arithmetic	**1234**	*1190*	188-4722	**977**	*671*	412-1719
Studies with Children (<15years only)						
N studies	**4**	-	-	**1**	-	-
Length of study (weeks) - weighted	**199**	*45*	0.9-209	**37**	-	-
Length of study (weeks) - arithmetic	**61**	*100*	0.9-209	**37**		
Low Phe - weighted	**601**	*144*	*476-726*	**381**	-	-
Low Phe - arithmetic	**643**	*144*	*476-726*	**381**		
High Phe - weighted	**1153**	*630*	1027-5448	**793**	-	-
High Phe - arithmetic	**3155**	*1978*	1027-5448	**793**		
Phe Difference - weighted	**2819**	*1506*	1611-4722	**412**	-	-
Phe Difference- arithmetic	**3221**	*1558*	1611-4722	**412**		
Studies with Adults (>=15 years only)						

N studies	**15**	-	-	**2**	-	-
Length of study (weeks) - weighted	**88**	*33*	4-108	**8**	-	-
Length of study (weeks) - arithmetic	**42**	*28*	4-108	**8**	-	-
Low Phe - weighted	**593**	*59*	121-940	**517**	*43*	424-536
Low Phe - arithmetic	**589**	*209*	121-940	**480**	*79*	424-536
High Phe - weighted	**1349**	*205*	1105-1705	**1476**	*310*	1337-2143
High Phe - arithmetic	**1401**	*234*	1105-1705	**1740**	*570*	1337-2143
Phe Difference - weighted	**763**	*214*	297-1443	**959**	*353*	801-1719
Phe Difference- arithmetic	**839**	*375*	297-1443	**1260**	*650*	801-1719
Studies decreasing Phe						
N studies	**16**	-	-	**1**	-	-
Length of study (weeks) - weighted	**111**	*57*	11-209	**8**	-	-
Length of study (weeks) - arithmetic	**55**	*50*	11-209	**8**		
Low Phe - weighted	**594**	*67*	121-940	**424**	-	-
Low Phe - arithmetic	**609**	*196*	121-940	**424**		
High Phe - weighted	**1278**	*284*	690-4056	**2143**	-	-
High Phe - arithmetic	**1520**	*759*	690-4056	**2143**		
Phe Reduction - weighted	**747**	*280*	188-3330	**1719**	-	-
Phe Reduction- arithmetic	**988**	*801*	188-3330	**1719**		
Studies increasing Phe						
N studies	**3**	-	-	**2**	-	-
Length of study (weeks) - weighted	**3**	*1*	0.9-4	**15**	*12*	8-37
Length of study (weeks) - arithmetic	**2**	*2*	0.9-4	**23**	*21*	8-37
Low Phe - weighted	**627**	*74*	476-726	**501**	*66*	381-536
Low Phe - arithmetic	**617**	*128*	476-726	**459**	*110*	381-536
High Phe - weighted	**1717**	*1222*	1220-5448	**1214**	*231*	793-1337
High Phe - arithmetic	**2918**	*2233*	1220-5448	**1065**	*385*	793-1337
Phe Increase - weighted	**1090**	*1212*	571-4722	**713**	*165*	412-801
Phe Increase arithmetic	**2301**	*2160*	571-4722	**606**	*275*	412-801

Table 6.2 compares numbers and parameters of studies reporting better well-being with lower Phe with those reporting no benefit. For the age comparison, mixed-age studies were removed from analyses. When all studies were considered together, significantly more studies (and more participants) showed benefits than a null result: number of studies showing a benefit over total number 19/22 = 86% ($\chi^2(1)=11.6$, p=.001). There were more studies showing benefits for all sub-comparisons. However, because of small numbers, we only assessed separately the rate of studies showing benefits for adult/adolescent studies: 15/17= 88% ($\chi^2(1) = 9.9$, p=.002) and for studies decreasing Phe: 16/17; 94.1% ($\chi^2(1) =13.2$, p=<.001). Parameters could not be statistically compared because of the low number of studies showing no benefit. In numerical terms, there were no notable differences between high or low Phe levels or magnitude of Phe differences between studies with benefit vs. no benefit. Length of study, however, was much higher in studies reporting a benefit (Mean=80, SD=65) compared to those reporting no benefit (Mean=14, SD=12).

Neurophysiological outcomes: results and discussion

We reviewed 14 studies reporting neurophysiological outcomes after a Phe change. Ten were conducted with mixed-age cohorts (5 of which were with children only) and three with adults. Measures were obtained using MRI (N=5), fMRI (N=1), visual field examination (N=1), EEG (N=5), diffusion tensor imaging (DTI; N=2) and automated fibre-tract quantification (AFQ; N=1). We included Jaulent et al.'s (2020) review as a single study, summarising parameters and outcomes for the eight new participants and 22 pre-existing case studies. This study considered neurological symptoms as well as MRI results. Tables 6.A6 and 6.A7 in Appendix H show results when Phe was respectively decreased and increased. Both weighted and arithmetic means are reported.

Table 6.A6 shows parameters and effects for studies looking at effects of decreasing Phe through diet resumption (N=4) and Sapropterin treatment (N=2). Most studies (5/6) involved early-treated participants only. Study-length was between 26 and 126 weeks (mean=39, SD=15). Weighted Phe at baseline was **1240** µmol/L; after decrease it was **770** µmol/L, with an average reduction of **463** µmol/L.

Table 6.3. - Comparison of studies reporting a benefit vs. no benefit of lower Phe on neurophysiological outcomes, in terms of their number and treatment parameters. 'Phe difference' is the difference in Phe mol/L achieved within studies. Note: values for Phe difference do not exactly match differences calculated from the table values because Phe post manipulation value was missed for one adult study decreasing Phe so difference could not be calculated. *significant at the .05 level, **significant at the .01 level. + = 2 studies reported a benefit of lower Phe levels in adult/adolescents, but only 1 provided details of Phe levels post-manipulation (reported in table)

	Neurophysiological Studies (N=14)						
	Benefit			No benefit			T-test/X2
Overall	**Value**	**SD**	**Range**	**Value**	**SD**	**Range**	**p value**
N studies	**9**	-	-	**5**	-	-	.29
Sum participants	**105**	-	-	**50**	-	-	**<.001****
Mean N participants	**17**	*7*	1-25	**15**	*6.3*	2-21	.72
N studies with early treated ppts	**8**	-	-	**5**	-	-	.41
Length of study (weeks) - weighted	**94**	*87*	0.01-209	**31**	*29*	0.01-104	.19
Length of study (weeks) - arithmetic	**99**	*89*	0.01-209	**38**	*43*	0.01-104	.18
Low Phe - weighted	**588**	*222*	300-969	**982**	*198*	476-1180	**.01***
Low Phe - arithmetic	**490**	*196*	300-969	**871**	*290*	476-1180	**.01***
High Phe - weighted	**1211**	*331*	653-1613	**1635**	*390*	1300-2170	.09
High Phe - arithmetic	**1121**	*356*	653-1613	**1690**	*412*	1300-2170	**.02***
Phe Difference - weighted	**620**	*214*	244-1000	**653**	*378*	300-1611	.85
Phe Difference- arithmetic	**607**	*268*	244-1000	**819**	*544*	300-1611	.36
Studies with Children (<15years only)							
N studies	**3**	-	-	**2**	-	-	-
Length of study (weeks) - weighted	**209**	*0*	*209*	**42**	*22*	*1-52*	-
Length of study (weeks) - arithmetic	**209**	*0*	*209*	**27**	*36*	*1-52*	
Low Phe - weighted	**476**	*33*	400-509	**628**	*80*	476-666	-
Low Phe - arithmetic	**462**	*56*	400-509	**571**	*134*	476-666	
High Phe - weighted	**1244**	*167*	835-1362	**1653**	*229*	1544-2087	-
High Phe - arithmetic	**1160**	*284*	835-1362	**1816**	*384*	1544-2087	
Phe Difference - weighted	**768**	*135*	435-853	**1025**	*309*	878-1611	-
Phe Difference- arithmetic	**698**	*229*	435-853	**1245**	*518*	878-1611	
Studies with Adults (>=15 years only)							

N studies	**2**[+]	-	-	**2**	-	-	-
Length of study (weeks) - weighted	**55**	*15*	*52-126*	**22**	*44*	*0.01-104*	-
Length of study (weeks) - arithmetic	**89**	*52*	*52-126*	**52**	*74*	*0.01-104*	
Low Phe - weighted	**943**	*131*	300-969	**1149**	*62*	1033-1180	-
Low Phe - arithmetic	**635**	*473*	300-969	**1107**	*104*	1033-1180	
High Phe - weighted	**1613**	-	-	**1997**	*344*	1348-2170	-
High Phe - arithmetic	**1613**	-	-	**1759**	*581*	1348-2170	
Phe Difference - weighted	**644**	-	-	**848**	*283*	315-990	-
Phe Difference- arithmetic	**644**	-	-	**653**	*477*	315-990	
Studies decreasing Phe							
N studies	**5**	-	-	**1**	-	-	-
Length of study (weeks) - weighted	**42**	*17*	26-126	**33**	-	-	-
Length of study (weeks) - arithmetic	**53**	*42*	26-126	**33**			
Low Phe - weighted	**675**	*292*	300-969	**1000**	-	-	-
Low Phe - arithmetic	**491**	*271*	300-969	**1000**			
High Phe - weighted	**1214**	*454*	653-1613	**1300**	-	-	-
High Phe - arithmetic	**1090**	*489*	653-1613	**1300**			
Phe Decrease- weighted	**531**	*238*	244-1000	**300**	-	-	-
Phe Decrease- arithmetic	**552**	*346*	244-1000	**300**			
Studies increasing Phe							
N studies	**4**	-	-	**4**	-	-	-
Length of study (weeks) - weighted	**143**	*98*	0.01-209	**29**	*38*	0.01-104	-
Length of study (weeks) - arithmetic	**157**	*105*	0.01-209	**39**	*50*	0.01-104	
Low Phe - weighted	**506**	*53*	400-572	**969**	*261*	476-1180	-
Low Phe - arithmetic	**490**	*71*	400-572	**839**	*324*	476-1180	
High Phe - weighted	**1208**	*148*	835-1362	**1878**	*348*	1348-2170	-
High Phe - arithmetic	**1152**	*233*	835-1362	**1787**	*404*	1348-2170	
Phe Increase- weighted	**702**	*149*	435-853	**909**	*299*	315-1611	-
Phe Increase- arithmetic	**663**	*200*	435-853	**949**	*531*	315-1611	

Table 6.A7 shows the results for studies where Phe was increased either through diet discontinuation (N=6) or Phe loading (N=2). All studies involved early-treated participants only. Study-length varied from 0.01 to 209 weeks (Mean=103, SD=99). Weighted average Phe at baseline was **668** µmol/L; after increase, it was **1442** µmol/L, with an average increase of **774** µmol/L. Overall, 4/8 (50%) of studies showed a negative effect of increasing Phe levels (thus a positive effect of lower Phe levels).

Among the studies which demonstrated a benefit, the smallest difference in Phe for studies decreasing Phe was **244** µmol/L in a mixed age study (White et al., 2013). In studies increasing Phe it was **435 µmol**/L in a child study (Cabalska et al., 1977c; – Group C with mild PKU or Hyperphenylalanemia).

Table 6.3 compares the number of studies and the parameters of studies reporting better neurophysiological health with lower Phe levels with those reporting no benefit. For the age comparison, mixed-age studies were removed from analyses. Numerically, all comparisons showed more studies reporting benefit. However, even considering all studies together, the difference did not reach significance: 64% ($\chi^2(1) = 1.1$, p=.29). The only marginal comparison was with studies decreasing Phe. 5/6 studies showed a benefit (83.3%; $\chi^2(1) = 2.7$, p=.10). Statistical comparisons of parameters were not carried out because of the low number of studies. At least numerically, studies reporting a benefit were longer, both when considered all together and in the individual sub-comparisons. Studies with a benefit did not show a larger difference in Phe (if anything, results were in the opposite direction).

Discussion

The current literature surrounding the effects of metabolic control on outcomes in PKU is strongly led by between-participant reports. These studies demonstrate the impact of Phe levels on cognition in children (see Canton et al., 2019; Christ et al., 2010; Hofman et al., 2018) and in adolescents and adults (Aitkenhead et al., 2021; Jahja, van Spronsen et al., 2017; Romani et al., 2017; 2019). However, they suffer from important limitations related to the fact that individuals with the best control may be those with better education and socio-economic status (variables which may also affect cognitive abilities) and with the best general health (which would affect brain health, mental well-being, and cognitive health). Studies with adults/adolescents are particularly problematic since metabolic control at these ages is strongly correlated with metabolic control at younger ages. Therefore, the correlations seen between current Phe and cognitive abilities, mental health, or brain

measures may be due to collinearity with Phe levels during childhood. The strongest evidence to demonstrate an impact of Phe on outcomes, therefore, would come from within-participant studies, where outcomes are measured in the same participants at different times after the level of Phe has been altered. These studies are few, but a systematic review can allow us to collate results and reach firmer conclusions. In this paper, we reviewed studies investigating the effects of changes to Phe that take place within-participants, either through dietary changes or pharmacological interventions. We considered effects on cognitive, well-being, and neurophysiological measures. Where possible, we considered adult/adolescent studies, and studies with children or mixed-age cohorts separately. Because of difficulties in computing quantitative effects sizes, most of our analyses have involved a qualitative classification of studies showing vs. not showing a benefit when Phe levels were lower.

Our first, most important result was that, across domains, significantly more studies reported benefits of lower Phe than no benefit. This was true for cognition (27/37=73%), and well-being (19/22=86%), with a non-significant result in the same direction for neurophysiological /neurological outcomes (9/14 studies=64%).

Cognitive benefits of lower Phe were reported in the domains of attention, executive function, and speed of processing, whilst more mixed outcomes were reported for general intelligence and working memory. A meta-analysis of 20 quantitative cognitive measures across 12 articles/publications demonstrated an overall estimated ES of 0.55, which was highly significant and indicates a sizeable effect. To illustrate what this means, this corresponds to a drop of 21 positions in a normally distributed population of 100 individuals from a starting position at the mean. These results support, and extend, reports from between-participant studies (for adults see Aitkenhead et al., 2021 Romani et al., 2017; for children see Huijbregts et al., 2002; Jahja, van Spronsen et al., 2017; Schmidt et al. 1996).

In terms of mental well-being, benefits of lower Phe levels were reported on measures of anxiety, depression, aggression, and hyperactivity. The great majority of studies reported benefits. This is consistent with between-participant and correlational studies which have found that poor metabolic control is associated with more behavioural difficulties and personality problems in children and adults with PKU (Burton et al., 2013; Holtzman et al., 1986; Jahja, Huijbregts et al., 2017; Koch et al., 2002; Smith et al., 1988). It is important to consider, however, that most assessments were carried out through questionnaires/self-report or observations. Our reviewed studies involved caregiver reports (e.g., by parents, teachers, clinicians; N=6), self-report via questionnaire or interview (N=10), and systematic observations (N=9; note some studies used multiple methods). Among the

observational studies, four relied on observations by nursing staff/key workers at residential institutions (Giffin et al., 1980; Harper & Reid, 1987; Hoskin et al., 1992; Lee et al., 2009), one on observations by "trained inconspicuous observers" (Marholin et al., 1978), one on observations made by teachers, parents, play-leaders, psychiatrists and/or nurses (Leuzzi et al., 1997), and three did not specify how behavioural observations were made (Anwar et al., 2013; Riordan et al., 1955 - conditions A & B; Dion et al., 2001). These measures are more subject to bias and expectation than quantitative, controlled, performance measures. Only three studies were carried out with measures taken under double-blind conditions. Of these, one reported some positive outcomes following diet initiation in late-treated AwPKU (Lee et al., 2009), one reported negative outcomes following Phe loading in early-treated AwPKU (ten Hoedt et al., 2011), and one reported no meaningful changes in a cohort of untreated adults (Marholin et al., 1978). Therefore, results on well-being should be interpreted with more caution. Not all between-participant studies have reported difficulties with well-being in AwPKU or correlations with Phe levels (e.g., Aitkenhead et al. 2021; Channon et al., 2007; Romani et al., 2021). In fact, some studies have highlighted that following a PKU diet can have not only positive, but also negative effects on well-being since the diet is unsociable and time-demanding (e.g., Bik-Multanowski et al., 2008; Palermo et al., 2020; Pietz et al., 1997).

Finally, neurophysiological/neurological benefits of lower Phe levels were reported in terms of reduced white matter abnormalities, increased diffusivity, reduced EEG abnormalities, improved motor responses, and reduced neurological symptoms (including upper limb tremors, brisk tendon reflexes, dysarthria, and blurred/loss of vision). This is consistent with findings from between-participant and correlational studies which have reported that white matter lesions and abnormalities in a number of brain areas are associated with elevated Phe levels (Cleary et al., 1994; Hawks et al., 2019; Hellewell et al., 2021; Mastrangelo et al., 2015; Nardecchia et al., 2015; Thompson et al., 1993).

A secondary aim of our review was to gather information on the metabolic parameters which elicited a change in outcomes. We had some indication that benefits were easier to demonstrate when differences in Phe between conditions were larger and when this difference was maintained for longer, but effects were not always consistent. With cognitive outcomes: 1. studies reporting a benefit of lower Phe (compared to no benefit) showed larger Phe differences between conditions, 2. Phe difference was a predictor of benefit/no-benefit in a binary logistic regression (but only with unweighted means) and 3. Phe difference predicted the size of effect in the meta-analysis. The same effects were not shown with well-being and neurophysiological studies, but this could be due to the small

number of studies included in this comparison. We could only provide a rough estimate for the size of the Phe difference which elicited benefit. We can note that it was =>600 μmol/L in all three domains. The average weighted difference was **691** μmol/L (SD=278) for cognitive outcomes, **758** μmol/L (SD=354) for well-being outcomes, and **620** μmol/L (SD=214) for neurophysiological/neurological outcomes. Higher Phe levels for studies showing a benefit were also similar across domains: 1230 μmol/L (SD=278), 1290 μmol/L (SD=348) and 1211 μmol/L (SD =356) for cognitive, well-being and neurophysiological outcomes respectively, showing the negative impact of Phe above 1000 μmol/L

In terms of length of studies, across domains and sub-comparisons, studies demonstrating a benefit were numerically longer than studies showing no benefit (weighted average length in weeks, cognitive: 104 (SD=78) vs. 54 (SD=89); well-being: 107 (SD=59) vs. 14 (SD=12); neurophysiological: 94 (SD=87) vs. 31 (SD=29)). This was true for all sub-comparisons (except for studies looking at the effects of increasing Phe on well-being outcomes) although differences were either not statistically significant, or there were not enough results to assess significance.

These results give us an indication of the level of Phe difference that may be needed to demonstrate benefits. However, studies showed great variability and we really do not know whether smaller Phe reductions, or reductions occurring over a smaller time frame, would also show benefits. Positive differences have been reported in a few cases with Phe differences <400 μmol/L (cognition: Burton et al., 2015, well-being: Bik-Multanowski et al., 2008; Douglas et al., 2013). Some studies did compare performance across a span of only a few weeks, days, or hours, but they were very few (see for cognition: Hujibregts et al., 2002; Krause et al., 1985; for well-being: ten Hoedt et al., 2011; for neurophysiology: Leuzzi et al., 2014).

A final aim of our review was to assess outcomes separately for studies involving older PKU cohorts. Whilst the importance of maintaining low-Phe levels in childhood is undisputed, the importance of metabolic control in adulthood remains contested, with many adults and adolescents choosing to relax a strict diet after early adolescence (Berry et al., 2013; Crone et al., 2005; Walter et al., 2002; Walter & White, 2004). Our results show that poor metabolic control, even in later life, can have a measureable negative impact. Studies with participants =>15 years old demonstrated better cognition with lower Phe levels (11/14 = 79% of studies, with the great majority of studies involving early-treated participants) as well as better well-being (15/17= 88% of studies). There were not enough studies focusing

solely on older cohorts to assess effects on neurophysiological outcomes. Our meta-analysis, assessing the effect of changing Phe levels on cognition, demonstrated a substantial and significant ES (0.57) when studies with adolescents/adults were considered separately. There was no indication that there were differences in the parameters used by adult/adolescent studies compared to child studies nor that adult/adolescent studies required a higher difference in Phe to show a benefit. There is too much variability between studies, however, to reach any firm conclusion. In adult/adolescent studies, cognition and well-being were detrimentally impacted when levels rose above above 1200 µmol/L. This does not mean, however, that lower levels are safe. There is a lack of studies with Phe levels in a lower range. Similarly, our results show that it is possible to undo some of damage caused by previously high levels of Phe by decreasing Phe. The extent of these reversals remains unclear, however, and it is unclear whether they can be achieved even when Phe has remained elevated for extensive periods of time.

Limitations. A main limitation in our review is the low number of studies reporting outcomes in a quantitative way, such that they could be effectively accrued using meta-analysis. We had to limit most of our analyses to a qualitative and binary outcome: whether or not the study showed benefits with lower Phe. Due to the paucity of within-participant studies in this field, we did not put a date limit on our literature search. The earliest studies included in this review were conducted in 1954, with a number of studies also conducted in the 60s and 70s. These studies often reported outcomes in qualitative terms, and studies reporting quantitative results often had only results in term of IQ. Results on a more comprehensive set of cognitive abilities would, therefore, be desirable. Studies assessing effects on well-being often included late-treated participants (14/22=64%) limiting their generalisability to early-treated participants that now constitute the great majority of the PKU population (although 7/8 with early treated participants did indicate a positive effect of lowering Phe). Moreover, well-being outcomes were generally <u>not</u> assessed blind to the dietary status of the participants. This may inflate positive results where behaviour and mood are more positively judged when Phe is lower, consistent with expectations. This is less of a concern for cognitive studies where outcomes are evaluated in terms of test performance. Lastly, only very few studies assessed neurophysiological changes in early-treated AwPKU after changes in Phe. These studies are important and hopefully more will be carried out in the future.

Conclusions. Our review showed significant benefits of lower Phe when Phe levels are changed within participants. There is some variability in the results reported by the literature, but the great majority of the reviewed studies reported benefits rather than null

results and differences in numbers were highly statistically significant for cognitive and well-being outcomes. Results were also in the same direction, although not statistically significant, for neurophysiological outcomes. Moreover, meta-analyses comparing cognitive performance with lower vs. higher Phe showed a highly significant effect of moderate size (.55). Importantly, the same results were obtained when cohorts of adults/adolescents were considered separately. These results add to consistent results from between-participant studies without having the same limitations. In between-participant studies, cognitive differences between groups with higher vs. lower Phe can potentially be due to differences in education, socio-economic status, and the underlying cognitive abilities of the participants and their families. These confounding influences are eliminated in studies where the same participants are assessed in conditions when Phe has changed.

Our conclusions are more limited in terms of the parameters necessary to see benefits (e.g., size of Phe difference, length of study). We only had some indication that the size of the Phe difference and the duration that new Phe levels are maintained affected outcomes (in particular cognitive outcomes). In the studies we reviewed, Phe changed substantially between low and high conditions (in general >500 μmol) and most of the studies had long durations, often spanning months. Thus we have no, or very limited, information about the effect of Phe changes which are smaller and/or are achieved and maintained within a smaller time span. Our review strongly highlights the need for studies that manipulate these parameters and evaluate effects, not only in IQ, but also in other, possibly more sensitive, tasks. This is important so that future intervention studies can provide a better understanding of the minimum metabolic changes that need to be achieved to see benefits. Our review also highlights the need for such studies to understand the lower time limit for a change in Phe to result in a meaningful change in outcomes.

Chapter 7: General Discussion

Aims of thesis

This thesis had two key aims. The first was to better understand the potential interactions between increasing age and early-treated PKU in later adulthood, primarily with regards to cognition, but also in terms of quality of life. The second was to investigate the extent to which cognitive, well-being, neurophysiological and neuropathological impairments caused by PKU are directly impacted by metabolic control throughout the lifespan. The first aim was addressed through empirical investigations in Chapters 3, 4 and 5. Chapters 3 and 4 sought to increase our understanding of how PKU is likely to be affected by increasing age, in terms of overlapping cognitive impairments and neurological deterioration caused by the natural ageing process and by excess Phe in PKU. Chapter 5, meanwhile, considered how cognition and well-being in early-treated AwPKU is impacted in middle-age compared to young adulthood, further investigating the potential impact of increasing age on outcomes. The second aim was addressed through a systematic review and meta-analysis in Chapter 6, in which the direct impact of changing Phe levels on individuals' cognition, well-being, and brain health, as measured through within-subject studies, was explored.

This General Discussion will first summarise the rationales and key findings of each chapter. It will then discuss the overall implications of findings presented in these chapters, as well as considering their limitations and suggesting future directions for research in this field. Finally, the chapter will culminate in an overall Conclusion.

Chapter summaries

Chapters 1 and 2: Introduction and Methodology.

The first chapter of this thesis outlined the existing literature surrounding cognitive, well-being, and neurological impairments in early-treated PKU. Outcomes in all three of these domains in early-treated adults and children with PKU were explored, as well as associations between metabolic control and outcomes in these domains during different life stages. Chapter 2 then laid out the comprehensive cognitive assessment battery used throughout this thesis to investigate the impact of ageing and PKU on cognition across a number of different domains. Tasks and delivery methods used in face-to-face and remote assessment conditions were described in detail.

Chapter 3. Investigating the Impacts of PKU and Healthy Ageing on Cognition

This chapter considered performance of healthy older controls (aged 53-88) on a cognitive assessment battery previously delivered to healthy younger controls and young AwPKU (aged 18-41). Due to similarities in white matter abnormalities and neurotransmitter depletions caused by natural ageing and by early-treated PKU, we hypothesised that cognitive impairments in these two populations would overlap. In particular, we expected impairments in the domains of speed of processing, executive function, and visuo-motor coordination to be apparent in both populations. Meanwhile, differences in performance between the two populations were expected in the domain of long-term memory and learning, as impairments in this domain are associated with hippocampal neurotransmitter depletion caused by ageing (Gorbach et al., 2017; Papenberg et al., 2014), but not typically associated with PKU.

In line with our hypotheses, a significant overlap was found between cognitive domains exhibiting impaired performance, and those apparently spared, in older controls and AwPKU. Within impaired domains, however, the severity of deficits was notably different between groups, with older controls demonstrating less impairment than AwPKU in tasks assessing executive function, whilst AwPKU demonstrated less impairment than older controls in measures of speed of processing and visuo-motor coordination. Double dissociations were also apparent, with only AwPKU performing below average in measures of IQ, and only older controls showing impaired performance in measures of memory. These findings suggest that the neurological circuits impacted by PKU and ageing show some overlap (i.e., myelin depletion causing significantly slowed speed of processing) but also some distinct differentiation, with PKU having a stronger impact on frontal lobe function, resulting in impaired executive function, and ageing having a stronger impact on the hippocampal and cerebellar systems, resulting in impaired memory and learning, and visuo-motor coordination.

Chapter 4. Speed of Processing and Executive Function in Ageing and PKU

This chapter delved deeper into the slowed response times demonstrated by both older controls and AwPKU, investigating the interaction between speed of processing and cognitive load in the domains of visuo-spatial attention and language. Distributions of reaction times across difficulty conditions within each of these domains were explored to assess the nature of the speed impairments demonstrated by both populations, and to compare the similarities and differences in the underlying cause of slowed response times for both groups. Based on the findings of the previous chapter, it was hypothesised that older controls and AwPKU would demonstrate a similar pattern of impairment across these

domains. Specifically, it was postulated that older controls would demonstrate the same patterns of impairment reported in young AwPKU following similar explorations by Romani et al. (2018), that is, a differential effect of task difficulty on speed of response in visuo-spatial compared to language tasks. It was expected that both cohorts would demonstrate a fixed delay in responses to language tasks, but an exponential impact of task difficulty on response times in the visuo-spatial domain. As speed of processing and executive function have previously been shown to be closely linked (e.g., Christ et al., 2001; Fry & Hale, 1996), this chapter also aimed to disentangle the relative contributions of executive function and speed of processing deficits to the slowed response times exhibited by both cohorts. Correlations between executive function impairments and slowed response times were, therefore, also carried out.

General linear and Brinley plot analyses of results from language and visuo-spatial tasks supported the hypotheses for this chapter, with both AwPKU and older controls demonstrating an exponential effect of task difficulty on response times in the visuo-spatial domain but a fixed delay across difficulty conditions for language tasks. Similarly to the previous chapter, however, the enhanced impact of task difficulty in the domain of visuo-spatial processing was more pronounced in older controls than in young AwPKU. These findings suggest that, rather than being due to a generalised speed deficit, processing speed in both of these populations is affected in a domain-specific manner, whilst fixed delays across domains are instead reflective of an additional level of caution in executive mechanisms. This was supported by findings of correlations between response times for visuo-spatial and language tasks of increasing difficulty, and performance in measures of executive function for both cohorts. These findings, then, support those of the previous chapter, suggesting that both AwPKU and older adults demonstrate similar patterns of impairment, therefore indicating some overlap in the mechanisms underlying impaired performance in both of these cohorts. This, in turn, suggests that impairments caused by ageing and PKU may interact with one another as AwPKU reach older age.

Chapter 5. Cognitive and Well-Being Outcomes in Middle-Aged AwPKU

In Chapters 3 and 4, we predicted how PKU may interact with normal neural degradation when early-treated AwPKU start to reach old age. Evidence of an overlap in the patterns of impairment in young AwPKU and healthy older controls suggests that AwPKU may demonstrate an accelerated effect of ageing on cognition, causing deficits associated with older age to appear earlier in this population. Chapter 5, therefore, recruited middle-aged AwPKU and age-matched healthy controls to assess performance across a range of cognitive domains and measures of well-being. Performance on cognitive measures was

then compared to previously observed impairments in a cohort of younger AwPKU. The impact of metabolic control on outcomes in middle-aged AwPKU was also investigated by comparing scores across measures of cognition and well-being between middle-aged AwPKU with good vs. poor metabolic control and carrying out correlations between blood-Phe levels and outcomes in both domains.

Findings indicated some cognitive slowing and poorer quality of life in middle-aged AwPKU compared to controls, however cognitive impairments were less severe than expected, with middle-aged AwPKU performing statistically better than young AwPKU in composite executive function scores, and demonstrating non-significantly, but numerically better, performance in composite scores of speed of processing and accuracy. No associations between concurrent Phe, or Phe over the last 10 years, and cognitive performance were found in our cohort of middle-aged AwPKU. These findings suggest that there is no evidence of accelerated ageing impacting cognitive health in AwPKU, however the small sample size of AwPKU included in this study limits the conclusions that can be drawn from these analyses. The potential impact of differences in testing modalities between young and middle-aged cohorts (i.e., remote vs. face-to-face assessment administration) must also be taken into consideration when interpreting these findings. With regards to well-being, differences between middle-aged AwPKU and controls were significant, with AwPKU demonstrating significantly higher depression and anxiety than controls, as well as a more significant impact of the pandemic on their mood. Emotional well-being and quality of life, however, were not found to be significantly related to concurrent Phe, or Phe over the last 10 years. Importantly, however, only 14 AwPKU completed well-being questionnaires, limiting the power of these findings.

Chapter 6. The Impact of Metabolic Control on Cognition, Neurophysiology, and Well-Being in PKU: A Systematic Review and Meta-Analysis of the Within-Participant Literature

The existing literature surrounding outcomes in PKU has established that early-treated AwPKU do show some evidence of impaired neurological health, and that this may interact to some extent with current metabolic control in adulthood (see Canton et al., 2019; Christ et al., 2010; Hofman et al., 2018 for reviews). Chapter 6 reviewed the within-participant literature regarding the impact of altering blood-Phe concentrations on cognitive, well-being, and neurological outcomes. This is significant because within-participant studies allow us to investigate the direct impact of blood-Phe levels on outcomes, whilst removing the potential confounding factors that may be present in between-participant and correlational studies (i.e., genetic endowment, educational opportunities, socio-economic

status, etc). This review investigated how metabolic control can influence outcomes in both early- and late-treated children, adolescents, and adults with PKU, as well as considering the parameters necessary to elicit changes in outcomes. In particular, the impact of the magnitude of change of blood-Phe concentrations, and the amount of time that they are maintained for, on the likelihood of observing cognitive, well-being, and neurological outcomes was explored.

Results from the review and meta-analysis suggest that metabolic control has a significant impact on cognitive, well-being, and neurological outcomes, with lower levels associated with better outcomes in both adults and children with PKU. Consistent associations between cognitive ability and metabolic control were found in the domains of attention, executive function, and speed of processing, supporting the existing between-participant literature, as well as our own findings. Findings across measures of well-being found differences in the domains of anxiety, depression, aggression, and hyperactivity, suggesting that metabolic control may be linked to well-being in both childhood and adulthood, despite the findings presented in Chapter 5 of this thesis. Finally, differences in white matter and EEG abnormalities were found to be associated with metabolic control in both adults and children with PKU. Cognitive and neurological impairments were also found to be reliably reversible through decreasing Phe levels, although this effect was evident at lower blood-Phe levels in CwPKU than in AwPKU. Results concerning the impact of magnitude and duration of blood-Phe changes on outcomes were varied. Larger changes in Phe levels were found to be more likely to elicit changes in cognitive outcomes, but not in well-being or neurological outcomes, whilst longer durations of altered Phe levels appeared to be more likely to impact all measured outcomes both in terms of worsening and ameliorating Phe-related impairments, although differences were either not statistically significant or there were not enough results to assess significance.

Implications

The findings presented throughout this thesis hold some key implications with regards to the cognitive impact of ageing and how this may differ between healthy older adults, and those with PKU. Differences in impairments observed between older controls and young AwPKU in cognitive domains associated with grey matter atrophy and hippocampal or cerebellar abnormalities, such as long-term memory loss, indicate that these brain regions are impacted by normal ageing, but not by PKU and, as such, would not be differentially impacted by ageing in adults with and without PKU. It remains unclear however, whether the chronic impact of high Phe levels on the brains of individuals with PKU will make this

population more susceptible to degenerative diseases primarily associated with increasing age in later life. A significant overlap in speed of processing impairments between these two cohorts is apparent, with clear evidence of demyelination and white matter atrophy in both groups. Despite ageing being found to affect both the central and peripheral nervous systems, the impact of PKU is seemingly limited to the central nervous system, suggesting that this system, specifically, may be more severely impacted by increasing age in AwPKU than in adults without PKU. Similarly, prefrontal and striatal neurotransmitter depletion with increasing age have been associated with a decrease in executive function abilities in older adults. In AwPKU, these age-related depletions may, therefore, further interact with pre-existing neurotransmitter depletions caused by excess neurological Phe, resulting in more significant executive function impairments than may otherwise be associated with the normal ageing process. The findings of this thesis indicate that cognitive impairments due to ageing and PKU are likely to interact to some extent as AwPKU reach old age leading to accelerated cognitive decline in this population. As such, the inclusion of cognitive testing in the clinical management of PKU in middle- and older-aged AwPKU would be prudent. For example, the use of a condensed version of the assessment battery used in these studies could be carried out, focussing on the domains highlighted here as being particularly sensitive to the impacts of PKU and ageing (response times, visuo-spatial processing, executive function, and sustained attention). This could allow rapid identification of accelerated cognitive ageing in AwPKU and provide the opportunity for early intervention.

This increased understanding of the similarities and differences between the profiles of cognitive ageing likely to be observed in older adults with and without PKU, as well as the potential impact of metabolic control on this process, further holds important implications for the future treatment of PKU, and the best recommendations for clinicians to present to AwPKU as they start to reach middle- and older age. Chapter 5 demonstrates that cognitive and well-being impairments in early-treated PKU persist through to adulthood, although they do not appear to increase in middle-age. Analyses investigating the extent to which impairments may be impacted by current metabolic control in middle-age suggest that there is little impact of Phe levels on outcomes in this age group, however small participant numbers mean that these results must be interpreted with caution. Chapter 6 challenges these findings, suggesting that metabolic control can impact cognitive, well-being, and neurological outcomes both in childhood and adulthood, and that existing impairments may be reversible in adulthood through a reduction of blood-Phe levels. This review found that a reduction in current Phe levels in adulthood has an ameliorating effect on well-being in severely intellectually disabled, late-treated AwPKU, as well as reducing cognitive impairment and white matter abnormalities in early-treated AwPKU. Importantly, decreased

levels are required to be maintained for at least a few months to allow reversal of neurological abnormalities to become apparent. Adoption of a relatively relaxed, but continuous, low-Phe diet in late adulthood, therefore, may serve to mitigate the impact of ageing on neurological health in ageing AwPKU. A key consideration, however, must also be the socially restrictive nature of a low-Phe diet, and the lack of palatable options currently available to those who follow it. The potentially negative impact of these factors on individual's well-being must also be taken into consideration when weighing up the potential benefits of resuming a low-Phe diet in later life.

Limitations

Remote vs. In-Person Testing

It is impossible to ignore the impact of the global COVID-19 pandemic on research practices during work for this thesis (2018-2022). A national UK-wide lockdown was declared in March 2020, causing all research to be re-designed to allow remote assessment of participants. This was achieved through a variety of methods, with some assessments rewritten in PsychoPy to allow online hosting of tasks via Pavlovia, whilst other assessments were only slightly adjusted to allow remote delivery via video call. Some assessments, however, had to be removed completely as they were not suitable for online or video call delivery. Moving assessments to remote delivery had both advantages and disadvantages. The lack of travel required, and the ability to conduct testing at participants' homes (therefore removing any issues finding appropriate spaces for testing across different collaborating sites) made it much easier to recruit and assess participants in a flexible manner to best suit their schedules. Unfortunately, whilst increased flexibility was a benefit, completion of remote assessments required participants to have appropriate equipment available (i.e., a laptop or tablet with sufficient memory to run online tasks and video call capabilities) and this will have necessarily biased recruitment of participants to those with a higher socioeconomic status, and those with an appropriate level of comfort with technology. As such, participants without access to, or understanding of, technology such as video calls, were not able to take part in assessments, likely biasing the results gathered through this manner of testing to those with higher cognitive functioning. This will have had a particularly significant effect on recruitment of AwPKU, as this is a population that has previously been found to demonstrate significant variability in cognitive impairments (Romani et al., 2019).

Participant numbers

Beyond impacting the modality of testing, the COVID-19 pandemic also had a dramatic effect on participant recruitment, in particular for participants with PKU. During the peak of the pandemic in 2020 (a key recruitment period for this thesis) all non-COVID related research was placed on hold by the NHS for approximately 4 months. Once participant recruitment was re-opened, non-COVID research remained a low priority for NHS institutions. This, in combination with staffing issues due to redeployment, off-site working, self-isolation, staff illness etc., meant that significantly less patients were recruited from clinics than had previously been intended. Despite the circumstances, every effort was made to maximise recruitment, including weekly communications with on-site dieticians to identify eligible participants, and researchers offering as much flexibility as possible for assessment times, to fit in with potential participants' schedules. Overall recruitment of AwPKU was 20 participants, one of whom was later excluded as they were not started on a low-Phe diet until 3 months old and therefore did not meet our criteria for early-treatment. This sample size is not especially small when recruiting individuals with a rare disease, with multiple previously published studies of adults and children with PKU reporting similar sample sizes between 14 and 25 participants (e.g., Channon et al., 2004; Jaulent et al., 2020; Nardecchia et al., 2015; Thompson et al., 1993; Weglage et al., 1995). However, we originally intended to recruit a minimum of 30 AwPKU to try and match previous comparisons made between 37 young AwPKU and controls, and to ensure sufficient statistical power as calculated through a-priori power analyses. As such, findings from comparisons between middle-aged adults with and without PKU must be interpreted with caution due to low statistical power. Similarly, comparisons between middle-aged AwPKU with high- and low-Phe, as well as correlations with Phe-levels, are considerably limited in power.

Historical data

Finally, it was not possible to acquire paediatric metabolic information for the middle-aged AwPKU included in the study, aside from two participants. This prevented any analyses investigating the impact of childhood metabolic control, and Phe variability, on outcomes in adulthood. All participants consented to having their childhood metabolic records accessed by researchers, however staffing limitations due to the COVID-19 pandemic meant that clinical collaborators were unable to assist with accessing these. All participants were therefore contacted to consent to 'Freedom of Information Request' letters being sent to their paediatric clinics requesting access to all childhood metabolic records. Letters were approved and sent to paediatric hospital on behalf of 10 participants. Two clinics responded with participants' paediatric records, one hospital was unable to locate the participant's paediatric records, and three hospitals requested further information from

participants but then did not respond again once information was provided. Previous research has found that childhood Phe levels can have a key impact on outcomes in later life (e.g., Aitkenhead et al., 2021; Koch, Burton, Hoganson, Peterson, Rhead, Rouse, Scott, Wolff, Stern, Guttler, & et al., 2002; Smith et al., 1996), therefore this is a necessary aspect to consider when assessing outcomes in AwPKU, and trying to disentangle the effects of childhood, adulthood, and concurrent Phe levels on performance.

Future Directions

As is often the case with scientific research, the studies carried out for this thesis raised a number of questions. There is a definite need for future studies to build upon the research conducted here. Further investigation of outcomes in middle- and older-aged AwPKU is needed. Assessing more clinical participants using the same materials is necessary to increase the power of comparisons with middle-aged controls and allow further correlation analyses to be carried out investigating the relationship between blood-Phe levels and outcomes in PKU for adults in this age range. Additional research considering the impact of childhood metabolic control on middle-aged outcomes would also be beneficial. Hopefully, as the impact of the pandemic begins to reduce, this will be something that future researchers can work with clinical collaborators to access and include in analyses moving forward.

In addition to extra information regarding cognitive and well-being outcomes in middle- and older-aged AwPKU, studies utilising neuroimaging techniques with this cohort of AwPKU are also needed to increase our understanding of the impact of PKU on neurological health in AwPKU of increasing age. This research could also be conducted with middle-aged AwPKU, and then continued as part of a longitudinal study assessing outcomes across cognitive, well-being, and neurological dimensions as AwPKU move from middle- into older-age and the effects of PKU begin to interact with those of normal ageing (in particular with regards to white matter integrity and neurotransmitter depletion). This could then support, or contradict, the predictive findings of our research with young AwPKU and healthy older controls. Longitudinal research with this cohort of AwPKU would also allow within-participant comparisons of the impact of changing Phe levels in later adulthood on outcomes to be carried out. This would further our knowledge of the potential negative impacts of increasing Phe in later life, and the potential ameliorating effects of increasing metabolic control during this stage of life.

Finally, as the effects of the COVID-19 pandemic have been felt across the globe, a significant proportion of research has been moved online to allow for remote delivery. The impact of this move has yet to be fully investigated. Future research considering the comparative efficacy of cognitive assessments delivered in remote and face-to-face conditions is going to be vital to allow appropriate interpretation of results gathered in remote settings and to inform future research design to ensure the most effective data gathering methods are being employed. As such, a future study comparing results gathered from one cohort of participants assessed using the remote version of our cognitive assessment battery and a matched cohort assessed using the face-to-face version is essential. Such a study could also consider the differential outcomes of remote vs. face-to-face conditions for participants of different age groups, as one might expect that older participants may struggle to adjust to online tasks more than younger cohorts who are more familiar with remote setups and online working.

Conclusion

The introduction of a low-Phe diet to treat PKU in the neonatal period has dramatically changed what it means to live with this condition. Early-treatment of PKU prevents the severe intellectual disability, microcephaly, seizures, and behavioural difficulties previously associated with the disorder. A low-Phe diet does not, however, prevent individuals with PKU from experiencing any negative impact of their condition, nor is it a simple (or palatable) diet to follow. As such, many individuals with PKU choose to abandon their low-Phe diet in adolescence or adulthood, once the critical period of neurodevelopment is past. This thesis investigated the impact of metabolic control on outcomes in AwPKU, in particular focusing on the potential interactions between older age and PKU in terms of cognition and well-being, and the extent to which cognitive, well-being, neuropathological and neurophysiological impairments may be reversible through decreasing concurrent blood-Phe levels, in particular in adulthood.

Our empirical investigations comparing performance in AwPKU and healthy older controls found a notable overlap in impairment, in particular in the domains of speed of processing and executive function. This study suggests that white matter degradation and neurotransmitter depletion due to increasing age may have a particularly dramatic effect on speed of processing and executive function abilities in AwPKU, as already damaged neurological systems begin to degrade further. Initial investigations of cognition in middle-aged AwPKU did not find any more severe impairments in this cohort compared to young AwPKU, however a number of limitations including small participant numbers and differential

testing modalities mean that these findings warrant further investigation before clear conclusions about the effect of middle-age on cognition in PKU can be drawn. Our systematic review and meta-analysis demonstrated that cognitive, well-being, and neurological outcomes can all be mediated by metabolic control, both in childhood and adulthood. Of particular interest, this review indicates that neurological damage caused by previously high Phe levels can be reversed through decreasing concurrent blood Phe levels in adulthood, if these lower levels are maintained for a suitable length of time. As such, a recommendation of maintaining Phe levels below a given threshold in adulthood to mitigate the potentially negative impact of ageing on outcomes may seem warranted. It is important, however, to also consider the potentially negative impact of maintaining a socially restrictive, and largely unpalatable, diet upon well-being, beyond the direct impact of metabolic control on neurotransmitter levels and white matter integrity. As such, it will be key to ensure that individual needs and tolerances are considered when making recommendations with regards to lifestyle and metabolic control for AwPKU as they start to reach older age.

Key Takeaways

- There are notable similarities between the impact of ageing and PKU on white matter integrity and neurotransmitter availability, with both AwPKU and older adults demonstrating impairments in the cognitive domains of speed of processing, executive function, sustained attention, and visuo-spatial processing. Thus, AwPKU may be at risk of accelerated cognitive ageing in these domains in later life, and therefore would benefit from additional cognitive monitoring as part of the clinical management of their PKU.
- Differences in grey matter associated cognitive abilities, such as long-term memory and learning, between older adults and AwPKU suggest that these abilities are unlikely to be at risk of accelerated decline with age due to PKU.
- Evidence of accelerated ageing is not apparent in middle-aged AwPKU, nor is there any clear cross-sectional effect of concurrent Phe levels on cognition or well-being at this age.
- Systematic review of within-participant studies suggests that metabolic control does affect cognitive, well-being, and neurological outcomes in both children and adults with PKU, with evidence that reducing Phe levels in adulthood can reverse Phe-related neuropathological damage if reduced levels are maintained for at least a few months.

References

Abdolahi, A., Bull, M. T., Darwin, K. C., Venkataraman, V., Grana, M. J., Dorsey, E. R., & Biglan, K. M. (2016). A feasibility study of conducting the Montreal Cognitive Assessment remotely in individuals with movement disorders. *Health Informatics Journal*, *22*(2), 304–311. https://doi.org/10.1177/1460458214556373

Agresti, A. (2013). *Categorical data analysis* (3rd ed). Wiley.

Ahring, K., Bélanger-Quintana, A., Dokoupil, K., Gokmen Ozel, H. G., Lammardo, A. M., MacDonald, A., Motzfeldt, K., Nowacka, M., Robert, M., & van Rijn, M. (2009). Dietary management practices in phenylketonuria across European centres. *Clinical Nutrition*, *28*(3), 231–236. https://doi.org/10.1016/j.clnu.2009.03.004

Aitkenhead, L., Krishna, G., Ellerton, C., Moinuddin, M., Matcham, J., Shiel, L., Hossain, S., Kiffin, M., Foley, J., Skeath, R., Cleary, M., Lachmann, R., & Skeath, R. J. J.. I. M. D. (2021). (2021). Long-term cognitive and psychosocial outcomes in adults with phenylketonuria. *Journal of Inherited Metabolic Disease*, *44*(6), 1353–1368. https://doi.org/10.1002/jimd.12413

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*(3), 414–421. https://doi.org/10.1016/j.neubiorev.2008.11.001

Alexander, G. E., Ryan, L., Bowers, D., Foster, T. C., Bizon, J. L., Geldmacher, D. S., & Glisky, E. L. (2012). Characterizing cognitive aging in humans with links to animal models. *Frontiers in Aging Neuroscience*, *4*, 21. https://doi.org/10.3389/fnagi.2012.00021

Allen, P. A., Bucur, B., Grabbe, J., Work, T., & Madden, D. J. (2011) Influence of encoding difficulty, word frequency, and phonological regularity on age differences in word naming. *Experimental Aging Research*, *37*(3), 261–292. https://doi.org/10.1080/0361073X.2011.568805

Alvord, Jr., E. C., Stevenson, L. D., Vogel, F. S., & Engle, Jr., R. L. (1950). Neuropathological findings in phenyl-pyruvic oligophrenia (phenyl-ketonuria). *Journal of Neuropathology and Experimental Neurology*, *9*(3), 298–310. https://doi.org/10.1097/00005072-195007000-00004

Anderson, P. J., & Leuzzi, V. (2010). White matter pathology in phenylketonuria. *Molecular Genetics and Metabolism*, *99*, Suppl. 1, S3–S9. https://doi.org/10.1016/j.ymgme.2009.10.005

Anderson, V. E., Siegel, F. S., & Bruhl, H. H. (1976). Behavioral and biochemical correlates of diet change in phenylketonuria. *Pediatric Research*, *10*(1), 10–17. https://doi.org/10.1203/00006450-197601000-00003

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*(2), 645–668. https://doi.org/10.1080/87565640701375963

Anderson, P. J., Wood, S. J., Francis, D. E., Coleman, L., Warwick, L., Casanelia, S., Anderson, V. A., & Boneh, A. (2004). Neuropsychological functioning in children with early-treated phenylketonuria: Impact of white matter abnormalities. *Developmental Medicine and Child Neurology*, *46*(4), 230–238. https://doi.org/10.1017/s0012162204000386

Antenor-Dorsey, J. A., Hershey, T., Rutlin, J., Shimony, J. S., McKinstry, R. C., Grange, D. K., Christ, S. E., & White, D. A. (2013). White matter integrity and executive abilities in individuals with phenylketonuria. *Molecular Genetics and Metabolism*, *109*(2), 125–131. https://doi.org/10.1016/j.ymgme.2013.03.020

Antshel, K. M., & Waisbren, S. E. (2003). Timing is everything: Executive functions in children exposed to elevated levels of phenylalanine. *Neuropsychology*, *17*(3), 458–468. http://doi.org/10.1037/0894-4105.17.3.458

Anwar, M. S., Waddell, B., & O'Riordan, J. (2013). Neurological improvement following reinstitution of a low phenylalanine diet after 20 years in established phenylketonuria. *BMJ Case Reports*, *2013*. https://doi.org/10.1136/bcr-2013-010509

Archibald, C. J., Wei, X., Scott, J. N., Wallace, C. J., Zhang, Y., Metz, L. M., & Mitchell, J. R. (2004). Posterior fossa lesion volume and slowed information processing in multiple sclerosis. *Brain*, *127*(7), 1526–1534. https://doi.org/10.1093/brain/awh167

Arighi, A., Fumagalli, G. G., Carandini, T., Pietroboni, A. M., De Riz, M. A., Galimberti, D., & Scarpini, E. J. N. S. (2021). Facing the digital divide into a dementia clinic during COVID-19 pandemic: Caregiver age matters. *Neurological Sciences*, *42*(4), 1247–1251. https://doi.org/10.1007/s10072-020-05009-w

Ashworth, M., Palikara, O., Burchell, E., Purser, H., Nikolla, D., & Van Herwegen, J. (2020). Online and face-to-face performance on two cognitive tasks in children with Williams syndrome. *Frontiers in Psychology*, *11*, 594465. https://doi.org/10.3389/fpsyg.2020.594465

Azen, C. G., Koch, R., Friedman, E. G., Berlow, S., Coldwell, J., Krause, W., Matalon, R., McCabe, E., O'Flynn, M., & Peterson, R. (1991). Intellectual development in 12-year-

old children treated for phenylketonuria. *American Journal of Diseases of Children*, *145*(1), 35–39. https://doi.org/10.1001/archpedi.1991.02160010037012

Bäckman, L., Lindenberger, U., Li, S. C., & Nyberg, L. (2010). Linking cognitive aging to alterations in dopamine neurotransmitter functioning: Recent data and future avenues. *Neuroscience and Biobehavioral Reviews*, *34*(5), 670–677. https://doi.org/10.1016/j.neubiorev.2009.12.008

Bannon, M. J., & Whitty, C. J. (1997). Age-related and regional differences in dopamine transporter mRNA expression in human midbrain. *Neurology*, *48*(4), 969–977. https://doi.org/10.1212/wnl.48.4.969

Bartus, A., Palasti, F., Juhasz, E., Kiss, E., Simonova, E., Sumanszki, C., & Reismann, P. (2018). The influence of blood phenylalanine levels on neurocognitive function in adult PKU patients. *Metabolic Brain Disease*, *33*(5), 1609–1615. https://doi.org/10.1007/s11011-018-0267-6

Bauman, M. L., & Kemper, T. L. (1982). Morphologic and histoanatomic observations of the brain in untreated human phenylketonuria. *Acta Neuropathologica*, *58*(1), 55–63. https://doi.org/10.1007/BF00692698

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897. https://doi.org/10.1037//0022-006x.56.6.893

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory (BDI-II)*, *10*. Pearson.

Begg, C. B. (1994). Publication bias. *The handbook of research synthesis, 25*, 299–409.

Bekhof, J., Van Spronsen, F. J., Crone, M. R., Van Rijn, M., Oudshoorn, C. G., & Verkerk, P. H. (2003). Influence of knowledge of the disease on metabolic control in phenylketonuria. *European Journal of Pediatrics*, *162*(6), 440–442. https://doi.org/10.1007/s00431-003-1197-8

Berry, S. A., Brown, C., Grant, M., Greene, C. L., Jurecki, E., Koch, J., Moseley, K., Suter, R., van Calcar, S. C., Wiles, J., & Cederbaum, S. (2013). Newborn screening 50 years later: Access issues faced by adults with PKU. *Genetics in Medicine*, *15*(8), 591–599. https://doi.org/10.1038/gim.2013.10

Bick, U., Ullrich, K., Stöber, U., Möller, H., Schuierer, G., Ludolph, A. C., Oberwittler, C., Weglage, J., & Wendel, U. (1993). White matter abnormalities in patients with treated hyperphenylalaninaemia: Magnetic resonance relaxometry and proton spectroscopy findings. *European Journal of Pediatrics*, *152*(12), 1012–1020. https://doi.org/10.1007/BF01957228

Bickel, H., Gerrard, J., & Hickmans, E. M. (1954). The influence of phenylalanine intake on the chemistry and behaviour of a phenylketonuria child. *Acta Paediatrica*, *43*(1), 64–77. https://doi.org/10.1111/j.1651-2227.1954.tb04000.x

Bik-Multanowski, M., Didycz, B., Mozrzymas, R., Nowacka, M., Kaluzny, L., Cichy, W., Schneiberg, B., Amilkiewicz, J., Bilar, A., Gizewska, M., Lange, A., Starostecka, E., Chrobot, A., Wojcicka-Bartlomiejczyk, B. I., & Milanowski, A. (2008). Quality of life in noncompliant adults with phenylketonuria after resumption of the diet. *Journal of Inherited Metabolic Disease*, *31*, Suppl. 2, S415–S418. https://doi.org/10.1007/s10545-008-0978-7

Bik-Multanowski, M., & Pietrzyk, J. J. (2011). Blood phenylalanine clearance and BH4-responsiveness in classic phenylketonuria. *Molecular Genetics and Metabolism*, *103*(4), 399–400. https://doi.org/10.1016/j.ymgme.2011.04.014

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*(2), 210–213. https://doi.org/10.1016/j.ymgme.2010.10.003

Bilder, D. A., Arnold, G. L., Dimmock, D., Grant, M. L., Janzen, D., Longo, N., Nguyen-Driver, M., Jurecki, E., Merilainen, M., Amato, G., & Waisbren, S. (2022). Improved attention linked to sustained phenylalanine reduction in adults with early-treated phenylketonuria. *American Journal of Medical Genetics. Part A*, 188(3), 768–778. https://doi.org/10.1002/ajmg.a.62574

Bilder, D. A., Burton, B. K., Coon, H., Leviton, L., Ashworth, J., Lundy, B. D., Vespa, H., Bakian, A. V., & Longo, N. (2013). Psychiatric symptoms in adults with phenylketonuria. *Molecular Genetics and Metabolism*, *108*(3), 155–160. https://doi.org/10.1016/j.ymgme.2012.12.006

Bilder, D. A., Kobori, J. A., Cohen-Pfeffer, J. L., Johnson, E. M., Jurecki, E. R., & Grant, M. L. (2017). Neuropsychiatric comorbidities in adults with phenylketonuria: A retrospective cohort study. *Molecular Genetics and Metabolism*, *121*(1), 1–8. https://doi.org/10.1016/j.ymgme.2017.03.002

Bilder, D. A., Noel, J. K., Baker, E. R., Irish, W., Chen, Y. P., Merilainen, M. J., Prasad, S., & Winslow, B. J. (2016). Systematic review and meta-analysis of neuropsychiatric symptoms and executive functioning in adults with phenylketonuria. *Developmental Neuropsychology*, *41*(4), 245–260. https://doi.org/10.1080/87565641.2016.1243109

Bilginsoy, C., Waitzman, N., Leonard, C. O., & Ernst, S. L. (2005). Living with phenylketonuria: Perspectives of patients and their families. *Journal of Inherited Metabolic Disease*, *28*(5), 639–649. https://doi.org/10.1007/s10545-005-4478-8

Bjornson, J. (1964). Behavior in phenylketonuria: Case with schizophrenia. *Archives of General Psychiatry*, *10*(1), 65–70. https://doi.org/10.1001/archpsyc.1964.01720190067009

Blau, N. (2013). Sapropterin dihydrochloride for the treatment of hyperphenylalaninemias. *Expert Opinion on Drug Metabolism and Toxicology*, *9*(9), 1207–1218. https://doi.org/10.1517/17425255.2013.804064

Blau, N., Bélanger-Quintana, A., Demirkol, M., Feillet, F., Giovannini, M., MacDonald, A., Trefz, F. K., van Spronsen, F., & European PKU centers. (2010). Management of phenylketonuria in Europe: Survey results from 19 countries. *Molecular Genetics and Metabolism*, *99*(2), 109–115. https://doi.org/10.1016/j.ymgme.2009.09.005

Blau, N., Hennermann, J. B., Langenbeck, U., & Lichter-Konecki, U. (2011). Diagnosis, classification, and genetics of phenylketonuria and tetrahydrobiopterin (BH4) deficiencies. *Molecular Genetics and Metabolism*, *104*, Suppl., S2–S9. https://doi.org/10.1016/j.ymgme.2011.08.017

Blau, N., MacDonald, A., & van Spronsen, F. (2011). There is no doubt that the early identification of PKU and prompt and continuous intervention prevents mental retardation in most patients. *Molecular Genetics and Metabolism*, *104*, Suppl., S1. https://doi.org/10.1016/j.ymgme.2011.10.007

Blau, N., van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *Lancet*, *376*(9750), 1417–1427. https://doi.org/10.1016/S0140-6736(10)60961-0

Bohnen NI, Albin RL, Müller ML, Petrou M, Kotagal V, Koeppe RA, Scott PJ, Frey KA. (2015). Frequency of cholinergic and caudate nucleus dopaminergic deficits across the predemented cognitive spectrum of Parkinson disease and evidence of interaction effects. *JAMA Neurology. 72*(2),194-200. https://doi.org/10.1001/jamaneurol.2014.2757

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. *The handbook of research synthesis and meta-analysis*, *2*, 221–235.

Bosch, A. M., Burlina, A., Cunningham, A., Bettiol, E., Moreau-Stucker, F., Koledova, E., Benmedjahed, K., & Regnault, A. (2015). Assessment of the impact of phenylketonuria and its treatment on quality of life of patients and parents from seven European countries. *Orphanet Journal of Rare Diseases*, *10*, 80. https://doi.org/10.1186/s13023-015-0294-x

Bosch, A. M., Maurice-Stam, H., Wijburg, F. A., & Grootenhuis, M. A. (2009). Remarkable differences: The course of life of young adults with galactosaemia and PKU. *Journal of Inherited Metabolic Disease*, *32*(6), 706. https://doi.org/10.1007/s10545-009-1253-2

Botwinick, J. (1977). Aging and intelligence. *Handbook of the psychology of aging*, 580–605.

Bowersox, J., & Panel, N. I. H. C. D. (2001). National Institutes of Health Consensus Development Conference Statement: Phenylketonuria: Screening and management,

October 16–18, 2000. *Pediatrics*. National Institutes of Health Consensus Development Conference Statement: Phenylketonuria: Screening and Management, *108*(4), 972–982. https://doi.org/10.1542/peds.108.4.972

Bracco, L., Piccini, C., Moretti, M., Mascalchi, M., Sforza, A., Nacmias, B., Cellini, E., Bagnoli, S., & Sorbi, S. (2005). Alzheimer's disease: Role of size and location of white matter changes in determining cognitive deficits. *Dementia and Geriatric Cognitive Disorders*, *20*(6), 358–366. https://doi.org/10.1159/000088562

Bradberry, C. W., Karasic, D. H., Deutch, A. Y., & Roth, R. H. (1989). Regionally specific alterations in mesotelencephalic dopamine synthesis in diabetic rats: Association with precursor tyrosine. *Journal of Neural Transmission. General Section*, *78*(3), 221–229. https://doi.org/10.1007/BF01249231

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review*, *27*(2), 174–186. https://doi.org/10.1007/s11065-017-9349-1

Brown, E. S., & Warner, R. (1976). Mental development of phenylketonuric children on or off diet after the age of six. *Psychological Medicine*, *6*(2), 287–296. https://doi.org/10.1017/s0033291700013842

Bruinenberg, V. M., van der Goot, E., van Vliet, D., de Groot, M. J., Mazzola, P. N., Heiner-Fokkema, M. R., van Faassen, M., van Spronsen, F. J., & van der Zee, E. A. (2016). The behavioral consequence of phenylketonuria in mice depends on the genetic background. *Frontiers in Behavioral Neuroscience*, *10*, 233. https://doi.org/10.3389/fnbeh.2016.00233

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*(5), 549–566. https://doi.org/10.1023/b:boli.0000042985.02049.ff

Bucur, B., & Madden, D. J. (2010). Effects of adult age and blood pressure on executive function and speed of processing. *Experimental Aging Research*, *36*(2), 153–168. https://doi.org/10.1080/03610731003613482

Burgard, P. (2000). Development of intelligence in early treated phenylketonuria. *European Journal of Pediatrics*, *159*, Suppl. 2, S74–S79. https://doi.org/10.1007/pl00014388

Burgard, P., Rey, F., Rupp, A., Abadie, V., & Rey, J. (1997). Neuropsychologic functions of early treated patients with phenylketonuria, on and off diet: Results of a cross-national and cross-sectional study. *Pediatric Research*, *41*(3), 368–374. https://doi.org/10.1203/00006450-199703000-00011

Burgess, N. M., Kelso, W., Malpas, C. B., Winton-Brown, T., Fazio, T., Panetta, J., De Jong, G., Neath, J., Atherton, S., Velakoulis, D., & Walterfang, M. (2021). The effect of improved dietary control on cognitive and psychiatric functioning in adults with phenylketonuria: The ReDAPT study. *Orphanet Journal of Rare Diseases*, *16*(1), 35. https://doi.org/10.1186/s13023-020-01668-2

Burlina, A. P., Lachmann, R. H., Manara, R., Cazzorla, C., Celato, A., van Spronsen, F. J., & Burlina, A. (2019). The neurological and psychological phenotype of adult patients with early-treated phenylketonuria: A systematic review. *Journal of Inherited Metabolic Disease*, *42*(2), 209–219. https://doi.org/10.1002/jimd.12065

Burton, B., Grant, M., Feigenbaum, A., Singh, R., Hendren, R., Siriwardena, K., Phillips, J., Sanchez-Valle, A., Waisbren, S., Gillis, J., Prasad, S., Merilainen, M., Lang, W., Zhang, C., Yu, S., & Stahl, S. (2015). A randomized, placebo-controlled, double-blind study of sapropterin to treat ADHD symptoms and executive function impairment in children and adults with sapropterin-responsive phenylketonuria. *Molecular Genetics and Metabolism*, *114*(3), 415–424. https://doi.org/10.1016/j.ymgme.2014.11.011

Burton, B. K., Leviton, L., Vespa, H., Coon, H., Longo, N., Lundy, B. D., Johnson, M., Angelino, A., Hamosh, A., & Bilder, D. (2013). A diversified approach for PKU treatment: Routine screening yields high incidence of psychiatric distress in phenylketonuria clinics. *Molecular Genetics and Metabolism*, *108*(1), 8–12. https://doi.org/10.1016/j.ymgme.2012.11.003

Cabalska, B., Duczyńska, N., Borzymowska, J., Zorska, K., Koślacz-Folga, A., & Bożkowa, K. (1977). Termination of dietary treatment in phenylketonuria. *European Journal of Pediatrics*, *126*(4), 253–262. https://doi.org/10.1007/BF00477051

Canton, M., Le Gall, D. L., Feillet, F., Bonnemains, C., & Roy, A. (2019). Neuropsychological profile of children with early and continuously treated phenylketonuria: Systematic review and future approaches. *Journal of the International Neuropsychological Society*, *25*(6), 624–643. https://doi.org/10.1017/S1355617719000146

Cazzorla, C., Cegolon, L., Burlina, A. P., Celato, A., Massa, P., Giordano, L., Polo, G., Daniele, A., Salvatore, F., & Burlina, A. B. (2014). Quality of Life (QoL) assessment in a cohort of patients with phenylketonuria. *BMC Public Health*, *14*, 1243. https://doi.org/10.1186/1471-2458-14-1243

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*(1), 67–83. https://doi.org/10.1037/0033-2909.98.1.67

Channon, S., German, E., Cassina, C., & Lee, P. (2004). Executive functioning, memory, and learning in phenylketonuria. *Neuropsychology*, *18*(4), 613–620. https://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*(3), 213–218. https://doi.org/10.1136/adc.2006.104786

Channon, S., Mockler, C., & Lee, P. (2005). Executive functioning and speed of processing in phenylketonuria. *Neuropsychology*, *19*(5), 679–686. https://doi.org/10.1037/0894-4105.19.5.679

Chapman, J. E., Cadilhac, D. A., Gardner, B., Ponsford, J., Bhalla, R., & Stolwyk, R. J. (2021). Comparing face-to-face and videoconference completion of the Montreal Cognitive Assessment (MoCA) in community-based survivors of stroke. *Journal of Telemedicine and Telecare*, *27*(8), 484–492. https://doi.org/10.1177/1357633X19890788

Choo, K. H., Cotton, R. G., Danks, D. M., & Jennings, I. G. (1979). Genetics of the mammalian phenylalanine hydroxylase system. Studies of human liver phenylalanine hydroxylase subunit structure and of mutations in phenylketonuria. *Biochemical Journal*, *181*(2), 285–294. https://doi.org/10.1042/bj1810285

Christ, S. E. (2003). Asbjørn Følling and the discovery of phenylketonuria. *Journal of the History of the Neurosciences*, *12*(1), 44–54. https://doi.org/10.1076/jhin.12.1.44.13788

Christ, S. E., Huijbregts, S. C., de Sonneville, L. M., & White, D. A. (2010). Executive function in early-treated phenylketonuria: Profile and underlying mechanisms. *Molecular Genetics and Metabolism*, *99*, Suppl. 1, S22–S32. https://doi.org/10.1016/j.ymgme.2009.10.007

Clarke, J. T., Gates, R. D., Hogan, S. E., Barrett, M., & MacDonald, G. W. (1987). Neuropsychological studies on adolescents with phenylketonuria returned to phenylalanine-restricted diets. *American Journal of Mental Retardation*, *92*(3), 255–262.

Cleary, M. A., Walter, J. H., Wraith, J. E., Jenkins, J. P., Alani, S. M., Tyler, K., & Whittle, D. (1994). Magnetic resonance imaging of the brain in phenylketonuria. *Lancet*, *344*(8915), 87–90. https://doi.org/10.1016/s0140-6736(94)91281-5

Cleary, M. A., Walter, J. H., Wraith, J. E., White, F., Tyler, K., & Jenkins, J. P. (1995). Magnetic resonance imaging in phenylketonuria: Reversal of cerebral white matter change. *Journal of Pediatrics*, *127*(2), 251–255. https://doi.org/10.1016/s0022-3476(95)70303-9

Clocksin, H. E., Hawks, Z. W., White, D. A., & Christ, S. E. (2021). Inter- and intra-tract analysis of white matter abnormalities in individuals with early-treated phenylketonuria (PKU). *Molecular Genetics and Metabolism*, *132*(1), 11–18. https://doi.org/10.1016/j.ymgme.2020.12.001

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science*, *3*(5), 155–158. https://doi.org/10.1111/1467-8721.ep10770653

Crone, M. R., van Spronsen, F. J., Oudshoorn, K., Bekhof, J., van Rijn, G., & Verkerk, P. H. (2005). Behavioural factors related to metabolic control in patients with phenylketonuria. *Journal of Inherited Metabolic Disease*, *28*(5), 627–637. https://doi.org/10.1007/s10545-005-0014-0

Crosson, B. A. (1992). *Subcortical functions in language and memory*. Guilford Press.

Curtius, H. C., Baerlocher, K., & Völlmin, J. A. (1972b). Pathogenesis of phenylketonuria: Inhibition of DOPA and catecholamine synthesis in patients with phenylketonuria. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, *42*(1), 235–239. https://doi.org/10.1016/0009-8981(72)90406-8

Dauphinot, V., Boublay, N., Moutet, C., Achi, S., Bathsavanis, A., & Krolak-Salmon, P. (2020). Comparison of Instrumental Activities of Daily Living assessment by face-to-face or telephone interviews: A randomized, crossover study. *Alzheimer's Research and Therapy*, *12*(1), 24. https://doi.org/10.1186/s13195-020-00590-w

Dawson, C., Murphy, E., Maritz, C., Chan, H., Ellerton, C., Carpenter, R. H. S., & Lachmann, R. H. (2011). Dietary treatment of phenylketonuria: The effect of phenylalanine on reaction time. *Journal of Inherited Metabolic Disease*, *34*(2), 449–454. https://doi.org/10.1007/s10545-010-9276-2

De Beni, R., & Palladino, P. (2004). Decline in working memory updating through ageing: Intrusion error analyses. *Memory*, *12*(1), 75–89. https://doi.org/10.1080/09658210244000568

De Felice, S., Romani, C., Geberhiwot, T., MacDonald, A., & Palermo, L. (2018). Language processing and executive functions in early treated adults with phenylketonuria (PKU). *Cognitive Neuropsychology*, *35*(3–4), 148–170. https://doi.org/10.1080/02643294.2017.1422709

de Groot, M. J., Hoeksma, M., Blau, N., Reijngoud, D. J., & van Spronsen, F. J. (2010). Pathogenesis of cognitive dysfunction in phenylketonuria: Review of hypotheses. *Molecular Genetics and Metabolism*, *99*, Suppl. 1, S86–S89. https://doi.org/10.1016/j.ymgme.2009.10.016

de la Parra, A., García, M. I., Hamilton, V., Arias, C., Cabello, J. F., & Cornejo, V. (2017). First-year metabolic control guidelines and their impact on future metabolic control and neurocognitive functioning in children with PKU. *Molecular Genetics and Metabolism Reports*, *13*, 90–94. https://doi.org/10.1016/j.ymgmr.2017.09.003

de Quervain, D. J., Henke, K., Aerni, A., Coluccia, D., Wollmer, M. A., Hock, C., Nitsch, R. M., & Papassotiropoulos, A. (2003). A functional genetic variation of the 5-HT2a receptor affects human memory. *Nature Neuroscience*, *6*(11), 1141–1142. https://doi.org/10.1038/nn1146

Delaloye, C., Moy, G., Baudois, S., De Bilbao, F., Dubois Remund, C. D., Hofer, F., Ragno Paquier, C., Weber, K., Urben, S., & Giannakopoulos, P. (2009). The contribution of aging to the understanding of the dimensionality of executive functions. *Archives of Gerontology and Geriatrics*, *49*(1), e51–e59. https://doi.org/10.1016/j.archger.2008.08.011

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*(1), 62–73. https://doi.org/10.1037/0882-7974.21.1.62

DeRoche, K., & Welsh, M. (2008). Twenty-five years of research on neurocognitive outcomes in early-treated phenylketonuria: Intelligence and executive function. *Developmental Neuropsychology*, *33*(4), 474–504. https://doi.org/10.1080/87565640802101482

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997a). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development*, 62(4), i-206. https://doi.org/10.2307/1166208

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997b). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU – Introduction. *Monographs of the Society for Research in Child Development*, *62*(4), 1–+

Dion, E., Prévost, M. J., Carrière, S., Babin, C., & Goisneau, J. (2001). Phenylalanine restricted diet treatment of the aggressive behaviours of a person with mental retardation. *British Journal of Development Disabilities*, *47*(92), 21–29. https://doi.org/10.1179/096979501799155639

Dixon, R. A., Kramer, D., & Baltes, P. B. (1985). Intelligence: A life-span developmental perspective. In *Handbook of intelligence: Theories, measurements, and applications* (pp. 301–350). Wiley.

Douglas, T. D., Ramakrishnan, U., Kable, J. A., & Singh, R. H. (2013). Longitudinal quality of life analysis in a phenylketonuria cohort provided sapropterin dihydrochloride. *Health and Quality of Life Outcomes*, *11*, 218. https://doi.org/10.1186/1477-7525-11-218

Drag, L. L., & Bieliauskas, L. A. J. J. (2010). Contemporary review 2009: Cognitive aging. *Journal of Geriatric Psychiatry and Neurology*, *23*(2), 75–93. https://doi.org/10.1177/0891988709358590

Drag, L. L., Light, S. N., Langenecker, S. A., Hazlett, K. E., Wilde, E. A., Welsh, R., Steinberg, B. A., & Bieliauskas, L. A. (2016). Patterns of frontoparietal activation as a marker for unsuccessful visuospatial processing in healthy aging. *Brain Imaging and Behavior*, *10*(3), 686–696. https://doi.org/10.1007/s11682-015-9428-y

Dyer, C. A. (2000). Comments on the neuropathology of phenylketonuria. *European Journal of Pediatrics*, *159*, Suppl. 2, S107–S108. https://doi.org/10.1007/pl00014369

Dyer, C. A., Kendler, A., Philibotte, T., Gardiner, P., Cruz, J., & Levy, H. L. (1996). Evidence for central nervous system glial cell plasticity in phenylketonuria. *Journal of Neuropathology and Experimental Neurology*, *55*(7), 795–814. https://doi.org/10.1097/00005072-199607000-00005

Eckert, M. A., Keren, N. I., Roberts, D. R., Calhoun, V. D., & Harris, K. C. (2010). Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*, *4*, 10. https://doi.org/10.3389/neuro.09.010.2010

Egger, M., Davey Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Elbaz, S., Cinalioglu, K., Sekhon, K., Gruber, J., Rigas, C., Bodenstein, K., Naghi, K., Lavin, P., Greenway, K. T., Vahia, I., Rej, S., & Sekhon, H. (2021). A systematic review of telemedicine for older adults with dementia during COVID-19: An alternative to in-person health services? *Frontiers in Neurology*, *12*, 761965. https://doi.org/10.3389/fneur.2021.761965

Enzinger, C., Fazekas, F., Matthews, P. M., Ropele, S., Schmidt, H., Smith, S., & Schmidt, R. (2005). Risk factors for progression of brain atrophy in aging: Six-year follow-up of normal subjects. *Neurology*, *64*(10), 1704–1711. https://doi.org/10.1212/01.WNL.0000161871.83614.BB

Erixon-Lindroth, N., Farde, L., Wahlin, T. B., Sovago, J., Halldin, C., & Bäckman, L. (2005). The role of the striatal dopamine transporter in cognitive aging. *Psychiatry Research*, *138*(1), 1–12. https://doi.org/10.1016/j.pscychresns.2004.09.005

Fearnley, J. M., & Lees, A. J. (1991). Ageing and Parkinson's disease: Substantia nigra regional selectivity. *Brain*, *114*(5), 2283–2301. https://doi.org/10.1093/brain/114.5.2283

Feldmann, R., Osterloh, J., Onon, S., Fromm, J., Rutsch, F., & Weglage, J. (2019). Neurocognitive functioning in adults with phenylketonuria: Report of a 10-year follow-up. *Molecular Genetics and Metabolism*, *126*(3), 246–249. https://doi.org/10.1016/j.ymgme.2018.12.011

Ferreira, B. K., Rodrigues, M. T., Streck, E. L., Ferreira, G. C., & Schuck, P. F. (2021). White matter disturbances in phenylketonuria: Possible underlying mechanisms. *Journal of Neuroscience Research*, 99(1), 349–360. https://doi.org/10.1002/jnr.24598

Fiege, B., & Blau, N. (2007). Assessment of tetrahydrobiopterin (BH4) responsiveness in phenylketonuria. *Journal of Pediatrics*, *150*(6), 627–630. https://doi.org/10.1016/j.jpeds.2007.02.017

Fitzgerald, B., Morgan, J., Keene, N., Rollinson, R., Hodgson, A., & Dalrymple-Smith, J. (2000). An investigation into diet treatment for adults with previously untreated

phenylketonuria and severe intellectual disability. *Journal of Intellectual Disability Research*, *44*(1), 53–59. https://doi.org/10.1046/j.1365-2788.2000.00260.x

Folk, C. L., & Lincourt, A. E. (1996). The effects of age on guided conjunction search. *Experimental Aging Research*, *22*(1), 99–118. https://doi.org/10.1080/03610739608254000

Ford, S., O'Driscoll, M., & MacDonald, A. (2018). Living with phenylketonuria: Lessons from the PKU community. *Molecular Genetics and Metabolism Reports*, *17*, 57–63. https://doi.org/10.1016/j.ymgmr.2018.10.002

Foster, J. K., Behrmann, M., & Stuss, D. T. (1995). Aging and visual search: Generalized cognitive slowing or selective deficit in attention? *Aging, Neuropsychology, and Cognition*, *2*(4), 279–299. https://doi.org/10.1080/13825589508256604

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., Démonet, J. F., Garibotto, V., Giannakopoulos, P., Gietl, A., Hansson, O., Herholz, K., Jack, C. R., Nobili, F., Nordberg, A., Snyder, H. M., Ten Kate, M., Varrone, A., Albanese, E., . . . & Winblad, B. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet. Neurology*, *16*(8), 661–676. https://doi.org/10.1016/S1474-4422(17)30159-X

Galluzzi, S., Lanni, C., Pantoni, L., Filippi, M., & Frisoni, G. B. (2008). White matter lesions in the elderly: Pathophysiological hypothesis on the effect on brain plasticity and reserve. *Journal of the Neurological Sciences*, *273*(1–2), 3–9. https://doi.org/10.1016/j.jns.2008.06.023

Gassió, R., Campistol, J., Vilaseca, M. A., Lambruschini, N., Cambra, F. J., & Fusté, E. (2003). Do adult patients with phenylketonuria improve their quality of life after introduction/resumption of a phenylalanine-restricted diet? *Acta Paediatrica*, *92*(12), 1474–1478. https://doi.org/10.1080/08035250310006683

Gately, M. E., Tickle-Degnen, L., McLaren, J. E., Ward, N., Ladin, K., & Moo, L. R. (2022). Factors influencing barriers and facilitators to in-home video telehealth for dementia management. *Clinical Gerontologist*, *45*(4), 1020–1033. https://doi.org/10.1080/07317115.2021.1930316

Genova, H. M., Hillary, F. G., Wylie, G., Rypma, B., & Deluca, J. (2009). Examination of processing speed deficits in multiple sclerosis using functional magnetic resonance imaging. *Journal of the International Neuropsychological Society*, *15*(3), 383–393. https://doi.org/10.1017/S1355617709090535

Giffin, F. D., Clarke, J. T., & d'Entremont, D. M. (1980). Effect of dietary phenylalanine restriction on visual attention span in mentally retarded subjects with phenylketonuria. *Canadian Journal of Neurological Sciences. le Journal Canadien des Sciences Neurologiques*, *7*(2), 127–131. https://doi.org/10.1017/s0317167100023490

González, M. J., Gutiérrez, A. P., Gassió, R., Fusté, M. E., Vilaseca, M. A., & Campistol, J. (2011). Neurological complications and behavioral problems in patients with phenylketonuria in a follow-up unit. *Molecular Genetics and Metabolism*, *104*, Suppl., S73–S79. https://doi.org/10.1016/j.ymgme.2011.07.015

Goodman-Casanova, J. M., Dura-Perez, E., Guzman-Parra, J., Cuesta-Vargas, A., & Mayoral-Cleries, F. (2020). Telehealth home support during COVID-19 confinement for community-dwelling older adults with mild cognitive impairment or mild dementia: Survey study. *Journal of Medical Internet Research*, *22*(5), e19434. https://doi.org/10.2196/19434

Gorbach, T., Pudas, S., Lundquist, A., Orädd, G., Josefsson, M., Salami, A., de Luna, X., & Nyberg, L. (2017). Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiology of Aging*, *51*, 167–176. https://doi.org/10.1016/j.neurobiolaging.2016.12.002

Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science*, *27*(1), 1–40. https://doi.org/10.1207/s15516709cog2701_1

Greenwood, P. M., & Parasuraman, R. (1999). Scale of attentional focus in visual search. *Perception and Psychophysics*, *61*(5), 837–859. https://doi.org/10.3758/bf03206901

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*(5), 365–372. https://doi.org/10.1111/j.1365-2788.1995.tb00540.x

Griffiths, P., Ward, N., Harvie, A., & Cockburn, F. (1998). Neuropsychological outcome of experimental manipulation of phenylalanine intake in treated phenylketonuria. *Journal of Inherited Metabolic Disease*, *21*(1), 29–38. https://doi.org/10.1023/a:1005307229813

Grueter, B. E., & Schulz, U. G. (2012). Age-related cerebral white matter disease (leukoaraiosis): A review. *Postgraduate Medical Journal*, *88*(1036), 79–87. https://doi.org/10.1136/postgradmedj-2011-130307

Güttler, F., & Lou, H. (1986). Dietary problems of phenylketonuria: Effect on CNS transmitters and their possible role in behaviour and neuropsychological function. *Journal of Inherited Metabolic Disease*, *9*, Suppl. 2, 169–177. https://doi.org/10.1007/BF01799701

Hackney, I. M., Hanley, W. B., Davidson, W., & Lindsao, L. (1968). Phenylketonuria: Mental development, behavior, and termination of low phenylalanine diet. *Journal of Pediatrics*, *72*(5), 646–655. https://doi.org/10.1016/s0022-3476(68)80007-1

Hale, S., Lima, S. D., & Myerson, J. (1991). General cognitive slowing in the nonlexical domain: An experimental validation. *Psychology and Aging*, *6*(4), 512–521. https://doi.org/10.1037//0882-7974.6.4.512

Hamilton, L. D., Thomas, E., Almuklass, A. M., & Enoka, R. M. (2017). A framework for identifying the adaptations responsible for differences in pegboard times between middle-aged and older adults. *Experimental Gerontology*, *97*, 9–16. https://doi.org/10.1016/j.exger.2017.07.003

Hamilton, V., Santa María, L., Fuenzalida, K., Morales, P., Desviat, L. R., Ugarte, M., Pérez, B., Cabello, J. F., & Cornejo, V. (2018). Characterization of phenyalanine hydroxylase gene mutations in Chilean PKU patients. *JIMD Reports*, *42*, 71–77. https://doi.org/10.1007/8904_2017_85

Harding, C. O., Amato, R. S., Stuy, M., Longo, N., Burton, B. K., Posner, J., Weng, H. H., Merilainen, M., Gu, Z., Jiang, J., Vockley, J., & PRISM-2 Investigators. (2018). Pegvaliase for the treatment of phenylketonuria: A pivotal, double-blind randomized discontinuation phase 3 clinical trial. *Molecular Genetics and Metabolism*, *124*(1), 20–26. https://doi.org/10.1016/j.ymgme.2018.03.003

Harper, M., & Reid, A. H. (1987). Use of a restricted protein diet in the treatment of behaviour disorder in a severely mentally retarded adult female phenylketonuric patient. *Journal of Mental Deficiency Research*, *31*(2), 209–212. https://doi.org/10.1111/j.1365-2788.1987.tb01357.x

Harpur, L. L., Scialfa, C. T., & Thomas, D. M. (1995). Age differences in feature search as a function of exposure duration. *Experimental Aging Research*, *21*(1), 1–15. https://doi.org/10.1080/03610739508254264

Hawks, Z., Hood, A. M., Lerman-Sinkoff, D. B., Shimony, J. S., Rutlin, J., Lagoni, D., Grange, D. K., & White, D. A. (2019). White and gray matter brain development in children and young adults with phenylketonuria. *NeuroImage. Clinical*, *23*, 101916. https://doi.org/10.1016/j.nicl.2019.101916

Hays, R. D., Sherbourne, C. D., & Mazel, R. M. (1993). The RAND 36-item health survey 1.0. *Health Economics*, *2*(3), 217–227. https://doi.org/10.1002/hec.4730020305

Hedman, A. M., van Haren, N. E., Schnack, H. G., Kahn, R. S., & Hulshoff Pol, H. E. (2012). Human brain changes across the life span: A review of 56 longitudinal magnetic resonance imaging studies. *Human Brain Mapping*, *33*(8), 1987–2002. https://doi.org/10.1002/hbm.21334

Hellewell, S. C., Welton, T., Eisenhuth, K., Tchan, M. C., & Grieve, S. M. (2021). Diffusion kurtosis imaging detects subclinical white matter abnormalities in phenylketonuria. *NeuroImage. Clinical*, *29*, 102555. https://doi.org/10.1016/j.nicl.2020.102555

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hoaglin, D. C. (2016). Misunderstandings about Q and "Cochran's Q test"in meta-analysis. *Statistics in Medicine*, *35*(4), 485–495. https://doi.org/10.1002/sim.6632

Hofman, D. L., Champ, C. L., Lawton, C. L., Henderson, M., & Dye, L. (2018). A systematic review of cognitive functioning in early treated adults with phenylketonuria. *Orphanet Journal of Rare Diseases*, *13*(1), 150. https://doi.org/10.1186/s13023-018-0893-4

Hogan, S. E., Gates, R. D., MacDonald, G. W., & Clarke, J. T. (1986). Experience with adolescents with phenylketonuria returned to phenylalanine-restricted diets. *Journal of the American Dietetic Association*, *86*(9), 1203–1207. https://doi.org/10.1016/S0002-8223(21)04094-3

Holtzman, N. A., Kronmal, R. A., van Doorninck, W., Azen, C., & Koch, R. (1986). Effect of age at loss of dietary control on intellectual performance and behavior of children with phenylketonuria. *New England Journal of Medicine*, *314*(10), 593–598. https://doi.org/10.1056/NEJM198603063141001

Holtzman, N. A., Welcher, D. W., & Mellits, E. D. (1975). Termination of restricted diet in children with phenylketonuria: A randomized controlled study. *New England Journal of Medicine*, *293*(22), 1121–1124. https://doi.org/10.1056/NEJM197511272932204

Hommel, B., Li, K. Z., & Li, S. C. (2004). Visual search across the life span. *Developmental Psychology*, *40*(4), 545–558. https://doi.org/10.1037/0012-1649.40.4.545

Hopkins, R. O., Beck, C. J., Burnett, D. L., Weaver, L. K., Victoroff, J., & Bigler, E. D. (2006). Prevalence of white matter hyperintensities in a young healthy population. *Journal of Neuroimaging*, *16*(3), 243–251. https://doi.org/10.1111/j.1552-6569.2006.00047.x

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*(2), 107–129. https://doi.org/10.1016/0001-6918(67)90011-x

Hoskin, R. G., Sasitharan, T., & Howard, R. (1992). The use of a low phenylalanine diet with amino acid supplement in the treatment of behavioural problems in a severely mentally retarded adult female with phenylketonuria. *Journal of Intellectual Disability Research*, *36*(2), 183–191. https://doi.org/10.1111/j.1365-2788.1992.tb00494.x

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Hoyer, W. J., Stawski, R. S., Wasylyshyn, C., & Verhaeghen, P. (2004). Adult age and digit symbol substitution performance: A meta-analysis. *Psychology and Aging*, *19*(1), 211–214. https://doi.org/10.1037/0882-7974.19.1.211

Huijbregts, S., de Sonneville, L., Licht, R., Sergeant, J., & van Spronsen, F. (2002). Inhibition of prepotent responding and attentional flexibility in treated phenylketonuria. *Developmental Neuropsychology*, *22*(2), 481–499. https://doi.org/10.1207/S15326942DN2202_4

Huijbregts, S. C., de Sonneville, L. M., Licht, R., van Spronsen, F. J., & Sergeant, J. A. (2002). Short-term dietary interventions in children and adolescents with treated phenylketonuria: Effects on neuropsychological outcome of a well-controlled population. *Journal of Inherited Metabolic Disease*, *25*(6), 419–430. https://doi.org/10.1023/a:1021205713674

Huijbregts, S. C. J., de Sonneville, L. M. J., van Spronsen, F. J., Licht, R., & Sergeant, J. A. (2002). The neuropsychological profile of early and continuously treated phenylketonuria: Orienting, vigilance, and maintenance versus manipulation-functions of working memory. *Neuroscience and Biobehavioral Reviews*, *26*(6), 697–712. https://doi.org/10.1016/s0149-7634(02)00040-4

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*(4), 275–292. https://doi.org/10.1111/1468-2389.00156

Ioannidis, J. (2017). Statistical biases in science communication: What we know about them and how they can be addressed. *The oxford handbook of the science of science communication*, 61–72.

Iyer, S., Mehta, P., Weith, J., Hoang-Gia, D., Moore, J., Carlson, C., Choe, P., Sakai, E., & Gould, C. (2021). Converting a geriatrics clinic to virtual visits during Covid-19: A case study. *Journal of Primary Care and Community Health*, *12*, 21501327211000235. https://doi.org/10.1177/21501327211000235

Jahja, R., Huijbregts, S. C., de Sonneville, L. M., van der Meere, J. J., Bosch, A. M., Hollak, C. E., Rubio-Gozalbo, M. E., Brouwers, M. C., Hofstede, F. C., de Vries, M. C., Janssen, M. C., van der Ploeg, A. T., Langendonk, J. G., & van Spronsen, F. J. (2013). Mental health and social functioning in early treated phenylketonuria: The PKU-COBESO study. *Molecular Genetics and Metabolism*, *110*, Suppl., S57–S61. https://doi.org/10.1016/j.ymgme.2013.10.011

Jahja, R., Huijbregts, S. C. J., De Sonneville, L. M. J., van der Meere, J. J., Legemaat, A. M., Bosch, A. M., Hollak, C. E. M., Rubio-Gozalbo, M. E., Brouwers, M. C. G. J., Hofstede, F. C., de Vries, M. C., Janssen, M. C. H., van der Ploeg, A. T., Langendonk, J. G., & van Spronsen, F. J. (2017). Cognitive profile and mental health in adult phenylketonuria: A PKU-COBESO study. *Neuropsychology*, *31*(4), 437–447. https://doi.org/10.1037/neu0000358

Jahja, R., Huijbregts, S. C. J., de Sonneville, L. M. J., van der Meere, J. J., Legemaat, A. M., Bosch, A. M., Hollak, C. E. M., Rubio-Gozalbo, M. E., Brouwers, M. C. G. J., Hofstede, F. C., de Vries, M. C., Janssen, M. C. H., van der Ploeg, A. T., Langendonk, J. G., & van Spronsen, F. J. (2017). Cognitive profile and mental health in adult phenylketonuria: A PKU-COBESO study. *Neuropsychology*, *31*(4), 437–447. https://doi.org/10.1037/neu0000358

Jahja, R., van Spronsen, F. J., de Sonneville, L. M. J., van der Meere, J. J., Bosch, A. M., Hollak, C. E. M., Rubio-Gozalbo, M. E., Brouwers, M. C. G. J., Hofstede, F. C., de Vries, M. C., Janssen, M. C. H., van der Ploeg, A. T., Langendonk, J. G., & Huijbregts, S. C. J. (2017). Long-term follow-up of cognition and mental health in adult phenylketonuria: A PKU-COBESO study. *Behavior Genetics*, *47*(5), 486–497. https://doi.org/10.1007/s10519-017-9863-1

Jancar, J. (1998). Increased life expectancy in people with untreated phenylketonuria. *Journal of Intellectual Disability Research*, *42*(1), 97–99. https://doi.org/10.1046/j.1365-2788.1998.00067.x

Janos, A. L., Grange, D. K., Steiner, R. D., & White, D. A. (2012). Processing speed and executive abilities in children with phenylketonuria. *Neuropsychology*, *26*(6), 735–743. http://doi.org/10.1037/a0029419

Jaulent, P., Charriere, S., Feillet, F., Douillard, C., Fouilhoux, A., & Thobois, S. (2020). Neurological manifestations in adults with phenylketonuria: New cases and review of the literature. *Journal of Neurology*, *267*(2), 531–542. https://doi.org/10.1007/s00415-019-09608-2

Jervis, G. A. (1954). Phenylpyruvic oligophrenia (phenylketonuria). *Research Publications – Association for Research in Nervous and Mental Disease*, *33*, 259–282.

Jiang, J., Paradise, M., Liu, T., Armstrong, N. J., Zhu, W., Kochan, N. A., Brodaty, H., Sachdev, P. S., & Wen, W. (2018). The association of regional white matter lesions with cognition in a community-based cohort of older individuals. *NeuroImage. Clinical*, *19*, 14–21. https://doi.org/10.1016/j.nicl.2018.03.035

Joy, S., Fein, D., Kaplan, E., & Freedman, M. (2000). Speed and memory in WAIS-R-NI Digit Symbol performance among healthy older adults. *Journal of the International*

Neuropsychological Society, *6*(7), 770–780. https://doi.org/10.1017/s1355617700677044

Kaasinen, V., & Rinne, J. O. (2002). Functional imaging studies of dopamine system and cognition in normal aging and Parkinson's disease. *Neuroscience and Biobehavioral Reviews*, *26*(7), 785–793. https://doi.org/10.1016/s0149-7634(02)00065-9

Kaasinen, V., Vilkman, H., Hietala, J., Någren, K., Helenius, H., Olsson, H., Farde, L., & Rinne, J. O. (2000). Age-related dopamine D2/D3 receptor loss in extrastriatal regions of the human brain. *Neurobiology of Aging*, *21*(5), 683–688. https://doi.org/10.1016/s0197-4580(00)00149-4

Kaufman, A. S., & Horn, J. L. (1996). Age changes on tests of fluid and crystallized ability for women and men on the Kaufman Adolescent and Adult Intelligence Test (KAIT) at ages 17–94 years. *Archives of Clinical Neuropsychology*, *11*(2), 97–121. https://doi.org/10.1093/arclin/11.2.97

Kaufman, A. S., & Kaufman, N. L. (1993). *KAIT: Kaufman adolescent and adult intelligence test* [Manual]. American Guidance Service.

Khan, A. A., Lodhi, F. S., Rabbani, U., Ahmed, Z., Abrar, S., Arshad, S., Irum, S., & Khan, M. I. (2020). Impact of coronavirus disease (COVID-19) pandemic on psychological well-being of the Pakistani general population. *Frontiers in Psychiatry*, *11*, 564364. https://doi.org/10.3389/fpsyt.2020.564364

Koch, R., Azen, C. G., Friedman, E. G., & Williamson, M. L. (1982). Preliminary report on the effects of diet discontinuation in PKU. *Journal of Pediatrics*, *100*(6), 870–875. https://doi.org/10.1016/s0022-3476(82)80503-9

Koch, R., Burton, B., Hoganson, G., Peterson, R., Rhead, W., Rouse, B., Scott, R., Wolff, J., Stern, A. M., Guttler, F., Nelson, M., de la Cruz, F., Coldwell, J., Erbe, R., Geraghty, M. T., Shear, C., Thomas, J., & Azen, C. (2002). Phenylketonuria in adulthood: A collaborative study. *Journal of Inherited Metabolic Disease*, *25*(5), 333–346. https://doi.org/10.1023/a:1020158631102

Koch, R., Burton, B., Hoganson, G., Peterson, R., Rhead, W., Rouse, B., Scott, R., Wolff, J., Stern, A. M., Guttler, F., Nelson, M., de la Cruz, F., Coldwell, J., Erbe, R., Geraghty, M. T., Shear, C., Thomas, J., & Azen, C. (2002). Phenylketonuria in adulthood: A collaborative study. *Journal of Inherited Metabolic Disease*, *25*(5), 333–346. https://doi.org/10.1023/a:1020158631102

Koch, R., Fishler, K., Schild, S., & Ragsdale, N. (1964). Clinical aspects of phenylketonuria. *Mental Retardation*, *2*(1), 47–54.

Koch, R., Friedman, E. G., Williamson, M. L., & Azen, C. G. (1982). Preliminary report of the effects of diet discontinuation in phenylketonuria. *Journal of Inherited Metabolic Disease*, *5*(S1), Suppl. 1, 63–64. https://doi.org/10.1007/BF01799830

Koppel, J., & Goldberg, T. (2009). The genetics of episodic memory. *Cognitive Neuropsychiatry*, *14*(4–5), 356–376. https://doi.org/10.1080/13546800902990438

Koppelmans, V., Hirsiger, S., Mérillat, S., Jäncke, L., & Seidler, R. D. (2015). Cerebellar gray and white matter volume and their relation with age and manual motor performance in healthy older adults. *Human Brain Mapping*, *36*(6), 2352–2363. https://doi.org/10.1002/hbm.22775

Kramer, A. F., Martin-Emerson, R., Larish, J. F., & Andersen, G. J. (1996). Aging and filtering by movement in visual search. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *51*(4), P201–P216. https://doi.org/10.1093/geronb/51b.4.p201

Krause, W., Halminski, M., McDonald, L., Dembure, P., Salvo, R., Freides, D., & Elsas, L. (1985). Biochemical and neuropsychological effects of elevated plasma phenylalanine in patients with treated phenylketonuria. A model for the study of phenylalanine and brain function in man. *Journal of Clinical Investigation*, *75*(1), 40–48. https://doi.org/10.1172/JCI111695

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychologist*, *59*(2), 105–117. https://doi.org/10.1037/0003-066X.59.2.105

Lai, F. H. Y., Yan, E. W. H., Yu, K. K. Y., Tsui, W. S., Chan, D. T. H., & Yee, B. K. (2020). The protective impact of telemedicine on persons with dementia and their caregivers during the COVID-19 pandemic. *American Journal of Geriatric Psychiatry*, *28*(11), 1175–1184. https://doi.org/10.1016/j.jagp.2020.07.019

Landolt, M. A., Nuoffer, J. M., Steinmann, B., & Superti-Furga, A. (2002). Quality of life and psychologic adjustment in children and adolescents with early treated phenylketonuria can be normal. *Journal of Pediatrics*, *140*(5), 516–521. https://doi.org/10.1067/mpd.2002.123663

Landvogt, C., Mengel, E., Bartenstein, P., Buchholz, H. G., Schreckenberger, M., Siessmeier, T., Scheurich, A., Feldmann, R., Weglage, J., Cumming, P., Zepp, F., & Ullrich, K. (2008). Reduced cerebral fluoro-L-dopamine uptake in adult patients suffering from phenylketonuria. *Journal of Cerebral Blood Flow and Metabolism*, *28*(4), 824–831. https://doi.org/10.1038/sj.jcbfm.9600571

Launer, L. J., Berger, K., Breteler, M. M., Dufouil, C., Fuhrer, R., Giampaoli, S., Nilsson, L. G., Pajak, A., de Ridder, M., van Dijk, E. J., Sans, S., Schmidt, R., & Hofman, A. (2006). Regional variability in the prevalence of cerebral white matter lesions: An MRI study in 9 European countries (Cascade). *Neuroepidemiology*, *26*(1), 23–29. https://doi.org/10.1159/000089233

Lefever, S., Dal, M., & Matthíasdóttir, Á. (2007). Online data collection in academic research: Advantages and limitations. *British Journal of Educational Technology*, *38*(4), 574–582. https://doi.org/10.1111/j.1467-8535.2006.00638.x

Lee, P. J., Amos, A., Robertson, L., Fitzgerald, B., Hoskin, R., Lilburn, M., Weetch, E., & Murphy, G. (2009). Adults with late diagnosed PKU and severe challenging behaviour: A randomised placebo-controlled trial of a phenylalanine-restricted diet. *Journal of Neurology, Neurosurgery, and Psychiatry*, *80*(6), 631–635. https://doi.org/10.1136/jnnp.2008.151175

Lee, J., Sung, J., & Choi, M. (2020). The factors associated with subjective cognitive decline and cognitive function among older adults. *Journal of Advanced Nursing*, *76*(2), 555–565. https://doi.org/10.1111/jan.14261

Leuzzi, V., Fois, D., Carducci, C., Antonozzi, I., & Trasimeni, G. (1997). Neuropsychological and neuroradiological (MRI) variations during phenylalanine load: Protective effect of valine, leucine, and isoleucine supplementation. *Journal of Child Neurology*, *12*(5), 338–340. https://doi.org/10.1177/088307389701200511

Leuzzi, V., Gualdi, G. F., Fabbrizi, F., Trasimeni, G., Di Biasi, C., & Antonozzi, I. (1993). Neuroradiological (MRI) abnormalities in phenylketonuric subjects: Clinical and biochemical correlations. *Neuropediatrics*, *24*(6), 302–306. https://doi.org/10.1055/s-2008-1071561

Leuzzi, V., Mannarelli, D., Manti, F., Pauletti, C., Locuratolo, N., Carducci, C., Carducci, C., Vanacore, N., & Fattapposta, F. (2014). Age-related psychophysiological vulnerability to phenylalanine in phenylketonuria. *Frontiers in Pediatrics*, *2*, 57. https://doi.org/10.3389/fped.2014.00057

Leuzzi, V., Pansini, M., Sechi, E., Chiarotti, F., Carducci, C., Levi, G., & Antonozzi, I. (2004). Executive function impairment in early-treated PKU subjects with normal mental development. *Journal of Inherited Metabolic Disease*, *27*(2), 115–125. https://doi.org/10.1023/B:BOLI.0000028781.94251.1f

Lindauer, A., Seelye, A., Lyons, B., Dodge, H. H., Mattek, N., Mincks, K., Kaye, J., & Erten-Lyons, D. (2017). Dementia care comes home: Patient and caregiver assessment via telemedicine. *Gerontologist*, *57*(5), e85–e93. https://doi.org/10.1093/geront/gnw206

Lochner, K. (2016). Online assessment and online surveys. In *Successful emotions* (pp. 7–20). Springer.

Lou, H. C. (1994). Dopamine precursors and brain function in phenylalanine hydroxylase deficiency. *Acta Paediatrica*, *407*, 86–88. https://doi.org/10.1111/j.1651-2227.1994.tb13461.x

Lou, H. C., Güttler, F., Lykkelund, C., Bruhn, P., & Niederwieser, A. (1985). Decreased vigilance and neurotransmitter synthesis after discontinuation of dietary treatment for

phenylketonuria in adolescents. *European Journal of Pediatrics*, *144*(1), 17–20. https://doi.org/10.1007/BF00491918

Lou, H. C., Lykkelund, C., Gerdes, A. M., Udesen, H., & Bruhn, P. (1987). Increased vigilance and dopamine synthesis by large doses of tyrosine or phenylalanine restriction in phenylketonuria. *Acta Paediatrica Scandinavica*, *76*(4), 560–565. https://doi.org/10.1111/j.1651-2227.1987.tb10521.x

Lou, H. C., Toft, P. B., Andresen, J., Mikkelsen, I., Olsen, B., Guldberg, P., & Güttler, F. (1994). Unchanged MRI of myelin in adolescents with PKU supplied with non-phe essential amino acids after dietary relaxation. *Acta Paediatrica*, *83*(12), 1312–1314. https://doi.org/10.1111/j.1651-2227.1994.tb13025.x

Luciana, M., Sullivan, J., & Nelson, C. A. (2001). Associations between phenylalanine-to-tyrosine ratios and performance on tests of neuropsychological function in adolescents treated early and continuously for phenylketonuria. *Child Development*, *72*(6), 1637–1652. https://doi.org/10.1111/1467-8624.00370

MacDonald, A. (2000). Diet and compliance in phenylketonuria. *European Journal of Pediatrics*, *159*, Suppl. 2, S136–S141. https://doi.org/10.1007/pl00014375

MacDonald, A., Davies, P., Daly, A., Hopkins, V., Hall, S. K., Asplin, D., Hendriksz, C., & Chakrapani, A. (2008). Does maternal knowledge and parent education affect blood phenylalanine control in phenylketonuria? *Journal of Human Nutrition and Dietetics*, *21*(4), 351–358. https://doi.org/10.1111/j.1365-277X.2008.00891.x

MacDonald, A., Rylance, G., Davies, P., Asplin, D., Hall, S. K., & Booth, I. W. (2003). Administration of protein substitute and quality of control in phenylketonuria: A randomized study. *Journal of Inherited Metabolic Disease*, *26*(4), 319–326. https://doi.org/10.1023/a:1025186217369

MacDonald, A., Van Wegberg, A., Ahring, K., Beblo, S., Bélanger-Quintana, A., Burlina, A., Campistol, J., Coşkun, T., Feillet, F., Giżewska, M., Huijbregts, S. C., Leuzzi, V., Maillot, F., Muntau, A. C., Rocha, J. C., Romani, C., Trefz, F., & Giżewska, M. (2020). PKU dietary handbook to accompany PKU guidelines. *Orphanet Journal of Rare Diseases*, *15*(1), 1–21.

MacDonald, A., & White, F. (2015). Amino acid disorders. In SV (Ed.), *Clinical paediatrics dietetics* (pp. 391–456). Wiley-Blackwell.

Macleod, E. L., & Ney, D. M. (2010). Nutritional management of phenylketonuria. *Annales Nestle [English Ed.]*, *68*(2), 58–69. https://doi.org/10.1159/000312813

Magnusson, K. R., Brim, B. L., & Das, S. R. (2010). Selective vulnerabilities of N-methyl-D-aspartate (NMDA) receptors during brain aging. *Frontiers in Aging Neuroscience*, *2*, 11. https://doi.org/10.3389/fnagi.2010.00011

Mainka, T., Fischer, J. F., Huebl, J., Jung, A., Lier, D., Mosejova, A., Skorvanek, M., de Koning, T. J., Kühn, A. A., Freisinger, P., Ziagaki, A., & Ganos, C. (2021). The neurological and neuropsychiatric spectrum of adults with late-treated phenylketonuria. *Parkinsonism and Related Disorders*, *89*, 167–175. https://doi.org/10.1016/j.parkreldis.2021.06.011

Manta-Vogli, P. D., Dotsikas, Y., Loukas, Y. L., & Schulpis, K. H. (2020). The phenylketonuria patient: A recent dietetic therapeutic approach. *Nutritional Neuroscience*, *23*(8), 628–639. https://doi.org/10.1080/1028415X.2018.1538196

Manti, F., Nardecchia, F., Chiarotti, F., Carducci, C., Carducci, C., & Leuzzi, V. (2016). Psychiatric disorders in adolescent and young adult patients with phenylketonuria. *Molecular Genetics and Metabolism*, *117*(1), 12–18. https://doi.org/10.1016/j.ymgme.2015.11.006

Marholin, D., Pohl, R. E., Stewart, R. M., Touchette, P. E., Townsend, N. M., & Kolodny, E. H. (1978). Effects of diet and behavior therapy on social and motor behavior of retarded phenylketonuric adults: An experimental analysis. *Pediatric Research*, *12*(3), 179–187. https://doi.org/10.1203/00006450-197803000-00004

Martynyuk, A. E., Glushakov, A. V., Sumners, C., Laipis, P. J., Dennis, D. M., & Seubert, C. N. (2005). Impaired glutamatergic synaptic transmission in the PKU brain. *Molecular Genetics and Metabolism*, *86*, Suppl. 1, S34–S42. https://doi.org/10.1016/j.ymgme.2005.06.014

Mastrangelo, M., Chiarotti, F., Berillo, L., Caputi, C., Carducci, C., Di Biasi, C., Manti, F., Nardecchia, F., & Leuzzi, V. (2015). The outcome of white matter abnormalities in early treated phenylketonuric patients: A retrospective longitudinal long-term study. *Molecular Genetics and Metabolism*, *116*(3), 171–177. https://doi.org/10.1016/j.ymgme.2015.08.005

Matarazzo, J. (2013). *Wechsler's measurement and appraisal of adult intelligence*. Baltimore, (MD).

Matthews, W. S., Barabas, G., Cusack, E., & Ferrari, M. (1986). Social quotients of children with phenylketonuria before and after discontinuation of dietary therapy. *American Journal of Mental Deficiency*, *91*(1), 92–94.

McAvinue, L. P., Habekost, T., Johnson, K. A., Kyllingsbæk, S., Vangkilde, S., Bundesen, C., & Robertson, I. H. (2012). Sustained attention, attentional selectivity, and attentional capacity across the lifespan. *Attention, Perception and Psychophysics*, *74*(8), 1570–1582. https://doi.org/10.3758/s13414-012-0352-6

McKean, C. M. (1972). The effects of high phenylalanine concentrations on serotinin and catecholamine metabolism in the human brain. *Brain Research*, *47*(2), 469–476. https://doi.org/10.1016/0006-8993(72)90653-1

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://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*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moers, C., Meyer, A., & Janse, E. (2017). Effects of word frequency and transitional probability on word reading durations of younger and older speakers. *Language and Speech*, *60*(2), 289–317. https://doi.org/10.1177/0023830916649215

Monchi, O., Ko, J. H., & Strafella, A. P. (2006). Striatal dopamine release during performance of executive functions: A [11C] raclopride PET study. *Neuroimage*, *33*(3), 907–912. https://doi.org/10.1016/j.neuroimage.2006.06.058

Morgan, C. J., Mofeez, A., Brandner, B., Bromley, L., & Curran, H. V. (2004). Acute effects of ketamine on memory systems and psychotic symptoms in healthy volunteers. *Neuropsychopharmacology*, *29*(1), 208–218. https://doi.org/10.1038/sj.npp.1300342

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*(2), 91–101. https://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*(4), 436–441. https://doi.org/10.1080/13803390600745829

Moyle, J. J., Fox, A. M., Bynevelt, M., Arthur, M., & Burnett, J. R. (2006). Changes in processing speed and executive function after resumption of treatment in adults with phenylketonuria. *Australian Journal of Psychology*, *58*, 228–228.

Muntau, A. C., Burlina, A., Eyskens, F., Freisinger, P., De Laet, C., Leuzzi, V., Rutsch, F., Sivri, H. S., Vijay, S., Bal, M. O., Gramer, G., Pazdírková, R., Cleary, M., Lotz-Havla, A. S., Munafo, A., Mould, D. R., Moreau-Stucker, F., & Rogoff, D. (2017). Efficacy, safety and population pharmacokinetics of sapropterin in PKU patients < 4 years: Results from the SPARK open-label, multicentre, randomized phase IIIb trial. *Orphanet Journal of Rare Diseases*, *12*(1), 47. https://doi.org/10.1186/s13023-017-0600-x

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*(2–3), 84–90. https://doi.org/10.1016/j.ymgme.2015.04.003

Nathan, P. J., Lim, Y. Y., Abbott, R., Galluzzi, S., Marizzoni, M., Babiloni, C., Albani, D., Bartres-Faz, D., Didic, M., Farotti, L., Parnetti, L., Salvadori, N., Müller, B. W., Forloni, G., Girtler, N., Hensch, T., Jovicich, J., Leeuwis, A., Marra, C., . . . & PharmaCog Consortium. (2017). Association between CSF biomarkers, hippocampal volume and cognitive function in patients with amnestic mild cognitive impairment (MCI). *Neurobiology of Aging*, *53*, 1–10. https://doi.org/10.1016/j.neurobiolaging.2017.01.013

Ogawa, S., & Ichinose, H. (2006). Effect of metals and phenylalanine on the activity of human tryptophan hydroxylase-2: Comparison with that on tyrosine hydroxylase activity. *Neuroscience Letters*, *401*(3), 261–265. https://doi.org/10.1016/j.neulet.2006.03.031

Oldendorf, W. H. (1973a). Saturation of blood brain barrier transport of amino acids in phenylketonuria. *Archives of Neurology*, *28*(1), 45–48. https://doi.org/10.1001/archneur.1973.00490190063008

Oldendorf, W. H. (1973b). Stereospecificity of blood–brain barrier permeability to amino acids. *American Journal of Physiology*, *224*(4), 967–969. https://doi.org/10.1152/ajplegacy.1973.224.4.967

Olsson, G. M., Montgomery, S. M., & Alm, J. (2007). Family conditions and dietary control in phenylketonuria. *Journal of Inherited Metabolic Disease*, *30*(5), 708–715. https://doi.org/10.1007/s10545-007-0493-2

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain and Language*, *103*(1–2), 175–176. https://doi.org/10.1016/j.bandl.2007.07.102

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*, 134, 178–189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Palermo, L., Geberhiwot, T., MacDonald, A., Limback, E., Hall, S. K., & Romani, C. (2017). Cognitive outcomes in early-treated adults with phenylketonuria (PKU): A comprehensive picture across domains. *Neuropsychology*, *31*(3), 255–267. https://doi.org/10.1037/neu0000337

Palermo, L., MacDonald, A., Limback, E., Robertson, L., Howe, S., Geberhiwot, T., & Romani, C. (2020). Emotional health in early-treated adults with phenylketonuria (PKU): Relationship with cognitive abilities and blood phenylalanine. *Journal of Clinical and*

Experimental Neuropsychology, *42*(2), 142–159. https://doi.org/10.1080/13803395.2019.1696753

Papenberg, G., Li, S. C., Nagel, I. E., Nietfeld, W., Schjeide, B. M., Schröder, J., Bertram, L., Heekeren, H. R., Lindenberger, U., & Bäckman, L. (2014). Dopamine and glutamate receptor genes interactively influence episodic memory in old age. *Neurobiology of Aging*, *35*(5), 1213.e3–1213.e8. https://doi.org/10.1016/j.neurobiolaging.2013.11.014

Pardridge, W. M. (1998). Blood–brain barrier carrier-mediated transport and brain metabolism of amino acids. *Neurochemical Research*, *23*(5), 635–644. https://doi.org/10.1023/a:1022482604276

Pare, C. M., Sandler, M., & Stacey, R. S. (1957). 5-hydroxytryptamine deficiency in phenylketonuria. *Lancet*, *272*(6968), 551–553. https://doi.org/10.1016/s0140-6736(57)90920-0

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Paus, T., Zijdenbos, A., Worsley, K., Collins, D. L., Blumenthal, J., Giedd, J. N., Rapoport, J. L., & Evans, A. C. (1999). Structural maturation of neural pathways in children and adolescents: In vivo study. *Science*, *283*(5409), 1908–1911. https://doi.org/10.1126/science.283.5409.1908

Pena, M. J., Almeida, M. F., van Dam, E., Ahring, K., Bélanger-Quintana, A., Dokoupil, K., Gokmen-Ozel, H., Lammardo, A. M., MacDonald, A., Robert, M., & Rocha, J. C. (2015). Special low protein foods for phenylketonuria: Availability in Europe and an examination of their nutritional profile. *Orphanet Journal of Rare Diseases*, *10*(1), 162. https://doi.org/10.1186/s13023-015-0378-7

Peng, S. S., Tseng, W. Y., Chien, Y. H., Hwu, W. L., & Liu, H. M. (2004). Diffusion tensor images in children with early-treated, chronic, malignant phenylketonuric: Correlation with intelligence assessment. *AJNR. American Journal of Neuroradiology*, *25*(9), 1569–1574.

Perfect, T. J. (1994). What can Brinley plots tell us about cognitive aging? *Journal of Gerontology*, *49*(2), P60–P64. https://doi.org/10.1093/geronj/49.2.p60

Pietz, J. (1998). Neurological aspects of adult phenylketonuria. *Current Opinion in Neurology*, *11*(6), 679–688. https://doi.org/10.1097/00019052-199812000-00012

Pietz, J., Fätkenheuer, B., Burgard, P., Armbruster, M., Esser, G., & Schmidt, H. (1997). Psychiatric disorders in adult patients with early-treated phenylketonuria. *Pediatrics*, *99*(3), 345–350. https://doi.org/10.1542/peds.99.3.345

Pietz, J., Kreis, R., Rupp, A., Mayatepek, E., Rating, D., Boesch, C., & Bremer, H. J. (1999). Large neutral amino acids block phenylalanine transport into brain tissue in patients

with phenylketonuria. *Journal of Clinical Investigation*, *103*(8), 1169–1178. https://doi.org/10.1172/JCI5017

Pietz, J., Landwehr, R., Kutscha, A., Schmidt, H., de Sonneville, L., & Trefz, F. K. (1995). Effect of high-dose tyrosine supplementation on brain function in adults with phenylketonuria. *Journal of Pediatrics*, *127*(6), 936–943. https://doi.org/10.1016/s0022-3476(95)70031-5

Pietz, J., Schmidt, E., Matthis, P., Kobialka, B., Kutscha, A., & De Sonneville, L. (1993). EEGs in phenylketonuria i: Follow-up to adulthood; ii: Short-term diet-related changes in EEGs and cognitive function. *Developmental Medicine and Child Neurology*, *35*(1), 54–64. https://doi.org/10.1111/j.1469-8749.1993.tb11552.x

Pievani, M., de Haan, W., Wu, T., Seeley, W. W., & Frisoni, G. B. (2011). Functional network disruption in the degenerative dementias. *Lancet. Neurology*, *10*(9), 829–843. https://doi.org/10.1016/S1474-4422(11)70158-2

Pilotto, A., Blau, N., Leks, E., Schulte, C., Deuschl, C., Zipser, C., Piel, D., Freisinger, P., Gramer, G., Kölker, S., Haas, D., Burgard, P., Nawroth, P., Georg, H., Scheffler, K., Berg, D., & Trefz, F. (2019). Cerebrospinal fluid biogenic amines depletion and brain atrophy in adult patients with phenylketonuria. *Journal of Inherited Metabolic Disease*, *42*(3), 398–406. https://doi.org/10.1002/jimd.12049

Plude, D. J., & Doussard-Roosevelt, J. A. (1989). Aging, selective attention, and feature integration. *Psychology and Aging*, *4*(1), 98–105. https://doi.org/10.1037/0882-7974.4.1.98

Prince, M. J. (2015). *World Alzheimer Report 2015: The global impact of dementia: An analysis of prevalence, incidence, cost and trends*. Alzheimer's Disease International.

Pueschel, S. M., Fogelson-Doyle, L., Kammerer, B., & Matsumiya, Y. (1983). Neurophysiological, psychological, and nutritional investigations during discontinuation of the phenylalanine-restricted diet in children with classic phenylketonuria. *Journal of Mental Deficiency Research*, *27*(1), 61–67. https://doi.org/10.1111/j.1365-2788.1983.tb00164.x

Pustejovsky, J. E. (2016). Alternative formulas for the standardized mean difference. https://www.jepusto.com/alternative-formulas-for-the-smd/

Randolph, C., Williams, J. B. W., Hannesdottir, K., Eureyecko, E., Langbaum, J. B., Tariot, P., Farlow, M. R., Galvin, J. E., Langlois, C., Hunt, C., Olsson, T., Poole, M., Weber, C., Boehm, P., Cohen, E., Garzio, L. M., & Alexander, R. (2014). P1-177: Telephone administration of the CDR: Excellent agreement with face-to-face administration. *Alzheimer's and Dementia*, *10*(4S_Part_9), 1–177. https://doi.org/10.1016/j.jalz.2014.05.415

Rao, G., Lopez-Jimenez, F., Boyd, J., D'Amico, F., Durant, N. H., Hlatky, M. A., Howard, G., Kirley, K., Masi, C., Powell-Wiley, T. M., Solomonides, A. E., West, C. P., Wessel, J., & American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Cardiovascular Surgery and Anesthesia; Council on Clinical Cardiology; Council on Functional Genomics and Translational Biology; and Stroke Council. (2017). Methodological standards for meta-analyses and qualitative systematic reviews of cardiac prevention and treatment studies: A scientific statement from the American Heart Association. *Circulation*, *136*(10), e172–e194. https://doi.org/10.1161/CIR.0000000000000523

Raven, J.-C. (1986). Court, JH, and Raven, J.. *Manual for Raven's progressive matrices and vocabulary scales*.

Reeves, S., Bench, C., & Howard, R. (2002). Ageing and the nigrostriatal dopaminergic system. *International Journal of Geriatric Psychiatry*, *17*(4), 359–370. https://doi.org/10.1002/gps.606

Regnault, A., Burlina, A., Cunningham, A., Bettiol, E., Moreau-Stucker, F., Benmedjahed, K., & Bosch, A. M. (2015). Development and psychometric validation of measures to assess the impact of phenylketonuria and its dietary treatment on patients' and parents' quality of life: The phenylketonuria – Quality of life (PKU-QOL) questionnaires. *Orphanet Journal of Rare Diseases*, *10*, 59. https://doi.org/10.1186/s13023-015-0261-6

Rey, F., Abadie, V., Plainguet, F., & Rey, J. (1996). Long-term follow up of patients with classical phenylketonuria after diet relaxation at 5 years of age. The Paris Study. *European Journal of Pediatrics*, *155*, Suppl. 1, S39–S44. https://doi.org/10.1007/pl00014246

Ris, M. D., Williams, S. E., Hunt, M. M., Berry, H. K., & Leslie, N. (1994). Early-treated phenylketonuria: Adult neuropsychologic outcome. *Journal of Pediatrics*, *124*(3), 388–392. https://doi.org/10.1016/s0022-3476(94)70360-4

Roberts, R., & Knopman, D. S. (2013). Classification and epidemiology of MCI. *Clinics in Geriatric Medicine*, *29*(4), 753–772. https://doi.org/10.1016/j.cger.2013.07.003

Robison, M. K., Diede, N. T., Nicosia, J., Ball, B. H., & Bugg, J. M. (2022). A multimodal analysis of sustained attention in younger and older adults. *Psychology and Aging*, *37*(3), 307–325. https://doi.org/10.1037/pag0000687

Rocha, J. C., & MacDonald, A. (2018). Treatment options and dietary supplements for patients with phenylketonuria. *Expert Opinion on Orphan Drugs*, *6*(11), 667–681. https://doi.org/10.1080/21678707.2018.1536541

Rocha, J. C., van Spronsen, F. J., Almeida, M. F., Ramos, E., Guimarães, J. T., & Borges, N. (2013). Early dietary treated patients with phenylketonuria can achieve normal

growth and body composition. *Molecular Genetics and Metabolism*, *110*, Suppl., S40–S43. https://doi.org/10.1016/j.ymgme.2013.10.009

Rohr, S., Arai, H., Mangialasche, F., Matsumoto, N., Peltonen, M., Raman, R., Riedel-Heller, S. G., Sakurai, T., Snyder, H. M., Sugimoto, T., Carrillo, M., Kivipelto, M., Espeland, M. A., & World Wide, F. S. G. (2021). Impact of the COVID-19 pandemic on statistical design and analysis plans for multidomain intervention clinical trials: Experience from World-Wide FINGERS. *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, *7*(1). https://doi.org/10.1002/trc2.12143

Romani. (2014). Order encoding and lexical learning: Contributions to developmental dyslexia. Olson, A, and Tsouknida, E. *Quarterly Journal of Experimental Psychology*, *68*(1), 99–128.

Romani, C. (2018). Cognitive impairments in inherited metabolic diseases: Promises and challenges. In Taylor & Francis.

Romani, C., MacDonald, A., De Felice, S., & Palermo, L. (2018). Speed of processing and executive functions in adults with phenylketonuria: Quick in finding the word, but not the ladybird. *Cognitive Neuropsychology*, *35*(3–4), 171–198. https://doi.org/10.1080/02643294.2017.1320278

Romani, C., Manti, F., Nardecchia, F., Valentini, F., Fallarino, N., Carducci, C., De Leo, S., MacDonald, A., Palermo, L., & Leuzzi, V. (2019). Adult cognitive outcomes in phenylketonuria: Explaining causes of variability beyond average Phe levels. *Orphanet Journal of Rare Diseases*, *14*(1), 273. https://doi.org/10.1186/s13023-019-1225-z

Romani, C., Palermo, L., MacDonald, A., Limback, E., Hall, S. K., & Geberhiwot, T. (2017). The impact of phenylalanine levels on cognitive outcomes in adults with phenylketonuria: Effects across tasks and developmental stages. *Neuropsychology*, *31*(3), 242–254. https://doi.org/10.1037/neu0000336

Rowe CC, Bourgeat P, Ellis KA, Brown B, Lim YY, Mulligan R, Jones G, Maruff P, Woodward M, Price R, Robins P, Tochon-Danguy H, O'Keefe G, Pike KE, Yates P, Szoeke C, Salvado O, Macaulay SL, O'Meara T, Head R, Cobiac L, Savage G, Martins R, Masters CL, Ames D, Villemagne VL. (2013) Predicting Alzheimer disease with β-amyloid imaging: results from the Australian imaging, biomarkers, and lifestyle study of ageing. *Annals of Neurology, 74*(6), 905-13. https://doi.org/10.1002/ana.24040

Rubin, S., Piffer, A. L., Rougier, M. B., Delyfer, M. N., Korobelnik, J. F., Redonnet-Vernhet, I., Marchal, C., Goizet, C., Mesli, S., Gonzalez, C., Gin, H., & Rigalleau, V. (2013). Sight-threatening phenylketonuric encephalopathy in a young adult, reversed by diet. *JIMD Reports*, *10*, 83–85. https://doi.org/10.1007/8904_2012_207

Salthouse, T. A. (1993a). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology*, *84*(2), 171–199. https://doi.org/10.1111/j.2044-8295.1993.tb02472.x

Salthouse, T. A. (1993b). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology*, *48*(1), P29–P36. https://doi.org/10.1093/geronj/48.1.p29

Salthouse, T. A. (1993c). Speed mediation of adult age differences in cognition. *Developmental Psychology*, *29*(4), 722–738. https://doi.org/10.1037/0012-1649.29.4.722

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295x.103.3.403

Salthouse, T. A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence*, *32*(6), 541–561. https://doi.org/10.1016/j.intell.2004.07.003

Salthouse, T. A., Toth, J., Daniels, K., Parks, C., Pak, R., Wolbrette, M., & Hocking, K. J. (2000). Effects of aging on efficiency of task switching in a variant of the trail making test. *Neuropsychology*, *14*(1), 102–111. https://doi.org/10.1037/0894-4105.14.1.102

Sathian, B., Asim, M., Banerjee, I., Pizarro, A. B., Roy, B., van Teijlingen, E. R., do Nascimento, I. J. B., & Alhamad, H. K. (2020). Impact of COVID-19 on clinical trials and clinical research: A systematic review. *Nepal Journal of Epidemiology*, *10*(3), 878–887. https://doi.org/10.3126/nje.v10i3.31622

Scala, I., Riccio, M. P., Marino, M., Bravaccio, C., Parenti, G., & Strisciuglio, P. (2020). Large neutral amino acids (LNAAs) supplementation improves neuropsychological performances in adult patients with phenylketonuria. *Nutrients*, *12*(4). https://doi.org/10.3390/nu12041092

Schmidt, E., Burgard, P., & Rupp, A. (1996). Effects of concurrent phenylalanine levels on sustained attention and calculation speed in patients treated early for phenylketonuria. *European Journal of Pediatrics*, *155*(1), Suppl. 1, S82–S86. https://doi.org/10.1007/pl00014258

Schmidt, H., Mahle, M., Michel, U., & Pietz, J. (1987). Continuation vs discontinuation of low-phenylalanine diet in pku adolescents. *European Journal of Pediatrics*, *146*, Suppl. 1, A17–A19. https://doi.org/10.1007/BF00442050

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*(5), 681–688. https://doi.org/10.1080/01688639408402681

Schuett, V. E., Brown, E. S., & Michals, K. (1985). Reinstitution of diet therapy in PKU patients from twenty-two US clinics. *American Journal of Public Health*, *75*(1), 39–42. https://doi.org/10.2105/ajph.75.1.39

Scriver, C., & Kaufman, S. (2001). Hyperphenylalaninemia: Phenylalanine hydroxylase deficiency. In. In C. R. Scriver, A. L. Beaudet, W. S. Sly & D. Valle (Eds.), *The metabolic and Molecular Bases of Inherited Disease*, *II*. McGraw-Hill.

Seashore, M. R., Friedman, E., Novelly, R. A., & Bapat, V. (1985). Loss of intellectual function in children with phenylketonuria after relaxation of dietary phenylalanine restriction. *Pediatrics*, *75*(2), 226–232. https://doi.org/10.1542/peds.75.2.226

Settle, J. R., Robinson, S. A., Kane, R., Maloni, H. W., & Wallin, M. T. (2015). Remote cognitive assessments for patients with multiple sclerosis: A feasibility study. *Multiple Sclerosis*, *21*(8), 1072–1079. https://doi.org/10.1177/1352458514559296

Sheline, Y. I., Mintun, M. A., Moerlein, S. M., & Snyder, A. Z. (2002). Greater loss of 5-HT2A receptors in midlife than in late life. *American Journal of Psychiatry*, *159*(3), 430–435. https://doi.org/10.1176/appi.ajp.159.3.430

Sigurdsson, S., Aspelund, T., Forsberg, L., Fredriksson, J., Kjartansson, O., Oskarsdottir, B., Jonsson, P. V., Eiriksdottir, G., Harris, T. B., Zijdenbos, A., van Buchem, M. A., Launer, L. J., & Gudnason, V. (2012). Brain tissue volumes in the general population of the elderly: The AGES-Reykjavik study. *Neuroimage*, *59*(4), 3862–3870. https://doi.org/10.1016/j.neuroimage.2011.11.024

Sliwinski, M., & Buschke, H. (1999). Cross-sectional and longitudinal relationships among age, cognition, and processing speed. *Psychology and Aging*, *14*(1), 18–33. https://doi.org/10.1037//0882-7974.14.1.18

Smith, I., Beasley, M. G., & Ades, A. E. (1990). Intelligence and quality of dietary treatment in phenylketonuria. *Archives of Disease in Childhood*, *65*(5), 472–478. https://doi.org/10.1136/adc.65.5.472

Smith, I., Beasley, M. G., Wolff, O. H., & Ades, A. E. (1988). Behavior disturbance in 8-year-old children with early treated phenylketonuria: Report from the MRC/DHSS phenylketonuria Register. *Journal of Pediatrics*, *112*(3), 403–408. https://doi.org/10.1016/s0022-3476(88)80320-2

Smith, I., & Knowles, J. (2000). Behaviour in early treated phenylketonuria: A systematic review. *European Journal of Pediatrics*, *159*, Suppl. 2, S89–S93. https://doi.org/10.1007/pl00014392

Smith, M. L., Klim, P., & Hanley, W. B. (2000). Executive function in schoolaged children with phenylketonuria. *Journal of Developmental and Physical Disabilities*, *12*(4), 317–332. https://doi.org/10.1023/A:1009480013237

Smith, M. L., Klim, P., Mallozzi, E., & Hanley, W. B. (1996). A test of the frontal-specificity hypothesis in the cognitive performance of adults with phenylketonuria. *Developmental Neuropsychology*, *12*(3), 327–341. https://doi.org/10.1080/87565649609540656

Smith, I., Lobascher, M. E., Stevenson, J. E., Wolff, O. H., Schmidt, H., Grubel-Kaiser, S., & Bickel, H. (1978). Effect of stopping low-phenylalanine diet on intellectual progress of children with phenylketonuria. *British Medical Journal*, *2*(6139), 723–726. https://doi.org/10.1136/bmj.2.6139.723

Solomons, G., Keleske, L., & Opitz, E. (1966). Evaluation of the effects of terminating the diet in phenylketonuria. *Journal of Pediatrics*, *69*(4), 596–602. https://doi.org/10.1016/s0022-3476(66)80046-x

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*(2), 225–231. https://doi.org/10.1037//0882-7974.15.2.225

Stillerova, T., Liddle, J., Gustafsson, L., Lamont, R., & Silburn, P. (2016). Could everyday technology improve access to assessments? A pilot study on the feasibility of screening cognition in people with Parkinson's disease using the Montreal Cognitive Assessment via Internet videoconferencing. *Australian Occupational Therapy Journal*, *63*(6), 373–380. https://doi.org/10.1111/1440-1630.12288

Studzinski, C. M., Christie, L. A., Araujo, J. A., Burnham, W. M., Head, E., Cotman, C. W., & Milgram, N. W. (2006). Visuospatial function in the beagle dog: An early marker of cognitive decline in a model of human aging and dementia. *Neurobiology of Learning and Memory*, *86*(2), 197–204. https://doi.org/10.1016/j.nlm.2006.02.005

Sullivan, J. E., & Chang, P. N. (1999). Review: Emotional and behavioral functioning in phenylketonuria. *Journal of Pediatric Psychology*, *24*(3), 281–299. https://doi.org/10.1093/jpepsy/24.3.281

Sundermann, B., Pfleiderer, 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*(3), 711–721. https://doi.org/10.1007/s10545-011-9318-4

Surtees, R., & Blau, N. (2000). The neurochemistry of phenylketonuria. *European Journal of Pediatrics*, *159*(2), Suppl. 2, S109–S113. https://doi.org/10.1007/pl00014370

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language*, *26*(6), 608–631. https://doi.org/10.1016/0749-596X(87)90105-7

Thomas, J., Levy, H., Amato, S., Vockley, J., Zori, R., Dimmock, D., Harding, C. O., Bilder, D. A., Weng, H. H., Olbertz, J., Merilainen, M., Jiang, J., Larimore, K., Gupta, S., Gu, Z., Northrup, H., & PRISM investigators. (2018). Pegvaliase for the treatment of phenylketonuria: Results of a long-term phase 3 clinical trial program

(PRISM). *Molecular Genetics and Metabolism*, *124*(1), 27–38. https://doi.org/10.1016/j.ymgme.2018.03.006

ten Hoedt, A. E., de Sonneville, L. M., Francois, B., ter Horst, N. M., Janssen, M. C., Rubio-Gozalbo, M. E., Wijburg, F. A., Hollak, C. E., & Bosch, A. M. (2011). High phenylalanine levels directly affect mood and sustained attention in adults with phenylketonuria: A randomised, double-blind, placebo-controlled, crossover trial. *Journal of Inherited Metabolic Disease*, *34*(1), 165–171. https://doi.org/10.1007/s10545-010-9253-9

Thiessen, B., Rajput, A. H., Laverty, W., & Desai, H. (1990). Age, environments, and the number of substantia nigra neurons. *Advances in Neurology*, *53*, 201–206.

Thompson-Schill, S. L., D'esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*(3), 513–522. https://doi.org/10.1016/s0896-6273(00)80804-1

Thompson, A. J., Smith, I., Brenton, D., Youl, B. D., Rylance, G., Davidson, D. C., Kendall, B., & Lees, A. J. (1990). Neurological deterioration in young adults with phenylketonuria. *Lancet*, *336*(8715), 602–605. https://doi.org/10.1016/0140-6736(90)93401-a

Thompson, A. J., Tillotson, S., Smith, I., Kendall, B., Moore, S. G., & Brenton, D. P. (1993). Brain MRI changes in phenylketonuria: Associations with dietary status. *Brain*, *116*(4), 811–821. https://doi.org/10.1093/brain/116.4.811

Tooyama, I., McGeer, E. G., Kawamata, T., Kimura, H., & McGeer, P. L. (1994). Retention of basic fibroblast growth factor immunoreactivity in dopaminergic neurons of the substantia nigra during normal aging in humans contrast with loss in Parkinson's disease. *Brain Research*, *656*(1), 165–168. https://doi.org/10.1016/0006-8993(94)91378-1

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology. Human Perception and Performance*, *16*(3), 459–478. https://doi.org/10.1037//0096-1523.16.3.459

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. *Attention and Performance*, *VI*, 333.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tuten, T. L., Urban, D. J., & Bosnjak, M. (2002). Internet surveys and data quality: A review *Online social sciences, 1*, 7–26.

Umucu, E., Tansey, T. N., Brooks, J., & Lee, B. J. R. C. B. (2021). The protective role of character strengths in COVID-19 stress and well-being in individuals with chronic conditions and disabilities: An exploratory study. *Rehabilitation Counseling Bulletin*, *64*(2), 67–74. https://doi.org/10.1177/0034355220967093

Vallesi, A., Tronelli, V., Lomi, F., & Pezzetta, R. (2021). Age differences in sustained attention tasks: A meta-analysis. *Psychonomic Bulletin and Review*, 28(6), 1755–1775. https://doi.org/10.3758/s13423-021-01908-x

van Spronsen, F. J., Hoeksma, M., & Reijngoud, D. J. (2009). Brain dysfunction in phenylketonuria: Is phenylalanine toxicity the only possible cause? *Journal of Inherited Metabolic Disease*, *32*(1), 46–51. https://doi.org/10.1007/s10545-008-0946-2

van Spronsen, F. J., Huijbregts, S. C. J., Bosch, A. M., & Leuzzi, V. (2011). Cognitive, neurophysiological, neurological and psychosocial outcomes in early-treated PKU-patients: A start toward standardized outcome measurement across development. *Molecular Genetics and Metabolism*, *104*, Suppl., S45–S51. https://doi.org/10.1016/j.ymgme.2011.09.036

van Spronsen, F. J., van Wegberg, A. M., Ahring, K., Bélanger-Quintana, A., Blau, N., Bosch, A. M., Burlina, A., Campistol, J., Feillet, F., Giżewska, M., Huijbregts, S. C., Kearney, S., Leuzzi, V., Maillot, F., Muntau, A. C., Trefz, F. K., van Rijn, M., Walter, J. H., & MacDonald, A. (2017). Key European guidelines for the diagnosis and management of patients with phenylketonuria. *Lancet. Diabetes and Endocrinology*, *5*(9), 743–756. https://doi.org/10.1016/S2213-8587(16)30320-5

van Vliet, D., van Wegberg, A. M. J., Ahring, K., Bik-Multanowski, M., Blau, N., Bulut, F. D., Casas, K., Didycz, B., Djordjevic, M., Federico, A., Feillet, F., Gizewska, M., Gramer, G., Hertecant, J. L., Hollak, C. E. M., Jørgensen, J. V., Karall, D., Landau, Y., Leuzzi, V., . . . & van Spronsen, F. J. (2018). Can untreated PKU patients escape from intellectual disability? A systematic review. *Orphanet Journal of Rare Diseases*, *13*(1), 149. https://doi.org/10.1186/s13023-018-0890-7

Van Wegberg, A. M. J., MacDonald, A., Ahring, K., Bélanger-Quintana, A., Blau, N., Bosch, A. M., Burlina, A., Campistol, J., Feillet, F., Giżewska, M., Huijbregts, S. C., Kearney, S., Leuzzi, V., Maillot, F., Muntau, A. C., van Rijn, M., Trefz, F., Walter, J. H., & van Spronsen, F. J. (2017). The complete European guidelines on phenylketonuria: Diagnosis and treatment. *Orphanet Journal of Rare Diseases*, *12*(1), 162. https://doi.org/10.1186/s13023-017-0685-2

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging*, *13*(1), 120–126. https://doi.org/10.1037//0882-7974.13.1.120

Vernooij, M. W., Ikram, M. A., Tanghe, H. L., Vincent, A. J., Hofman, A., Krestin, G. P., Niessen, W. J., Breteler, M. M., & van der Lugt, A. (2007). Incidental findings on brain MRI in the general population. *New England Journal of Medicine*, *357*(18), 1821–1828. https://doi.org/10.1056/NEJMoa070972

Villasana, D., Butler, I. J., Williams, J. C., & Roongta, S. M. (1989). Neurological deterioration in adult phenylketonuria. *Journal of Inherited Metabolic Disease*, *12*(4), 451–457. https://doi.org/10.1007/BF01802042

Volkow, N. D., Gur, R. C., Wang, G. J., Fowler, J. S., Moberg, P. J., Ding, Y. S., Hitzemann, R., Smith, G., & Logan, J. (1998). Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. *American Journal of Psychiatry*, *155*(3), 344–349. https://doi.org/10.1176/ajp.155.3.344

Waisbren, S. E., Noel, K., Fahrbach, K., Cella, C., Frame, D., Dorenbaum, A., & Levy, H. (2007). Phenylalanine blood levels and clinical outcomes in phenylketonuria: A systematic literature review and meta-analysis. *Molecular Genetics and Metabolism*, *92*(1–2), 63–70. https://doi.org/10.1016/j.ymgme.2007.05.006

Waisbren, S. E., & Zaff, J. (1994). Personality disorder in young women with treated phenylketonuria. *Journal of Inherited Metabolic Disease*, *17*(5), 584–592. https://doi.org/10.1007/BF00711596

Walter, J. H., & White, F. J. (2004). Blood phenylalanine control in adolescents with phenylketonuria. *International Journal of Adolescent Medicine and Health*, *16*(1), 41–45. https://doi.org/10.1515/ijamh.2004.16.1.41

Walter, J. H., White, F. J., Hall, S. K., MacDonald, A., Rylance, G., Boneh, A., Francis, D. E., Shortland, G. J., Schmidt, M., & Vail, A. (2002). How practical are recommendations for dietary control in phenylketonuria? *Lancet*, *360*(9326), 55–57. https://doi.org/10.1016/s0140-6736(02)09334-0

Walter, J. H., White, F., Wraith, J. E., Jenkins, J. P., & Wilson, B. P. M. (1997). Complete reversal of moderate/severe brain MRI abnormalities in a patient with classical phenylketonuria. *Journal of Inherited Metabolic Disease*, *20*(3), 367–369. https://doi.org/10.1023/A:1005330012574

Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., Ho, C. S., & Ho, R. C. (2020). Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *International Journal of Environmental Research and Public Health*, *17*(5), 1729. https://doi.org/10.3390/ijerph17051729

Wang, Y., Chan, G. L., Holden, J. E., Dobko, T., Mak, E., Schulzer, M., Huser, J. M., Snow, B. J., Ruth, T. J., Calne, D. B., & Stoessl, A. J. (1998). Age-dependent decline of dopamine D1 receptors in human brain: A PET study. *Synapse*, *30*(1), 56–61. https://doi.org/10.1002/(SICI)1098-2396(199809)30:1<56::AID-SYN7>3.0.CO;2-J

Wardlaw, J. M., Valdés Hernández, M. C., & Muñoz-Maniega, S. (2015). What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. *Journal*

of the American Heart Association, *4*(6), 001140. https://doi.org/10.1161/JAHA.114.001140

Ware, Jr., J. E., & Sherbourne, C. D. J. M. (1992). The MOS 36-item short-form health survey (SF-36): I. *Medical Care*, 30(6), 473–483.

Weglage, J., Fromm, J., Van Teeffelen-Heithoff, A., Möller, H. E., Koletzko, B., Marquardt, T., Rutsch, F., & Feldmann, R. (2013). Neurocognitive functioning in adults with phenylketonuria: Results of a long term study. *Molecular Genetics and Metabolism*, *110*, Suppl., S44–S48. https://doi.org/10.1016/j.ymgme.2013.08.013

Weglage, J., Pietsch, M., Fünders, B., Koch, H. G., & Ullrich, K. (1995). Neurological findings in early treated phenylketonuria. *Acta Paediatrica*, *84*(4), 411–415. https://doi.org/10.1111/j.1651-2227.1995.tb13661.x

Wen, W., & Sachdev, P. (2004). The topography of white matter hyperintensities on brain MRI in healthy 60- to 64-year-old individuals. *Neuroimage*, *22*(1), 144–154. https://doi.org/10.1016/j.neuroimage.2003.12.027

Weschler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. Psychological Corporation.

Wesonga, E., Shimony, J. S., Rutlin, J., Grange, D. K., & White, D. A. (2016). Relationship between age and white matter integrity in children with phenylketonuria. *Molecular Genetics and Metabolism Reports*, *7*, 45–49. https://doi.org/10.1016/j.ymgmr.2016.03.003

Weston, P. S., Nicholas, J. M., Lehmann, M., Ryan, N. S., Liang, Y., Macpherson, K., Modat, M., Rossor, M. N., Schott, J. M., Ourselin, S., & Fox, N. C. (2016). Presymptomatic cortical thinning in familial Alzheimer disease: A longitudinal MRI study. *Neurology*, *87*(19), 2050–2057. https://doi.org/10.1212/WNL.0000000000003322

White, D. A., Antenor-Dorsey, J. A., Grange, D. K., Hershey, T., Rutlin, J., Shimony, J. S., McKinstry, R. C., & Christ, S. E. (2013). White matter integrity and executive abilities following treatment with tetrahydrobiopterin (BH4) in individuals with phenylketonuria. *Molecular Genetics and Metabolism*, *110*(3), 213–217. https://doi.org/10.1016/j.ymgme.2013.07.010

White, D. A., Nortz, M. J., Mandernach, T., Huntington, K., & Steiner, R. D. (2002). Age-related working memory impairments in children with prefrontal dysfunction associated with phenylketonuria. *Journal of the International Neuropsychological Society*, *8*(1), 1–11. https://doi.org/10.1017/S135561770102001X

Williams, K. (1998). Benefits of normalizing plasma phenylalanine: Impact on behaviour and health. A case report. *Journal of Inherited Metabolic Disease*, *21*(8), 785–790. https://doi.org/10.1023/a:1005482732411

Wingfield, A., Tun, P. A., Koh, C. K., & Rosen, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology and Aging*, *14*(3), 380–389. https://doi.org/10.1037//0882-7974.14.3.380

Wolfe, J. M., & Cave, K. R. (1990). Deploying visual attention: The guided search model. *Ai and the Eye*, 79–103.

Wood, Jr., A. C., Friedman, C. J., & Steisel, I. M. (1967). Psychosocial factors in phenylketonuria. *American Journal of Orthopsychiatry*, *37*(4), 671–679. https://doi.org/10.1111/j.1939-0025.1967.tb00508.x

Woodcock, R. W., Johnson, M. B., & Mather, N. (1990). *Woodcock-Johnson psycho-educational battery—Revised*. DLM Teaching Resources.

Wright, S. W., & Tarjan, G. (1957). Phenylketonuria. *AMA Journal of Diseases of Children*, *93*(4), 405–419. https://doi.org/10.1001/archpedi.1957.02060040407009

Yannicelli, S., & Ryan, A. (1995). Improvements in behaviour and physical manifestations in previously untreated adults with phenylketonuria using a phenylalanine-restricted diet: A national survey. *Journal of Inherited Metabolic Disease*, *18*(2), 131–134. https://doi.org/10.1007/BF00711747

Yuwiler, A., Geller, E., & Slater, G. G. (1965). On the mechanism of the brain serotonin depletion in experimental phenylketonuria. *Journal of Biological Chemistry*, *240*(3), 1170–1174. https://doi.org/10.1016/S0021-9258(18)97557-5

Zagreda, L., Goodman, J., Druin, D. P., McDonald, D., & Diamond, A. (1999). Cognitive deficits in a genetic mouse model of the most common biochemical cause of human mental retardation. *Journal of Neuroscience*, *19*(14), 6175–6182. https://doi.org/10.1523/JNEUROSCI.19-14-06175.1999

Zeghari, R., Guerchouche, R., Tran Duc, M., Bremond, F., Lemoine, M. P., Bultingaire, V., Langel, K., De Groote, Z., Kuhn, F., Martin, E., Robert, P., & König, A. (2021). Pilot study to assess the feasibility of a mobile unit for remote cognitive screening of isolated elderly in rural areas. *International Journal of Environmental Research and Public Health*, *18*(11), 6108. https://doi.org/10.3390/ijerph18116108

Appendices

Appendix A. Beck Anxiety Inventory Questionnaire.

Beck Anxiety Inventory (BAI)

Below is a list of common symptoms of anxiety. Please carefully read each item in the list. Indicate how much you have been bothered by that symptom during the past month, including today, by circling the number in the corresponding space in the column next to each symptom.

	Not at all	Mildly, but it didn't bother me much	Moderately – it wasn't pleasant at times	Severely – it bothered me a lot
Numbness or tingling	0	1	2	3
Feeling hot	0	1	2	3
Wobbliness in legs	0	1	2	3
Unable to relax	0	1	2	3
Fear of worst happening	0	1	2	3
Dizzy or lightheaded	0	1	2	3
Heart pounding / racing	0	1	2	3
Unsteady	0	1	2	3
Terrified or afraid	0	1	2	3
Nervous	0	1	2	3
Feeling of choking	0	1	2	3
Hands trembling	0	1	2	3
Shaky / unsteady	0	1	2	3
Fear of losing control	0	1	2	3
Difficulty in breathing	0	1	2	3
Fear of dying	0	1	2	3
Scared	0	1	2	3
Indigestion	0	1	2	3
Faint / lightheaded	0	1	2	3
Face flushed	0	1	2	3
Hot / cold sweats	0	1	2	3

Appendix B. Beck Depression Inventory Questionnaire.

Please read each group of statements carefully, then fill in the circle next to the statement in each group that best describes the way you have been feeling **over the past two weeks, including today**.
Be sure to read all the statements in each group before making your choice.
Then move on to the next group of sentences.

Beck's Depression Inventory

This depression inventory can be self-scored. The scoring scale is at the end of the questionnaire.

1.
- 0 I do not feel sad.
- 1 I feel sad
- 2 I am sad all the time and I can't snap out of it.
- 3 I am so sad and unhappy that I can't stand it.

2.
- 0 I am not particularly discouraged about the future.
- 1 I feel discouraged about the future.
- 2 I feel I have nothing to look forward to.
- 3 I feel the future is hopeless and that things cannot improve.

3.
- 0 I do not feel like a failure.
- 1 I feel I have failed more than the average person.
- 2 As I look back on my life, all I can see is a lot of failures.
- 3 I feel I am a complete failure as a person.

4.
- 0 I get as much satisfaction out of things as I used to.
- 1 I don't enjoy things the way I used to.
- 2 I don't get real satisfaction out of anything anymore.
- 3 I am dissatisfied or bored with everything.

5.
- 0 I don't feel particularly guilty
- 1 I feel guilty a good part of the time.
- 2 I feel quite guilty most of the time.
- 3 I feel guilty all of the time.

6.
- 0 I don't feel I am being punished.
- 1 I feel I may be punished.
- 2 I expect to be punished.
- 3 I feel I am being punished.

7.
- 0 I don't feel disappointed in myself.
- 1 I am disappointed in myself.
- 2 I am disgusted with myself.
- 3 I hate myself.

8.
- 0 I don't feel I am any worse than anybody else.
- 1 I am critical of myself for my weaknesses or mistakes.
- 2 I blame myself all the time for my faults.
- 3 I blame myself for everything bad that happens.

9.
- 0 I don't have any thoughts of killing myself.
- 1 I have thoughts of killing myself, but I would not carry them out.
- 2 I would like to kill myself.
- 3 I would kill myself if I had the chance.

10.
- 0 I don't cry any more than usual.
- 1 I cry more now than I used to.
- 2 I cry all the time now.
- 3 I used to be able to cry, but now I can't cry even though I want to.

11.

0	I am no more irritated by things than I ever was.
1	I am slightly more irritated now than usual.
2	I am quite annoyed or irritated a good deal of the time.
3	I feel irritated all the time.

12.

0	I have not lost interest in other people.
1	I am less interested in other people than I used to be.
2	I have lost most of my interest in other people.
3	I have lost all of my interest in other people.

13.

0	I make decisions about as well as I ever could.
1	I put off making decisions more than I used to.
2	I have greater difficulty in making decisions more than I used to.
3	I can't make decisions at all anymore.

14.

0	I don't feel that I look any worse than I used to.
1	I am worried that I am looking old or unattractive.
2	I feel there are permanent changes in my appearance that make me look unattractive
3	I believe that I look ugly.

15.

0	I can work about as well as before.
1	It takes an extra effort to get started at doing something.
2	I have to push myself very hard to do anything.
3	I can't do any work at all.

16.

0	I can sleep as well as usual.
1	I don't sleep as well as I used to.
2	I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.
3	I wake up several hours earlier than I used to and cannot get back to sleep.

17.

0	I don't get more tired than usual.
1	I get tired more easily than I used to.
2	I get tired from doing almost anything.
3	I am too tired to do anything.

18.

0	My appetite is no worse than usual.
1	My appetite is not as good as it used to be.
2	My appetite is much worse now.
3	I have no appetite at all anymore.

19.

0	I haven't lost much weight, if any, lately.
1	I have lost more than five pounds.
2	I have lost more than ten pounds.
3	I have lost more than fifteen pounds.

20.

0	I am no more worried about my health than usual.
1	I am worried about physical problems like aches, pains, upset stomach, or constipation.
2	I am very worried about physical problems and it's hard to think of much else.
3	I am so worried about my physical problems that I cannot think of anything else.

21.

0	I have not noticed any recent change in my interest in sex.
1	I am less interested in sex than I used to be.
2	I have almost no interest in sex.
3	I have lost interest in sex completely.

Appendix C. PKU Quality of Life Questionnaire.

PKU-QOL

PKU and its effect on your everyday life - Adults -

Before filling in this questionnaire, please write today's date

/__/__/ /__/__/ /__/__/__/__/
day month year

- This questionnaire was designed with the help of people with Phenylketonuria (PKU). It was developed in order to understand the effect of PKU and its treatment on your everyday life.
- By supplement or amino acid mixture, we mean the product or liquid (for example milk, powder) prescribed by your doctor.
- Unless otherwise instructed, please **think about the past 7 days** when answering these questions.
- Please fill out this questionnaire in a quiet area and, if possible, by yourself. If you do not understand a question, please ask for some help.
- If you do not know how to answer, please choose the response that best applies to you.
- Please tell us about your experience with PKU; there are no "right" or "wrong" answers.
- Please take all the time you need.

Thank you very much for your participation.

The information in this questionnaire will remain strictly confidential and anonymous.

YOUR HEALTH

Following are some things that people with PKU may experience. For each sentence, please tick the box that best applied to you.

1. In the past 7 days, compared to others my age, I think my health in general was:

Poor	Fair	Good	Very good	Excellent
☐0	☐1	☐2	☐3	☐4

2a. In the past 7 days, I had headaches

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **2b. If you had this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

3a. In the past 7 days, I had stomach aches

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **3b. If you had this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

4a. In the past 7 days, I felt tired during the day

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **4b. If you felt this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

YOUR HEALTH (continued)

5a. In the past 7 days, I had trouble concentrating

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **5b. If you had this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

6a. In the past 7 days, I needed longer to think about things than others my age do

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **6b. If you experienced this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

7a. In the past 7 days, my hands were shaky

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **7b. If you had this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

8a. In the past 7 days, I was irritable

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **8b. If you were irritable, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

9a. In the past 7 days, I became aggressive

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **9b. If you became aggressive, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

YOUR HEALTH (continued)

10a. In the past 7 days, I had mood swings

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **10b. If you had this, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

11a. In the past 7 days, I felt sad

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **11b. If you felt sad, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

12a. In the past 7 days, I felt anxious

Never	A little of the time	Sometimes	Often	Very often
☐0	☐1	☐2	☐3	☐4

↳ **12b. If you felt anxious, do you think it was related to PKU?**

Yes	No	I don't know
☐0	☐1	☐2

YOUR PKU DIET AND SUPPLEMENTS

Following are some things that people may think about their PKU diet and supplements (formula or medical food). For each sentence, please tick the box that best applied to you.

By supplement or amino acid mixture, we mean the product or liquid (for example milk, powder) prescribed by your doctor.

13. In the past 7 days, it was hard to take my supplements several times a day

Not at all	A little	Somewhat	Very	Extremely	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

14. In the past 7 days, it was hard to only eat or drink what I should

Not at all	A little	Somewhat	Very	Extremely	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

15. In the past 7 days, I felt different because I couldn't eat or drink what others ate

Not at all	A little	Somewhat	Very	Extremely	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

16. In the past 7 days, my supplements tasted...

Very good	Good	Ok	Bad	Very bad	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

17. In the past 7 days, my PKU low protein food (for example biscuits, pasta, cereals, bread etc.) tasted...

Very good	Good	Ok	Bad	Very bad	I didn't eat PKU low protein food
☐0	☐1	☐2	☐3	☐4	☐5

18. In the past 7 days, it was annoying to weigh or measure, or estimate protein (exchanges) in my food

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

19. In the past 7 days, my family or my partner and I argued because of the PKU diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

YOUR PKU DIET AND SUPPLEMENTS (continued)

20. In the past 7 days, it was hard to see others eating food that I couldn't eat

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

21. In the past 7 days, I thought about eating food that was not in my diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

22. In the past 7 days, I wanted to eat things others could eat

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

23. In the past 7 days, I felt embarrassed about following my PKU diet in front of others

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

24. In the past 7 days, I felt left out because of the PKU diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

25. In the past 7 days, I felt less spontaneous (less able to do something unplanned) because of the PKU diet (for example eating out, visiting friends or family…)

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

26. In the past 7 days, it was hard for me to eat out because of my PKU diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

27. In the past 7 days, I followed my PKU diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

YOUR PKU DIET AND SUPPLEMENTS (continued)

28. In the past 7 days, I enjoyed eating even though I am following a PKU diet

Never	A little of the time	Sometimes	Most of the time	Always	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

29. In the past 7 days, my family or my partner and I argued because of my supplements

Never	A little of the time	Sometimes	Most of the time	Always	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

30. In the past 7 days, I felt embarrassed about taking my PKU supplements in front of others

Never	A little of the time	Sometimes	Most of the time	Always	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

31. In the past 7 days, I felt less spontaneous (less able to do something unplanned) because of the supplements (for example eating out, visiting friends or family…)

Never	A little of the time	Sometimes	Most of the time	Always	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

32. In the past 7 days, it was hard for me to eat out because of my supplements

Never	A little of the time	Sometimes	Most of the time	Always	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

33. In the past 7 days, I missed taking some supplements

Never	1 or 2 times	3 to 5 times	6 or 7 times	More than 7 times	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

34. In the past 7 days, I ate or drank something that I was not supposed to

Never	1 or 2 times	3 to 5 times	6 or 7 times	More than 7 times	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

35. In the past 7 days, I chose to eat or drink things I was not supposed to

Never	1 or 2 times	3 to 5 times	6 or 7 times	More than 7 times	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

YOUR DAILY LIFE WITH PKU

Following are some things that people may experience when living with PKU. For each sentence, please tick the box that best applied to you. If you do not plan or cook your meals, please tick the 'Does not apply' box.

36. In the past 7 days, it was easy to plan my meals in advance

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

37. In the past 7 days, it was time-consuming to prepare PKU meals (weighing, measuring, cooking)

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

38. In the past 7 days, it was easy to cook special PKU low protein food

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

39. In the past 7 days, it was hard to cook food for others that I can't eat myself

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

40. In the past 7 days, it was time consuming to do PKU-related management tasks (filling in forms, making phone calls, reimbursements, etc)

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

41. In the past 7 days, I missed taking supplements because of work constraints

Never	A little of the time	Sometimes	Most of the time	Always	I don't work because of PKU	I don't work	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5	☐6	☐7

42. In the past 7 days, I ate or drank something I shouldn't have because of work constraints

Never	A little of the time	Sometimes	Most of the time	Always	I don't work because of PKU	I don't work	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5	☐6	☐7

43. In the past 7 days, PKU impacted my work life (for example concentrating on work, getting things done, remembering things…)

Never	A little of the time	Sometimes	Most of the time	Always	I don't work because of PKU	I don't work
☐0	☐1	☐2	☐3	☐4	☐5	☐6

YOUR DAILY LIFE WITH PKU (continued)

44. In the past 7 days, I ate or drank something I shouldn't have because of college/university constraints

Never	A little of the time	Sometimes	Most of the time	Always	I don't go to college/ university	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5	☐6

45. In the past 7 days, it was hard to do everything I needed to do for my PKU

Never	A little of the time	Sometimes	Most of the time	Always
☐0	☐1	☐2	☐3	☐4

46. It is inconvenient to carry my supplements with me when I am going on business trips or on holiday

Not at all	A little	Somewhat	Very	Extremely	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

47. It is inconvenient to carry my PKU food with me when I am going on business trips or on holiday

Not at all	A little	Somewhat	Very	Extremely	I don't eat any special food
☐0	☐1	☐2	☐3	☐4	☐5

YOUR FEELINGS IN GENERAL ABOUT PKU

Following are some things that people may feel related to their PKU. For each sentence, please tick the box that best applies to how you feel in general. If you do not have blood tests, please tick the 'Does not apply' box.

Generally:

48. I am afraid of having blood taken from my arm

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

49. I am afraid of having blood taken from my finger (finger prick)

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

50. PKU negatively impacts my relationship with my partner

Not at all	A little	Somewhat	Very	Extremely	Does not apply
☐0	☐1	☐2	☐3	☐4	☐5

51. I am worried that my Phe levels are high

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

52. PKU expenses (for prescription co-payments or special food) negatively impact my daily life

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

53. Visiting my doctor/dietician for PKU bothers me

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

54. Having PKU makes me angry

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

55. It is hard having constantly to explain PKU to others

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

56. It is hard to talk with my family or my partner about PKU

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

YOUR FEELINGS IN GENERAL ABOUT PKU (continued)

Generally:

57. It is hard for me to make friends because of my PKU

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

58. I am worried about how PKU might affect my future health

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

59. I am afraid of passing PKU on to my future children

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

60. It is easy to live with PKU

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

61. I like the way I am

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

62. I am confident that I have enough information about PKU and its treatment

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

63. *For women only:* I worry that high Phe levels during pregnancy might cause problems for any future children I might have

Not at all	A little	Somewhat	Very	Extremely
☐0	☐1	☐2	☐3	☐4

64. I feel guilty if I miss taking a supplement

Not at all	A little	Somewhat	Very	Extremely	I don't take a supplement
☐0	☐1	☐2	☐3	☐4	☐5

YOUR FEELINGS IN GENERAL ABOUT PKU (continued)

Generally:

65. I feel guilty after eating or drinking something I am not supposed to

Not at all	A little	Somewhat	Very	Extremely	I don't follow a PKU diet
☐0	☐1	☐2	☐3	☐4	☐5

Please check that you have answered all the questions.

Thank you for taking the time to answer these questions.

Appendix D. SF36 Health Questionnaire.

SF-36 QUESTIONNAIRE

Name:_____ **Ref. Dr:**_____ **Date:** _____

ID#: _____ **Age:** _____ **Gender: M / F**

Please answer the 36 questions of the **Health Survey** completely, honestly, and without interruptions.

GENERAL HEALTH:
In general, would you say your health is:
◯ Excellent ◯ Very Good ◯ Good ◯ Fair ◯ Poor

Compared to one year ago, how would you rate your health in general now?
◯ Much better now than one year ago
◯ Somewhat better now than one year ago
◯ About the same
◯ Somewhat worse now than one year ago
◯ Much worse than one year ago

LIMITATIONS OF ACTIVITIES:
The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports.
◯ Yes, Limited a lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Lifting or carrying groceries
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Climbing several flights of stairs
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Climbing one flight of stairs
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Bending, kneeling, or stooping
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Walking more than a mile
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Walking several blocks
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

Walking one block
◯ Yes, Limited a Lot ◯ Yes, Limited a Little ◯ No, Not Limited at all

IRAS ID 255820, V1, 01/07/19

Bathing or dressing yourself

Yes, Limited a Lot | Yes, Limited a Little | No, Not Limited at all

PHYSICAL HEALTH PROBLEMS:

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

Cut down the amount of time you spent on work or other activities

Yes | No

Accomplished less than you would like

Yes | No

Were limited in the kind of work or other activities

Yes | No

Had difficulty performing the work or other activities (for example, it took extra effort)

Yes | No

EMOTIONAL HEALTH PROBLEMS:

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

Cut down the amount of time you spent on work or other activities

Yes | No

Accomplished less than you would like

Yes | No

Didn't do work or other activities as carefully as usual

Yes | No

SOCIAL ACTIVITIES:

Emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

Not at all | Slightly | Moderately | Severe | Very Severe

PAIN:

How much bodily pain have you had during the past 4 weeks?

None | Very Mild | Mild | Moderate | Severe | Very Severe

During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all | A little bit | Moderately | Quite a bit | Extremely

ENERGY AND EMOTIONS:
These questions are about how you feel and how things have been with you during the last 4 weeks. For each question, please give the answer that comes closest to the way you have been feeling.

Did you feel full of pep?
- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Have you been a very nervous person?
- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Have you felt so down in the dumps that nothing could cheer you up?
- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Have you felt calm and peaceful?
- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Did you have a lot of energy?
- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Have you felt downhearted and blue?

- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Did you feel worn out?

- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Have you been a happy person?

- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

Did you feel tired?

- ◯ All of the time
- ◯ Most of the time
- ◯ A good Bit of the Time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

SOCIAL ACTIVITIES:
During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little bit of the time
- ◯ None of the Time

GENERAL HEALTH:
How true or false is each of the following statements for you?

I seem to get sick a little easier than other people
Definitely true | Mostly true | Don't know | Mostly false | Definitely false

I am as healthy as anybody I know
Definitely true | Mostly true | Don't know | Mostly false | Definitely false

I expect my health to get worse
Definitely true | Mostly true | Don't know | Mostly false | Definitely false

My health is excellent
Definitely true | Mostly true | Don't know | Mostly false | Definitely false

Appendix E. COVID-19 Questionnaire

Q1 Participant Number:

_____

Q2 Please rate how you feel since the UK-wide lockdown was declared on 23rd March 2020:

	Significantly more (3)	Moderately more (2)	Slightly more (1)	No more or less (0)	Slightly less (-1)	Moderately less (-2)	Significantly less (-3)
Stressed	○	○	○	○	○	○	○
Isolated	○	○	○	○	○	○	○
Anxious	○	○	○	○	○	○	○
Exhausted	○	○	○	○	○	○	○
Sad	○	○	○	○	○	○	○
Worried for my health	○	○	○	○	○	○	○
Short-tempered / angry	○	○	○	○	○	○	○

Q3 Please rate how the UK-wide lockdown has impacted your ability to manage your PKU:
(If you are not a patient at a PKU clinic, please select "Not applicable")

	Significantly more difficult (3)	Moderately more difficult (2)	Slightly more difficult (1)	No more or less difficult (0)	Slightly less difficult (-1)	Moderately less difficult (-2)	Significantly less difficult (-3)	Not applicable
General management of my PKU has become (1)	○	○	○	○	○	○	○	○
Accessing my supplements has become (2)	○	○	○	○	○	○	○	○

Appendix F. Older Control Data Cleaning Tables

Table 3.A1 - *Across-participant non-valid data cleaning; Number of participants removed with ≥20% non-valid RTs.*

	Number of participants removed			
	Picture naming	**Word reading**	**Non-word reading**	**Stroop**
Younger Controls	1	4	8	2
AwPKU	1	4	6	1
Older Controls	10	7	1	2

Table 3.A2 - *Within-participant reaction time data cleaning. Mean number of errors removed for each task.*

	Mean number removed RT responses (errors)										
	Choice reaction	**Simple detection**	**Detection with distractors**	**Attentional switch**	**Feature search**	**Conjunction search**	**RVP**	**Stroop**	**Picture naming**	**Word reading**	**Non-word reading**
Younger Controls	1.6	0.0	1.2	1.0	1.4	1.9	9.3	1.8	11.2	1.7	2.7
AwPKU	1.4	0.0	1.3	1.0	1.6	1.5	11.6	2.0	12.2	2.1	2.6
Older Controls	1.9	0.0	1.2	1.0	1.1	2.3	15	1.0	9.8	2.0	2.2
Overall mean	1.9	0.0	1.2	1.0	1.1	2.2	14.8	1.0	9.4	2.0	2.3

Table 3.A3 - *Within-participant reaction time data cleaning. Mean number of RT outliers removed for each task: responses <100ms or ±3SD from participant mean.*

	Mean number removed RT responses (non-valid responses)										
	Choice reaction	**Simple detection**	**Detection with distractors**	**Attentional switch**	**Feature search**	**Conjunction search**	**RVP**	**Stroop**	**Picture naming**	**Word reading**	**Non-word reading**
Younger Controls	1.7	1.0	1.0	1.0	1.1	1.1	2.7	3.8	8.8	6.5	2.3
AwPKU	1.6	1.0	1.0	1.0	1.2	1.0	1.0	3.8	8.3	6.2	1.8
Older Controls	1.7	1.0	1.0	1.0	1.3	0.9	2.5	2.4	13.2	5.1	1.1
Overall mean	1.7	1.0	1.0	1.0	1.3	0.9	2.5	2.4	13.0	5.2	1.1

Appendix G. Additional analyses of fast and slow sub-groups for OC and AwPKU cohorts

Participants within experimental groups in Chapter 4 were split into 'fast' and 'slow' sub-groups using a performance-based median split. Fast and slow participants were identified by averaging scores across conditions within each task, and then sorting participants into the fastest 50% and the slowest 50%. To demonstrate appropriate rigour in analyses, general linear and Brinley plot analyses were then carried out between difficulty conditions and sub-groups for tasks in the visuo-spatial and lexical domains to investigate the impact of increasing difficulty on performance.

General Linear Analyses

Language tasks

Picture naming

In additional comparisons of performance of sub-groups (fast and slow participants within each group) on picture naming tasks, MML analyses again found a main effect of item position ($F(1,249)=88.8$, $p<.001$) and sub-group ($F(5,205)=24.0$, $p<.001$) on RTs but **no significant interaction** between the two ($F(5,249)=0.84$, $p=.52$). Post-hoc MML analyses between individual sub-group pairings found significant differences between all sub-group pairs (p-values=.028 to <.001) except for between fast YC and fast AwPKU ($p=.12$), between slow YC and slow AwPKU ($p=.09$), or between slow YC and fast OC ($p=.72$) (Fig.4.A1). **No interactions** between individual sub-group pairs and item position were found.

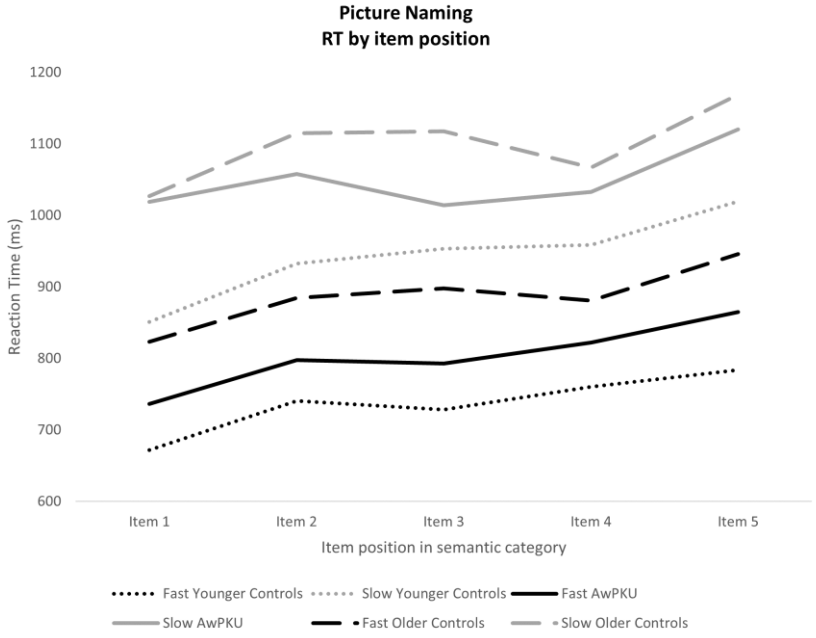


Figure 4.A1 – Error rates by item position in semantic category for fast and slow younger controls, older controls, and AwPKU in picture naming task.

Word Reading

Regularity. ANOVA analyses comparing RTs between fast and slow participants within each group found significant effects of regularity ($F(1,105)=23.1$, $p<.001$, $\eta p^2=.18$) and sub-group ($F(5,105)=21.2$, $p<.001$, $\eta p^2=.50$), **but no significant interaction between the two variables** ($F(5,105)=1.1$, $p=.34$, $\eta p^2=.05$), with all sub-groups demonstrating slower RTs for irregular, compared to regular, words (Fig.4.A2). Post-hoc ANOVA analyses of all sub-group pairings found that fast YC were significantly faster than slow YC ($p=.002$), slow AwPKU ($p<.001$), and slow OC ($p<.001$). Meanwhile slow OC were also significantly slower than slow YC ($p=.005$), fast AwPKU ($p<.001$) and fast OC ($p<.001$). Further significant differences were also found between slow AwPKU and: slow YC ($p=.03$), fast AwPKU ($p<.001$) and fast OC ($p<.001$). All other sub-group comparisons were non-significant, with **no interactions** found between any sub-group pairing and regularity effects.

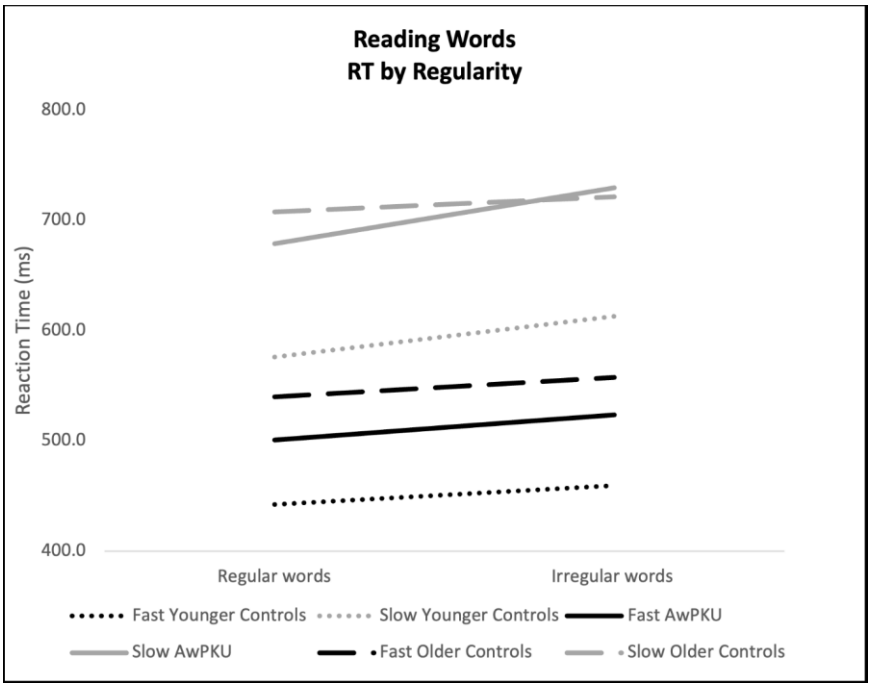


Figure 4.A2 – Error rates by word regularity for fast and slow younger controls, older controls, and AwPKU in word reading task.

Frequency. When RTs were compared between fast and slow participants in each group, significant main effects of frequency ($F(1,104)=12$, $p=.001$, $\eta p^2=.10$) and sub-group ($F(5,104)=29.8$, $p<.001$, $\eta p^2=.59$) were found, but **no interaction between the two** ($F(5,104)=1$, $p=.40$, $\eta p^2=.05$) (Fig.4.A3). The main effect of sub-group on RT was further investigated through post-hoc ANOVA analyses between individual sub-group pairings. These comparisons found that fast YC were significantly faster than slow YC ($p<.001$), slow AwPKU ($p<.001$), fast OC ($p=.004$), and slow OC ($p<.001$). Meanwhile slow OC were significantly slower than all sub-

groups (p<.001 for all) except for slow AwPKU (p=.097). Finally slow AwPKU were significantly slower than fast YC (p<.001), fast AwPKU (p<.001) and fast OC (=.001). All other sub-group pairings were non-significant. **No interactions between frequency effects and sub-group pairings** were found.

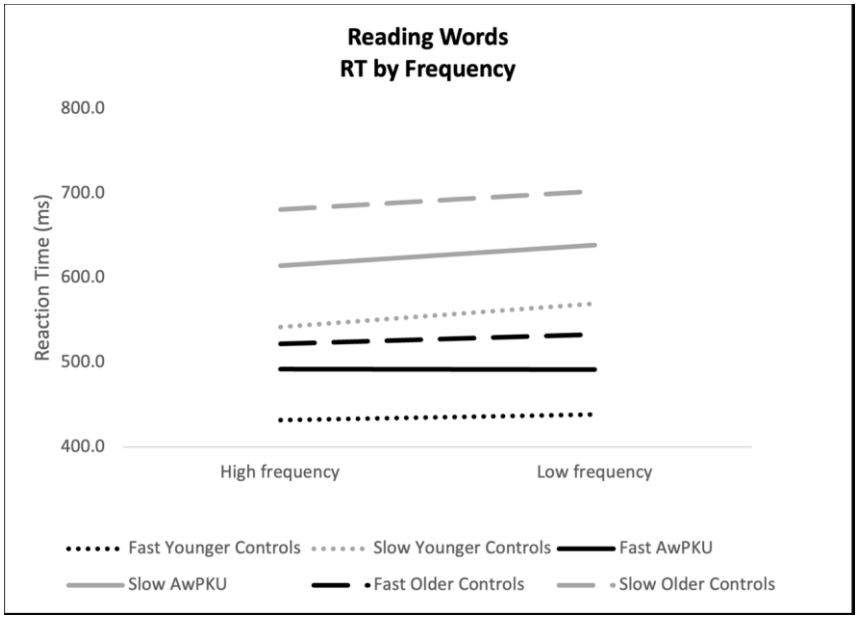


Figure 4.A3 – Error rates by word frequency for fast and slow younger controls, older controls, and AwPKU in word reading task.

Word Length. When populations were split into fast and slow groups, analyses found a significant main effect of sub-group (F(5,105)=31, p<.001) and a significant effect of word length (F(3,315)=9.4, p<.001) but **no interaction between these variables** (F(15,315)=0.96, p=.50). The main effect of sub-group on RT was further investigated through post-hoc ANOVA analyses between individual sub-group pairings. These comparisons found that fast YC were significantly faster than all other sub-groups (range p=.004 to p<.001) except for fast AwPKU (p=.50). Slow OC were significantly slower than all other sub-groups (p<.001 for all) except for slow AwPKU (p=.08). Fast AwPKU were also significantly faster than slow AwPKU (p<.001) and slow OC (p<.001), whilst fast OC were significantly faster than slow AwPKU (p=.001) (Fig.4.A4). **No significant interactions** between individual sub-group pairings and word length effects were found.

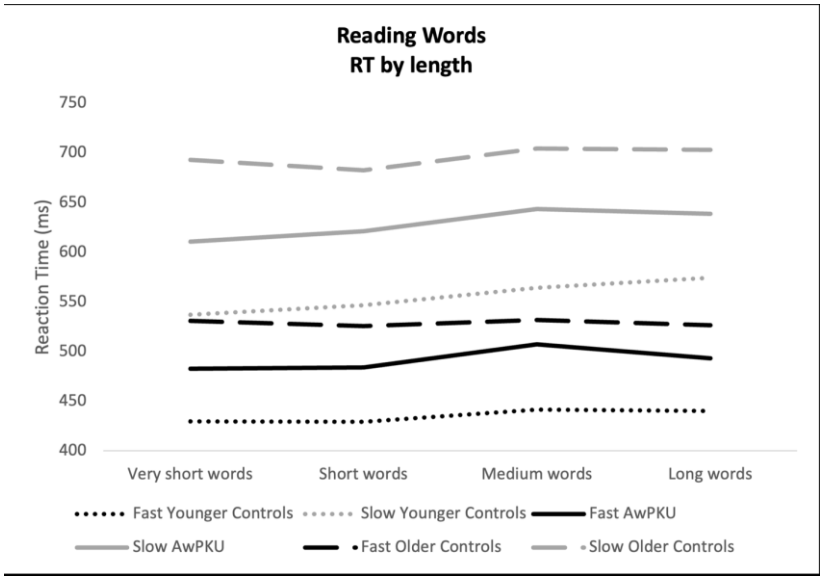


Figure 4.A4 – Error rates by word length for fast and slow younger controls, older controls, and AwPKU in word reading task.

Non-Word Length. Significant main effects of non-word length ($F(3,303)=81.9$, $p<.001$, $\eta p^2=.45$) and sub-group ($F(5,101)=28.2$, $p<.001$, $\eta p^2=.58$) on RTs were found, as well as a **significant interaction** ($F(15,303)=3.3$, $p<.001$, $\eta p^2=.14$). The main effect of sub-group on RT, and well as the significant interaction, were further investigated through post-hoc ANOVA analyses between individual sub-group pairings. These comparisons found that slow OC and slow AwPKU were significantly slower than all other sub-groups ($p<.001$ for all) except for each other ($p=1$) (Fig.4.A5). **Significant interactions** were found between non-word length effects and sub-group for comparisons of fast YC and: slow YC ($p=.02$), slow AwPKU ($p=.002$), and slow OC ($p=.003$), as well as for comparisons of fast OC and: slow YC ($p=.02$), slow AwPKU ($p<.001$), and slow OC ($p<.001$). **Significant interactions** were also found between non-word length effects in fast AwPKU and: slow AwPKU ($p=.02$), and slow OC ($p=.03$). These interactions suggest that, in general, slower sub-groups were more significantly impacted by increasing non-word length than faster sub-groups.

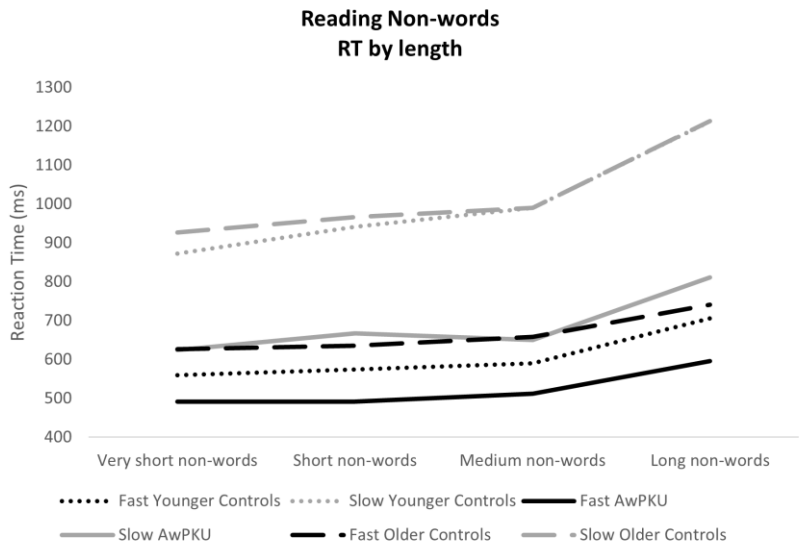


Figure 4.A5 – Error rates by non-word length for fast and slow younger controls, older controls, and AwPKU in word reading task.

Visuo-spatial tasks

Feature search

Analyses of feature search RTs between fast and slow sub-groups for displays where the **target was present** found significant main effects of sub-group ($F(5,154)=12.5$, $p<.001$) and number of distractors ($F(1,204)=4.9$, $p=.03$), but **no significant interaction** ($F(5,204)=0.5$, $p=.75$). This suggests that RTs increased in line with number of distractors, but that this increasing difficulty affected all sub-groups equally (Fig.4.A6). The main effect of sub-group on RT was further investigated through post-hoc MML analyses between individual sub-group pairings. These comparisons found significant differences in RT between all sub-groups ($p<.001$ for all) except for between fast YC and fast AwPKU ($p=.09$) or between slow YC and fast OC ($p=.10$). **No significant interactions between sub-group and number of distractors** were found for any individual sub-group pairing in target present displays.

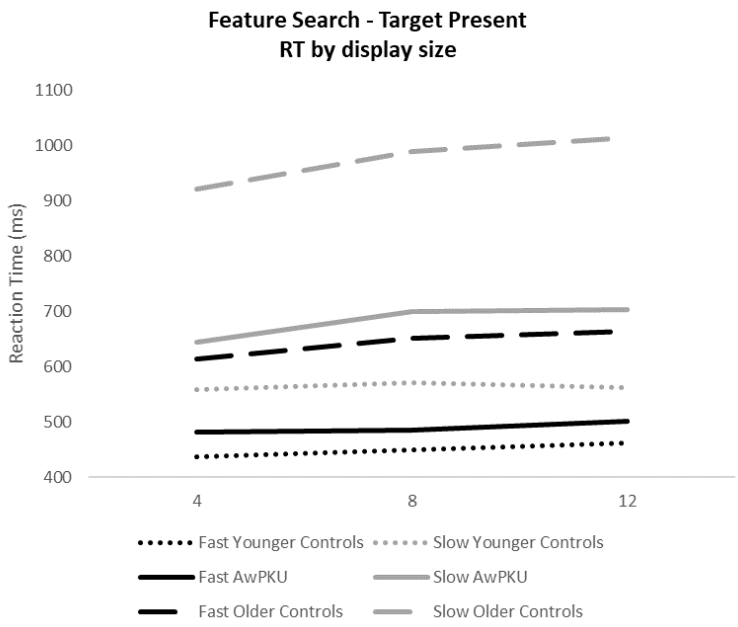


Figure 4.A6 – Error rates by number of distractors when target was present for fast and slow younger controls, older controls, and AwPKU (C) in feature search tasks.

RT analyses for sub-groups in displays where the **target was absent** found a significant main effect of sub-group ($F(5,158)=11.2$, $p<.001$) but no significant effect of number of distractors ($F(1,208)=0.9$, $p=.35$) **nor any interaction between the two** ($F(5,208)=0.5$, $p=.81$) with all sub-groups showing no significant effect of increasing numbers of distractors (Fig.4.A7). The main effect of sub-group on RT was further investigated through post-hoc MML analyses between individual sub-group pairings. These comparisons found significant differences in RT between all sub-groups ($p<.001$ for all) except for between slow YC and fast OC ($p=.10$). **No significant interactions between sub-group and number of distractors** were found for any individual sub-group pairing in target absent displays.

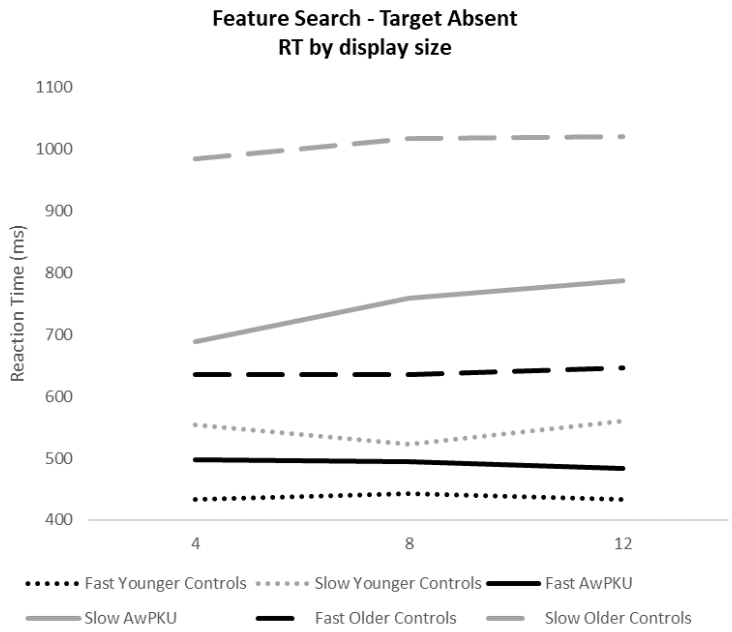


Figure 4.A7 – Error rates by number of distractors when target was absent for fast and slow younger controls, older controls, and AwPKU in feature search tasks.

Conjunction search

Analyses of conjunction search RTs between fast and slow sub-groups for **target-present** displays found significant main effects of sub-group ($F(5,158)=4.2$, $p=.001$) and number of distractors ($F(1,213)=50.8$, $p<.001$), but **no significant interaction** ($F(5,213)=1$, $p=.44$) with RTs increasing in line with number of distractors, but this increasing difficulty affecting all sub-groups relatively equally (Fig.4.A8). The main effect of sub-group on RT was further investigated through post-hoc MML analyses between individual sub-group pairings. These comparisons found significant differences in RT between fast AwPKU and: slow AwPKU ($p=.01$), slow YC ($p=.01$), fast OC ($p=.004$) and slow OC ($p=.004$), as well as between fast YC and both slow YC ($p=.04$) and slow OC ($p=.02$), and between fast and slow OC ($p=.049$). All other sub-group comparisons were non-significant. Only one interaction between sub-group and number of distractors was present, between slow AwPKU and fast OC ($F(1,74)=535$, $p=.02$), with fast OC significantly more negatively affected than slow AwPKU by the increase from 4 to 8 distractors in particular (Fig.4.A8).

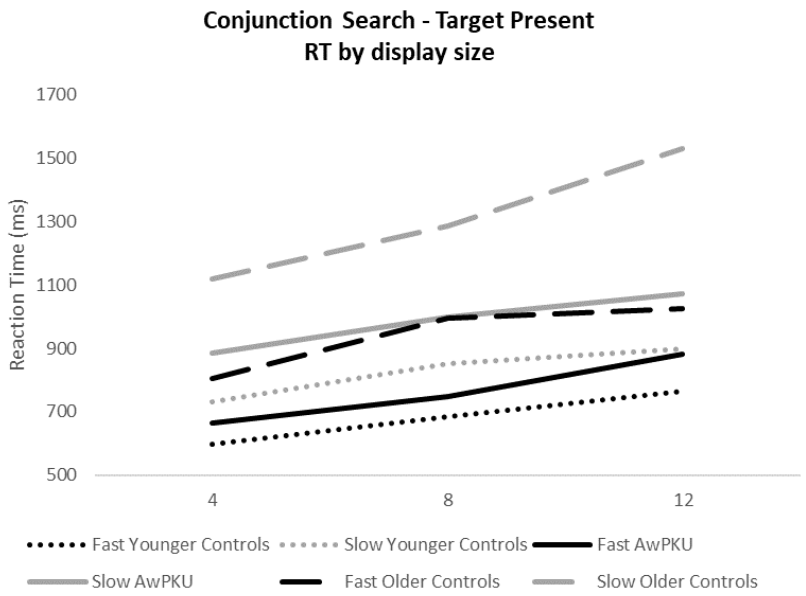


Figure 4.A8 – Error rates by number of distractors when target was present for fast and slow younger controls, older controls, and AwPKU in conjunction search task.

RT analyses between fast and slow sub-groups in displays where the **target was absent** found a significant main effect of sub-group ($F(5,143)=2.6$, $p=.03$) and number of distractors ($F(1,175)=54.7$, $p<.001$) as well as **an interaction between the two** ($F(5,175)=3.4$, $p=.006$) with increasing numbers of distractors producing negative effects of differing severities across sub-groups (Fig.4.A9). The main effect of sub-group on RT, as well as the interaction effect, were further investigated through post-hoc MML analyses between individual sub-group pairings. These comparisons found no significant main effect of sub-group for most pairings, except for slow OC who were significantly slower than both fast AwPKU ($F(1,51)=4.9$, $p=.03$) and fast YC ($F(1,56)=4.6$, $p=.04$). **Significant interactions between sub-group and number of distractors** were found between fast AwPKU and: fast YC ($p=.045$), slow YC ($p=.007$), and fast OC ($p<.001$), as well as a trend interaction with slow OC ($p=.059$). Significant interactions were also found between slow AwPKU and both slow YC ($p=.03$) and fast OC ($p=.002$) and between fast YC and fast OC ($p=.04$), demonstrating that these sub-groups were all differentially impacted by increasing task difficulty in terms of increasing number of distractors, with YC least impacted, then AwPKU, then OC, and slower sub-groups generally more impacted than faster sub-groups (Fig.4.A9).

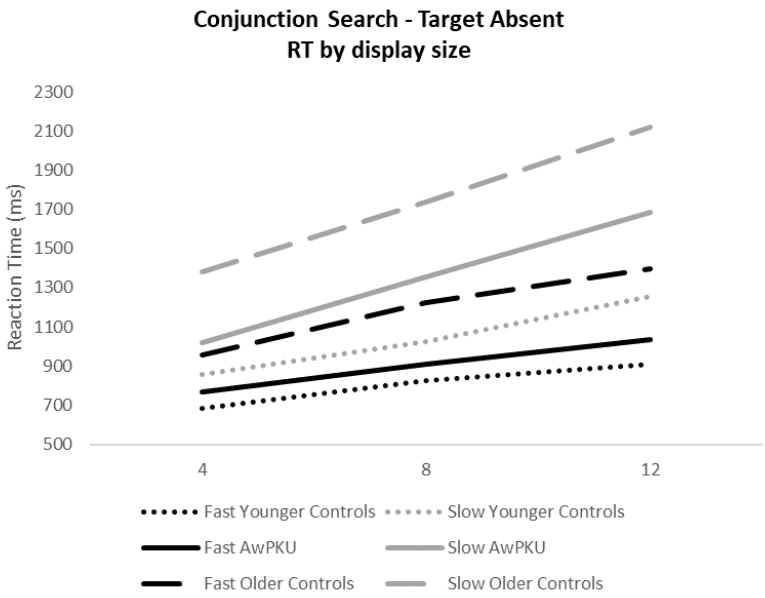
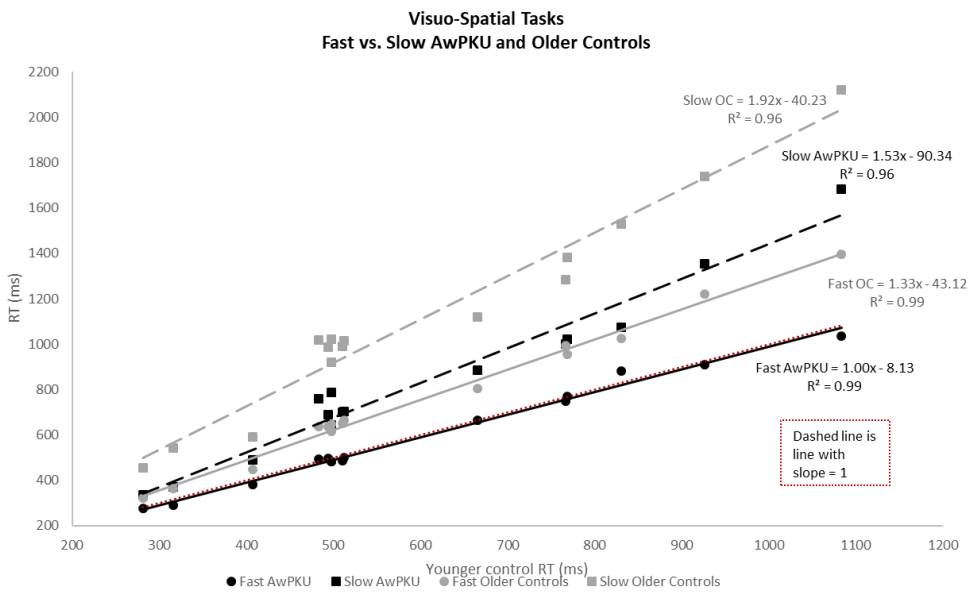


Figure 4.A9 – Error rates in by number of distractors when target was absent for fast and slow younger controls, older controls, and AwPKU in conjunction search task.

Brinley Plot Analyses

Slopes for visuo-spatial RTs of fast and slow AwPKUs and fast and slow OCs were plotted against RTs for YC (Fig.4.A10). These slopes explained between 80% and 99% of the variance in RTs exhibited by these sub-groups. RTs of fast AwPKU were almost identical to that of YC (y=1; R^2=0.99), however slower AwPKU show a 53% higher increase of RT with task difficulty than YC (y=1.53; R^2=0.96). The impact of task difficulty in visuo-spatial tasks appears to affect OC more dramatically than either of our other two populations, with fast OC showing a 33% greater increase of RT with increased difficulty, and slow OC demonstrating a 92% increase (Fast OC: y=1.33; R^2=0.99, Slow OC: y=1.92; R^2=0.96). In comparisons of language RTs, neither fast nor slow participants from either population show a notably greater impact of cognitive load on RT than YC, with both fast and slow AwPKU showing a lesser impact of task difficulty than YC (Fast AwPKU: y=0.89; R^2=0.99, Slow AwPKU: y=0.96; R^2=0.81, Fast OC: y=1.01; R^2=0.90, Slow OC: y=1.05; R^2=0.80). These comparisons, then, further support a domain-specific interaction between task complexity and visuo-spatial processing, but no such interaction with language performance.

A



B

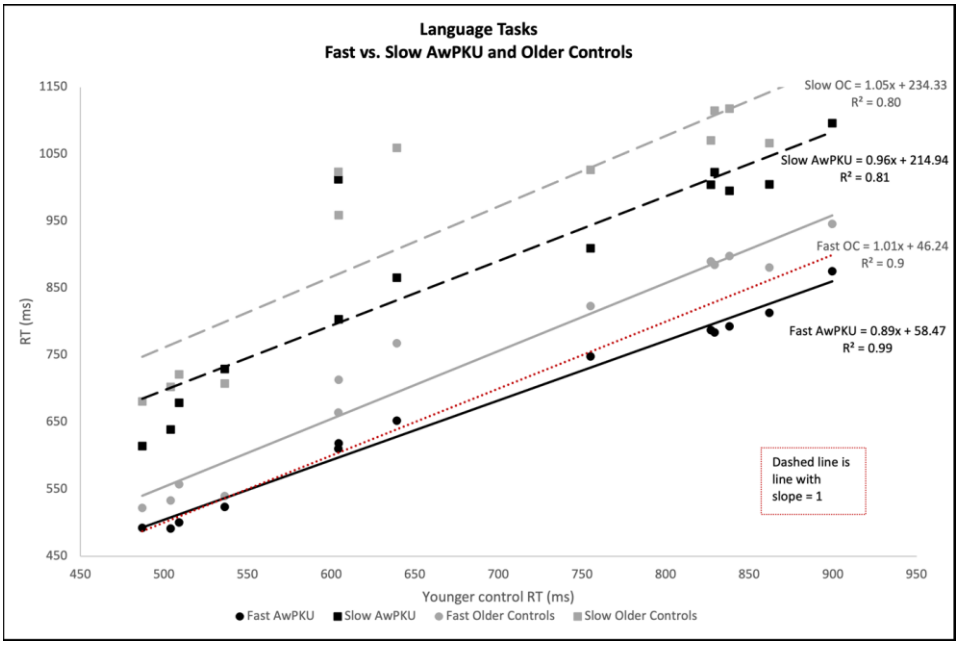


Figure 4.A10 – Brinley plots for language (A) and visuo-spatial (B) tasks. RTs of Fast and Slow AwPKU and OC plotted against the YC RTs for different conditions. Each point on the graph refers to a different condition. In both graphs, the red dotted line represents the equality line. If the RTs slope for a particular population was the same as that of controls, it would fall along the dotted line.

Appendix H. Systematic Review Tables

Table 6.A1 - Studies reporting effects of **decreasing Phe** on cognitive outcomes in A) mixed-age cohorts and B) adult cohorts. Outcomes which show a positive overall impact of lower Phe are highlighted. Weighted averages for parameters were calculated by weighting studies based on the number of participants.

Mixed-age cohorts

			Treatment		Baseline Age in yrs.				Phe in µmol/L					
									Baseline		After Manip.			
	Age Group	N ppts (Total = 194)	Early - treated	Until adolescence	Range	Mean	M/F	Mean Study Length (weeks)	Mean	SD	Mean	SD	Diff	Outcome
Diet Resumed/Restricted														
Anderson et al. (1976)a**	Child	4	No	No	7-15	12.3	3M, 1F	4	**1756**	*48*	**381**	*36*	**-1375**	Decreased choice RT
Hujibregts et al. (2002)a**	Child	21	Yes	Yes	7-14	11.3	-	1	**618**	*220*	**358**	*231*	**-260**	Improved RTs and sustained attention
Krause et al. (1985)a**	Child	5	Yes	Mixed	6-10	8.4	3M, 2F	1	**2365**	*1079*	**614**	*450*	**-1751**	Decreased choice RT
Clarke et al. (1987)	Mixed	9	Mixed (3 early infancy)	2-11y	11-18	14.4	3M, 6F	4.5	**1327**	*282*	**713**	*266*	**-614**	Significant decrease in choice RT only (no change in other tasks)*
Giffin et al. (1980)	Mixed	3	No (av. 2y)	No	9-21	15	3M	11	**1430**	-	**761**	-	**-669**	Improvement in visual attention span*
Hogan, et al. (1986)	Mixed	7	Yes	2-8y	13-18	14.5	1M, 6F	9	**1327**	*282*	**713**	*266*	**-614**	No systematic or consistent effects
Schuett et al. (1985)	Mixed	72	Mixed	No	3.5-19	5.3	36M, 36F	209	**1027**	*297*	**-**	-	**-**	Improved IQ
Sapropterin Treatment														
Burton et al. (2015)	Mixed	61	Yes	-	9-29	19.6	38M, 23F	13	**680**	-	**472**	-	**-208**	Improvement in ADHD-RS/ASRS scores in symptomatic ppts (N=19), improvement in BRIEF scores in all ppts
White et al. (2013)	Mixed	12	Yes	-	6-35	18.2	9M, 3F	26	**653**	*322*	**409**	*256*	**-244**	No change*
Arithmetic Average	**-**	**22**	**-**	**-**	**5-20**	**13**	**-**	**31**	**1243**	***578***	**553**	***166***	**-717**	**-**
Weighted Average	**-**		**-**	**-**	**5-20**	**12**	**-**	**84**	**931**	***351***	**488**	***115***	**-387**	**-**

A) Adult/adolescent cohorts.

	Age Group	N ppts (Total = 375)	Treatment		Baseline Age in yrs.		M/F	Mean Study Length (weeks)	Phe in µmol/L				Diff	Outcome
									Baseline		After Manip.			
			Early - treated	Until adolescence	Range	Mean			Mean	SD	Mean	SD		
Diet Resumed/Restricted														
Anwar et al (2013)	Adult	1	Yes	16y	41	41	1F	-	**1564**	*300*	**121**	-	**-1443**	General improvement on Adenbrook Scale
Burgess et al. (2021)	Adult	9	Yes	-	19-47	34	4M, 5F	54	**1108**	*293*	**-**	-	**-**	Increase in Mean Cognitive Performance Index
Dawson et al. (2011)	Adult	16	Yes	Yes	-	-	16F	-	**1223**	*330*	**277**	*108*	**-946**	Faster eye-movement RTs time
Schmidt et al. (1996)a**	Adult	15	Yes	-	17-24	20.5	7M, 8F	4.5	**1320**	*270*	**630**	*172*	**-690**	Improved sustained attention and calculation speed
Pegvaliase Treatment														
Bilder et al. (2021)	Adult	156	Yes	-	-	29.6	80M, 76F	164	**1263**	*29*	**584**	*47*	**-679**	Reduced score on ADHD-RS IV Inattention subscale
Thomas et al. (2018)	Adult	178	Yes	-	-	29.2	-	53	**1233****	*386*	**565****	*531*	**-668**	Reduced score on ADHD-RS IV Inattention subscale
Arithmetic Average	**-**	**63**	**-**	**-**	**21-41**	**31**	**-**	**69**	**1285**	***153***	**435**	***224***	**-885**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**21-41**	**29**	**-**	**99**	**1246**	***34***	**562**	***67***	**-688**	**-**

** = within-participant statistical comparisons conducted; **Phe levels estimated from larger group; ADHD-RS = ADHD Rating Scale; ASRS = Adult ADHD Self Report Scale; RT = Reaction Time; BRIEF = Behaviour Rating Inventory of Executive Function; Ppt = participant; RT = Reaction Time; ** Anderson et al. (1976)a = Low-Phe Phase, b = High-Phe phase; Huijbregts et al. (2002)a = 'Phe-Down' cohort; b = 'Phe-Up' cohort, Krause et al. (1985)a = Phe restriction cohort, b = Phe loading cohort; Schmidt et al. (1996)a = Phase 2 (diet restriction phase), b = Phase 1 (Diet relaxation phase)*

Table 6.A2 - Studies reporting effects of **increasing Phe** on cognitive outcomes in A) mixed-age cohorts and B) adult/adolescent cohorts. Outcomes which show an overall positive effect of lower Phe are highlighted. Weighted means for parameters were calculated by weighting studies based on the number of participants.

A) Mixed-age cohorts

			Treatment		Baseline Age in yrs.				Phe in µmol/L					
									Baseline		After Manip.			
	Age Group	N ppts (Total = 236)	Early - treated	Until adolescence	Range	Mean	M/F	Mean Study Length (weeks)	Mean	SD	Mean	SD	Diff	Outcome
Diet Discontinued														
Brown & Warner (1976)	Child	11	Late-treated (<2y)	Yes	6-8.8	7	4M, 7F	191	**433**	*200*	**1858**	*533*	**1425**	Decrease in IQ
Cabalska et al. (1977)a**	Child	22	Yes	Yes	-	4	-	209	**478**	*127*	**1283**	*276*	**805**	Decrease in IQ from 2y onwards*
Cabalska et al. (1977)b**	Child	10	Yes	Yes	-	2	-	209	**509**	*157*	**1362**	*218*	**853**	Steady decrease in IQ
Cabalska et al.(1977)c**	Child	5	Yes	Yes	-	3	-	209	**400**	*73*	**835**	*412*	**435**	No change
Holtzman et al. (1975)	Child	5	Yes (<1m)	Yes	-	4	3M, 2F	104	**993**	*545*	**1628**	*248*	**635**	Decreased in IQ
Huijbregts et al. (2002)b**	Child	19	Yes	Yes	7-14	11.3	-	1	**284**	*169*	**470**	*201*	**186**	No change
Koch et al. (1982)	Child	60	Yes (<4m)	Yes	-	6	-	104	**775**	*248*	**1556**	*333*	**781**	Decrease in scores on reading and spelling subtests of WRAT*
Leuzzi et al. (1997)	Child	2	Yes	Yes	11-12	11.5	2F	1	**476**	-	**2087**	*183*	**1611**	Decreased visual retention

														performance in 1 ppt
Pueschel et al. (1983)	Child	8	Yes	Yes	-	5	-	52	**666**	*109*	**1544**	*133*	**878**	No change
Smith et al. (1978)a**	Child	47	Mixed (21 ppts <4m, 26 ppts <5y)	Yes	7-10	8.5	-	209	**420**	-	**>1200**	-	**-**	Drop in IQ for some (not all) ppts*
Smith et al. (1978)b**	Child	22	Mixed (5 ppts <4m, 17 ppts <5y)	Yes	7-10	8.5	-	209	**420**	-	**1200**	-	**780**	No change*
Phe Loading														
Anderson et al. (1976)b**	Child	4	No	-	7-15	12.3	3M, 1F	4	**381**	*36*	**1774**	*604*	**1393**	Increased choice RT
Griffiths et al. (1998)	Mixed	16	Yes (<3w)	Yes	10-16	12.6	10M, 6F	13	**619**	*254*	**1177**	*255*	**558**	No significant changes on all but Rey verbal learning (Rey attributed to practice effects)*
Krause et al. (1985)b**	Mixed	5	Yes	Mixed	6-24	17.2	4M, 1F	1	**463**	*227*	**2192**	*672*	**1729**	Increased choice RT
Arithmetic Average	**-**	**17**	**-**	**-**	**2-17**	**8**	**-**	**108**	**523**	***185***	**1459**	***482***	**928**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**2-17**	**8**	**-**	**134**	**543**	***176***	**1358**	***390***	**784**	**-**

B) Adult/adolescent cohorts

	Age Group	N ppts (Total = 87)	Treatment		Baseline Age in yrs.		M/F	Mean Study Length (weeks)	Phe in µmol/L				Diff	Outcome
									Baseline		After Manip.			
			Early - treated	Until adolescence	Range	Mean			Mean	SD	Mean	SD		
Diet Discontinued														
Lou et al. (1985)	Adult	4	Mixed (1w-4y)	-	16-23	20	3M, 1F	3	**863**	*229*	**1590**	*509*	**727**	RT increase; more variability
Lou, et al. (1987)	Adult	9	Mixed (5ppts treated > 3m)	-	15-24	18.4	6M, 3F	3	**758**	*210*	**1410**	*460*	**652**	Increase in number of long RTs*
Lou et al. (1994)	Adult	4	Yes	Yes	15	15	-	104	**1033**	*34*	**1348**	*114*	**315**	No change
Schmidt et al. (1996)b**	Adult	15	Yes	-	17-24	20.5	7M, 8F	4.5	**630**	*172*	**1410**	*290*	**780**	Worse sustained attention, concentration, and calculation speed
Phe Loading														
Pietz, et al. (1993)	Adult	5	Yes	-	20	20	3M, 2F	4	**753**	*194*	**1600**	*290*	**847**	RT Increase
Sunderman n et al. (2011)	Adult	17	Yes	Yes	22-38	31	17M	0.01	**1180**	*265*	**2170**	*320*	**990**	No change
ten Hoedt et al. (2011)	Adult	9	Yes	Yes	19-34	23.5	3M, 6F	4	**649**	*222*	**1220**	*311*	**571**	More fluctuation in sustained attention; Slower RTs in a number of tasks
Pegvaliase Discontinued														
Harding et al. (2018)	Adult	24	Yes	-	19-51	31.2	16M, 12F	8	**536**	*433*	**1337**	-	**800**	No change in ADHD-RS IV Inattention subscale score*
Arithmetic Average	**-**	**11**	**-**	**-**	**15-31**	**22**	**-**	**16**	**800**	***216***	**1511**	***295***	**710**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**15-31**	**25**	**-**	**9**	**763**	***239***	**1535**	***328***	**772**	**-**

** = within-participant statistical comparisons conducted; ADHD-RS = ADHD Rating Scale; Ppt = participant; RT = Reaction Time; WRAT = Wide Range Achievement Test; ** Anderson et al. (1976)a = Low-Phe Phase, b = High-Phe phase; Cabalska et al. (1997)a = 'Group 1' (classic PKU, discontinued diet aged 4), b = 'Group 2' (classic PKU, discontinued diet aged 2), c = 'Group 3' (mild PKU/hyperphenylalaninemia); Huijbregts et al. (2002)a = 'Phe-Down' cohort; b = 'Phe-Up' cohort, Krause et al. (1985)a = Phe restriction cohort, b = Phe loading cohort; Schmidt et al. (1996)a = Phase 2 (diet restriction), b = Phase 1 (diet relaxation) ; Smith et al. (1978)a = 'Groups 1 & 2' (London cohort), b = 'Groups 3A & 4' (Heidelberg cohort)*

Table 6.A3 – Studies reporting quantitative cognitive outcomes included in meta-analysis

Study authors	N Ppts	Domain	Task	Score/RT Low Phe condition Mean(SD)	Score/RT High Phe condition Mean (SD)	Hedge's G for cognitive outcomes	Baseline Phe (µmol/L)	Phe difference (µmol/L)	Hedge's G for Phe levels	Length of study (weeks)	Difference (Yes/No)
Brown & Warner (1976)	11	IQ	Stanford-Binet IQ & WISC	75.4 (1.4)	66 (7.75)	1.61	433	1425	3.5	191	Y
Cabalska et al. (1977)a*	22	IQ	Brunet-Lezine, Terman-Merill and WISC	101.5 (10.5)	90.2 (9.7)	1.09	478	805	3.7	209	Y
Cabalska et al. (1977)b*	10	IQ	Brunet-Lezine, Terman-Merill and WISC	91.4 (16.2)	77.2 (6.5)	1.09	509	853	4.5	209	Y
Cabalska et al. (1977)c*	5	IQ	Brunet-Lezine, Terman-Merill and WISC	98.2 (10.7)	98.7 (4.9)	-0.05	400	435	1.5	209	N
Giffin et al. (1980)	2**	Visual attention	Picture fixation accuracy	50 (12)	32 (7.4)	0.80	1430	-669	-	11	Y
Griffiths et al. (1998)	16	Verbal memory	Rey Verbal Learning	47.5 (5.6)	52.9 (6)	-0.90	619	558	2.2	13	N
Griffiths et al. (1998)	16	Verbal memory	Paired-Associate Learning	26.4 (2.2)	26.4 (2.7)	0.00	619	558	2.2	13	N
Griffiths et al. (1998)	16	Verbal memory	Digit span	9.9 (2.1)	9.7 (2.3)	0.09	619	558	2.2	13	N
Griffiths et al. (1998)	16	Motor coordination	Rey-Davis Manual Labyrinth	15.9 (4.2)	15.2 (3.9)	0.17	619	558	2.2	13	N
Griffiths et al. (1998)	16	Motor coordination	Purdue Pegboard (Time to complete)	158.1 (9.7)	155.3 (14.2)	-0.22	619	558	2.2	13	N
Griffiths et al. (1998)	16	Motor coordination	Hole-type Steadiness Tester	19.3 (8.8)	18.8 (11.3)	-0.05	619	558	2.2	13	N

Griffiths et al. (1998)	16	Visuo-spatial memory	Matching Familiar Figures	6.1 (1.3)	6.2 (1.4)	-0.07	619	558	2.2	13	N
Griffiths et al. (1998)	16	Visuo-spatial memory	Corsi Block Tapping	7.9 (1.6)	8.1(2.1)	-0.10	619	558	2.2	13	N
Lou et al. (1985)	4	Speed of processing	Simple RT task (RT)	253.75 (22.6)	267.8 (23.9)	0.50	863	727	1.8	3	Y
Lou et al. (1987)	9	Speed of processing	Continuous detection task (RT)	304ms (29)	329ms (51)	0.57	758	652	1.8	3	Y
Pietz et al. (1993)	5	Sustained Attention	Dot Pattern Exercise	8.8s (0.9)	11.7s (2.3)	1.47	753	847	3.4	4	Y
Pueschel et al. (1983)	8	IQ	Stanford-Binet IQ	100 (16)	106 (8)	-0.44	666	878	7.2	52	N
Sundermann et al. (2011)	15	Inhibition	Stroop Incongruent (RT)	1016.9 (171.3)	1002.3 (164.7)ms	-0.08	1180	990	3.4	0.01	N
Sundermann et al. (2011)	15	Language	Stroop Neutral (RT)	871 (146)ms	854.3 (139.8)ms	-0.11	1180	990	3.4	0.01	N
Thomas et al. (2018)	178	Attention	ADHD RS - IA	9.8 (6.12)	5 (4.9)	0.86	565	668	1.4	53	Y

*Ppts = Participants, RT = Reaction time, WISC = Weschler Intelligence Scale for Children, WAIS(R) = Weschler Adult Intelligence Scale (-Revised), ADHD RS IA = ADHD Rating scale, Inattention subscale; *Cabalska et al. (1997)a = 'Group 1' (classic PKU, discontinued diet aged 4), b = 'Group 2' (classic PKU, discontinued diet aged 2), c = 'Group 3' (mild PKU/hyperphenylalaninemia); **Study included 3 participants, however quantitative data was only available for 2 participants*

Table 6.A4 - Studies reporting effects of **decreasing Phe** on well-being outcomes in mixed-age cohorts and adult cohorts. Outcomes which show an overall positive effect of lower Phe are highlighted. Weighted means for parameters were calculated by weighting studies based on the number of participants.

A) Mixed-age Cohorts:

	Age Group	N ppts (Total = 108)	Treatment: Early - treated	Treatment: Until adolescence	Baseline Age in yrs.: Range	Baseline Age in yrs.: Mean	M/F	Mean Study Length (weeks)	Phe in µmol/L: Baseline: Mean	Phe in µmol/L: Baseline: SD	Phe in µmol/L: After Manip.: Mean	Phe in µmol/L: After Manip.: SD	Diff	Outcome
Diet Resumed/Restricted														
Bickel et al. (1954)a***	Child	1	No	Untreated	-	2	1F	35	**4056**	-	**726**	-	**-3330**	Gained weight, brighter eyes, more interest in surroundings. Learned to crawl/stand.
Schuett et al. (1985)	Child	72	Mixed	No	6-6.5	-	36M, 36F	209	**1027**	*297*	**-**	-	**-**	42/72 ppts Improved school performance, mood, peer acceptance
Gassio et al (2003)	Mixed	15	Mixed (14 early, 3 late - 23-31y)	No	11-36.5	23	10F, 5M	52	**-**	-	**717**	*294*	-	Majority of ppts reported increased happiness, calmness, alertness, less easily upset. (BUT 7% ppts reported less vitality, more tiredness and more easily upset)
Giffin et al. (1980)	Mixed	3	No (av. 2y)	No	9-21	15	3M	11	**1430**	-	**761**	-	**-669**	1ppt - Increased affection. 2/3 ppts no change/no data
Sapropterin Treatment														
Douglas et al. (2013)	Mixed	17	Yes	-	10-49	22.1	-	52	**690**	-	**502**	-	**-188**	Improved Impact and Satisfaction PKU-QoL sub-scores
Arithmetic Average	**-**	**22**	**-**	**-**	**2-23**	**16**	**-**	**72**	**1801**	***1534***	**677**	***108***	**-1396**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**2-23**	**21**	**-**	**155**	**1011**	***355***	**619**	***113***	**-406**	**-**

B) Adult/adolescent cohorts

	Age Group	N ppts (Total = 336)	Treatment		Baseline Age in yrs.		M/F	Mean Study Length (weeks)	Phe in µmol/L				Diff	Outcome
									Baseline		After Manip.			
			Early - treated	Until adolescence	Range	Mean			Mean	SD	Mean	SD		
Diet Resumed/Restricted														
Anwar et al. (2013)	Adult	1	Yes	Until 16y	-	41	1F	-	**1564**	*300*	**121**	-	**-1443**	Improved affect and behaviour
Burgess et al. (2021)	Adult	9	Yes	-	19-47	34	4M, 5F	54	**1108**	*293*	**-**	-	-	Decreased anxiety and depression
Bik-Multanowski et al. (2008)	Adult	29	Yes	-	18-32	24	-	39	**1105**	*210*	**715**	*210*	**-390**	Decrease in distress and increased in well-being*
Dion et al. (2001)	Adult	1	No (av. 7y)	No	-	41	1F	52	**1705**	-	**503**	*135*	**-1202**	Decreased seclusion and aggression; improved mood
Fitzgerald et al. (2000)	Adult	5	No	-	30-56	41	2M, 3F	37	**1676**	*128*	**478**	*84*	**-1198**	Reduced problem behaviours and hyperactivity, increased sociability, alertness, and concentration
Harper & Reid (1987)	Adult	1	No	-	-	54	1F	30	**1450**	-	**940**	-	**-510**	Reduced aggression, increased calm
Hoskin et al. (1992)	Adult	1	No (av. 2y)	Until 4y	-	35	1F	18	**1539**	-	**523**	-	**-1016**	Reduced behavioural problems*
Lee et al. (2009)	Adult	17	No	-	21-69	-	-	48	**1444**	*255*	**553**	*158*	**-891**	More positive observations (from blinded observers) BUT no change in ABC or VABS scores*
Marholin et al. (1978)	Adult	5	No	-	19-53	33	3M, 2F	8	**2143**	*442*	**424**	*182*	**-1719**	No reliable impact on behavioural measures
Williams (1998)	Adult	1	No	-	-	73	1M	26	**1108**	-	**811**	-	**-297**	Increased social interest and smiling, dissipation of self-injury
Yannicelli & Ryan (1995)	Adult	88	No	-	50% ppts >35yrs	-	-	-	**1659**	*333*	**617**	*438*	**-1042**	Decreased irritability, hyperactivity, and aggression
Pegvaliase Treatment														
Thomas et al. (2018)	Adult	178	Yes	-	-	29	-	53	**1233**	*386*	**565**	*531*	**-668**	Reduced Profile of Mood States score
Arithmetic Average	**-**	**28**	**-**	**-**	**24-73**	**41**	**-**	**37**	**1478**	***309***	**568**	***214***	**-943**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**24-73**	**30**	**-**	**49**	**1365**	***226***	**588**	***62***	**-783**	**-**

** = within-participant statistical comparisons conducted; **Phe levels estimated from larger group; ABC = Aberrant Behaviour Checklist; Ppts = participant; VABS = Vine Adaptive Behaviour Scale; PKU-QoL = PKU Quality of Life Questionnaire; ***Bickel et al. (1954)a = Phase 1 (diet initiation), b = Phase 2 (Phe loading)*

Table 6.A5 - Studies reporting effects of **increasing Phe** on well-being outcomes in mixed-age and adult/adolescent cohorts. Outcomes which show a positive overall effect of lower Phe are highlighted. Weighted means for parameters were calculated by weighting studies based on the number of participants.

			Treatment		Baseline Age in yrs.				Phe in µmol/L					
									Baseline		After Manip.			
	Age Group	N ppts (Total = 43)	Early - treated	Until adolescence	Range	Mean	M/F	Mean Study Length (weeks)	Mean	SD	Mean	SD	Diff	Outcome
Diet Discontinued														
Bickel et al. (1954)b**	Child	1	No	Untreated	-	2	1F	0.9	**726**	-	**5448**	-	**4722**	Increased irritability and drowsiness, loss of interest in surroundings and food, loss of ability to stand/crawl
Leuzzi et al. (1997)	Child	2	Yes	Yes	11-12	11.5	2F	1	**476**	-	**2087**	*183*	**1611**	Increased restlessness, lability of attention and irritability
Solomons et al. (1966)	Child	7	Mixed	On diet for 1.5-7y	6-9	7.3	-	37	**381**	*351*	**793**	*375*	**412**	5/7 ppts increased sociability, decreased hyperactivity, tantrums, and self-abuse. 2/7 unchanged.
Phe Loading														
Ten-Hoedt et al. (2011)	Adult	9	Yes	Yes	19-34	23.5	3M, 6F	4	**649**	*222*	**1220**	*311*	**571**	Significant increase in depression and fatigue, plus decrease in vigour
Pegvaliase Discontinued														
Harding et al. (2018)	Adult	24	Yes	-	19-51	31.2	16M, 12F	8	**536.1**	*433*	**1337**	-	**801**	No change*
Arithmetic Average	**-**	**9**	**-**	**-**	**2-31**	**15**	**-**	**10**	**554**	***137***	**2177**	***1887***	**1623**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**2-31**	**24**	**-**	**11**	**536**	***88***	**1354**	***693***	**818**	**-**

** = within-participant statistical comparisons conducted; Ppt = participant; **Bickel et al. (1954)a = Phase 1 (diet initiation), b = Phase 2 (Phe loading)*

Table 6.A6 - Studies reporting effects of **decreasing Phe** on neurophysiological outcomes in mixed and adult/adolescent cohorts. Outcomes which show an overall positive effect of lower Phe are highlighted. Weighted means for parameters were calculated by weighting studies based on the number of participants.

			Treatment		Baseline Age in yrs.				Phe in µmol/L						
									Baseline		After manip.				
	Age group	N ppts (Total = 72)	Early - treated	Until adolescence	Range	Mean	M/F	Mean Study length (weeks)	Mean	SD	Mean	SD	Diff	Measure	Outcome
Diet Resumed/Restricted															
Jaulent et al. (2020)	Adult	25	Mixed	-	18-50	30.2	14M, 11F	52	**1613**	*491*	**969**	*304.7*	**-644**	Neurological exam & MRI	Improvement of symptoms in 24/25 ppts Improvement of MRI in 10/18 ppts
Walter et al. (1997)	Adult	1	Yes	14y	-	29	1F	126	**-**	-	**<300***	-	**-**	MRI	Reversal of WM abnormalities
Cleary et al. (1995)a**	Mixed	21	Yes	14y	14-49	-	-	33	**1300**	*75*	**1000**	*68*	**-300**	MRI	No change
Cleary et al. (1995)b**	Mixed	5	Yes	14y	14-49	-	-	33	**1400**	*175*	**400**	*183*	**-1000**	MRI	Reversal of WM abnormalities in 10/26 ppts
Sapropterin Treatment															
Clocksin et al. (2021)	Mixed	8	Yes	Yes	9-35	21.2	13M, 9F	28	**693**	*291*	**375**	*222*	**-318**	AFQ & DTI	WM improvement
White et al. (2013)	Mixed	12	Yes	-	7-35	18.2	9M, 3F	26	**653**	*322*	**409**	*256*	**-244**	DTI	Significant improvement in 3/10 ROIs Increased diffusivity in 8/10 ROIs
Arithmetic Average	**-**	**12**	**-**	**-**	**18-30**	**25**	**-**	**50**	**1132**	***434***	**576**	***319***	**-501**	**-**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**18-30**	**25**	**-**	**39**	**1240**	***382***	**770**	***286***	**-462**	**-**	**-**

**value of 300µmol/L used for means and analyses; Ppt = participant; ROI = Region of Interest; WM = White Matter; **Cleary et al. (1995)a = 'Group 1' (diet resumption – loose), b = 'Group 2' (diet resumption – strict)*

Table 6.A7 - Studies reporting effects of **increasing Phe on neurophysiological outcomes** in mixed-age and adult cohorts. Outcomes which show an overall positive effect of lower Phe are highlighted. Weighted means for parameters were calculated by weighting studies based on the number of participants.

			Treatment		Baseline Age in yrs.				Phe in µmol/L						
									Baseline		After manip.				
	Age group	N ppts (Total = 83)	Early - treated	Until adolescence	Range	Mean	M/F	Mean Study Length (weeks)	Mean	SD	Mean	SD	Diff	Measure	Outcome
Diet Discontinued															
Cabalska et al. (1977)a*	Child	22	Yes	Yes	-	4	-	209	**478**	*127*	**1283**	*276*	**805**	EEG	Increased abnormal EEGs
Cabalska et al. (1977)b*	Child	10	Yes	Yes	-	2	-	209	**509**	*157*	**1362**	*218*	**853**	EEG	Increased abnormal EEGs
Cabalska et al. (1977)c*	Child	5	Yes	Yes	-	3	-	209	**400**	*73*	**835**	*412*	**435**	EEG	Increased abnormal EEGs
Leuzzi et al. (1997)	Child	2	Yes	Yes	11-12	11.5	2F	1	**476**	-	**2087**	*183*	**1611**	MRI	No change
Pueschel et al. (1983)	Child	8	Yes	Yes	-	5	-	52	**666**	*109*	**1544**	*133*	**878**	EEG	No change
Lou et al. (1994)	Adult	4	Yes	Yes	-	15	-	104	**1033**	*34*	**1348**	*114*	**315**	MRI	No change
Phe Loading															
Sundermann et al. (2011)	Adult	15	Yes	Yes	22-38	31	15M	0.01	**1180**	*265*	**2170**	*320*	**990**	fMRI	No change
Leuzzi et al. (2014)	Mixed	17	Yes	>10y	10-20	14.3	11M, 6F	0.01	**572**	*193*	**1129**	*177.3*	**557**	EEG	Smaller CNV and delayed motor response in children only
Arithmetic Average	**-**	**10**	**-**	**-**	**2-31**	**11**	**-**	**98**	**664**	***286***	**1470**	***456***	**806**	**-**	**-**
Weighted Average	**-**	**-**	**-**	**-**	**2-31**	**12**	**-**	**103**	**668**	***273***	**1442**	***399***	**774**	**-**	**-**

*CNV = Contingent Negative Variation; * Cabalska et al. (1997)a = 'Group 1' (classic PKU, discontinued diet aged 4), b = 'Group 2' (classic PKU, discontinued diet aged 2), c = 'Group 3' (mild PKU/hyperphenylalaninemia)*