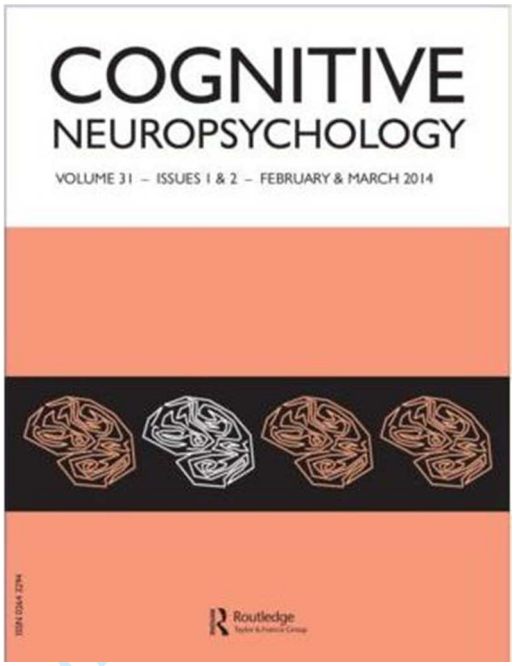
Markers of cognitive function in patients with metabolic disease: Morquio Syndrome and Tyrosinemia Type III

Blundell, James; Frisson, Steven; Chakrapani, Anupam; Kearney, Shauna; Vijay, Suresh; MacDonald, Anita; Gissen, Paul; Hendriksz, Chris; Olson, Andrew

DOI:
10.1080/02643294.2018.1443913

Document Version
Peer reviewed version

Citation for published version (Harvard):
Blundell, J, Frisson, S, Chakrapani, A, Kearney, S, Vijay, S, MacDonald, A, Gissen, P, Hendriksz, C & Olson, A 2018, 'Markers of cognitive function in patients with metabolic disease: Morquio Syndrome and Tyrosinemia Type III', *Cognitive Neuropsychology*, vol. 35, no. 3-4, pp. 120-147.
https://doi.org/10.1080/02643294.2018.1443913



Markers of cognitive function in patients with metabolic disease: Morquio Syndrome and Tyrosinemia Type III

Journal:	*Cognitive Neuropsychology*
Manuscript ID	CN-FA 38.17.R2
Manuscript Type:	Full Article
Keywords:	Tyrosinemia, inherited metabolic disease, language, attention, developmental disorder, MPS-IVa

Markers of cognitive function in individuals with metabolic disease: Morquio Syndrome and Tyrosinemia Type III

James Blundell[1], Dr Steven Frisson[1], Dr Anupam Chakrapani[2], Shauna Kearney[2], Dr Suresh Vijay[2], Professor Anita MacDonald[2], Professor Paul Gissen[3], Dr Chris Hendriksz[4], Dr Andrew Olson[1]

[1]University of Birmingham, Birmingham, UK

[2]Birmingham Children's Hospital, Birmingham, UK

[3]Great Ormond Street Hospital, London, UK

[4]University of Pretoria, Steve Biko Academic Unit, Pretoria, South Africa

Acknowledgements. The authors would like to thank all of the individuals with metabolic disease and their families who supported this research, the many school children who participated in the control sample and the teachers and school administrators who made this possible. We gratefully acknowledge the support of an ESRC CASE studentship award to James Blundell in collaboration with the MPS Society and by the Birmingham Children's Hospital Charities award BCHRF 230.

Correspondence concerning this article should be addressed to Dr James Blundell, School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

Abstract

We characterised cognitive function in two metabolic diseases. MPS-IVa (Morquio) and Tyrosinemia Type III individuals were assessed using tasks of attention, language and oculomotor function. MPS-IVa individuals were slower in visual search, but the display size effects were normal and slowing was not due to long reaction times (ruling out slow item processing or distraction). Maintaining gaze in an oculomotor task was difficult. Results implicated sustained attention and task initiation or response processing. Shifting attention, accumulating evidence and selecting targets were unaffected. Visual search was also slowed in Tyrosinemia Type III and patterns in visual search and fixation tasks pointed to sustained attention impairments, although there were differences from MPS-IVa. Language was impaired in Tyrosinemia Type III but not MPS-IVa. Metabolic diseases produced selective cognitive effects. Our results, incorporating new methods for developmental data and model selection, illustrate how cognitive data can contribute to understanding function in biochemical brain systems.

Keywords: Morquio, MPS-IVa, Tyrosinemia, inherited metabolic disease, language, attention, developmental disorder

Markers of cognitive function in individuals with mild metabolic disease: Morquio Syndrome and Tyrosinemia Type III

Inherited metabolic diseases (IMDs) are large and heterogeneous class of genetic disorders that are caused by dysfunction within a single pathway of intermediary metabolism. In these diseases the dysfunction of metabolic enzymes leads to the accumulation of metabolites, which are often toxic, disrupting the normal development of multiple systems. The severity of symptoms associated with IMDs can vary widely. Mild symptoms can include physiological abnormalities such as skeletal dysplasia and impaired endurance (Davison, Kearney, & Horton, 2013; Wraith, 2006). Severe consequences include mental retardation, central nervous system (CNS) complications, and reduced life expectancy (Bendadi et al., 2014; De Laet et al., 2011; Masurel-Paulet et al., 2008; Thimm et al., 2011, 2012). Research into the cognitive impact of IMDs has largely been limited to standardised intelligence tests, achievement tests and adaptive behaviour scales (Bax & Colville, 1995; Biernacka, Jakubowska-Winecka, & Tylki-Szymanska, 2010; Davison et al., 2013; Shapiro et al., 2009). These do not generally allow impairments in specific cognitive domains to be characterised or tracked over time (Martin et al., 2008). In the current study we present results from two IMD groups, Morquio syndrome and Tyrosinemia type III, where cognitive impairments have been considered mild (using standardised tests). We compare individuals with Morquio syndrome and Tyrosinemia type III to typically developing controls (Thomas, Annaz, & Ansari, 2009) to evaluate affected and preserved cognitive abilities in the domains of language, attention and oculomotor control.

Morquio Syndrome (MPS-IVa , OMIM 253000)

Morquio syndrome (MPS-IVa) is lysosomal storage disorder that is caused by the deficiency of the lysosomal enzyme N-acetylgalactosamine-6-sulfase (GALNS, EC 3.1.6.4; encoded by GALNS gene at 16q24.3) which has a role in degradation of the glycosaminoglycans (GAG) keratan sulphate and chondroitin-6-sulfate (Neufeld & Muenzer, 2001; Wraith, 2006). Both keratan sulphate and chondroitin-6-sulfate are essential constituents of connective tissue, including cartilage and vessel walls. The accumulation of these two GAGs leads to a classic phenotype defined by severe skeletal dysplasia, hip dysplasia, marked short stature, genu valgum, and cornea clouding (Hendriksz et al.,

2013; Wraith, 2006). Treatments to reduce substrate burden in MPS-IVa include enzyme replacement therapy (ERT) and haematopoietic stem cell transplantation (HSCT), which alleviate the majority of the skeletal and coronary complications.

In contrast to other lysosomal storage disorders (e.g. Hurler-Scheie syndrome, Hunter's syndrome, Niemann-Pick Type C), Individuals with MPS-IVa have not typically been reported to have neurological or neurocognitive impairments (Dvorak-Ewell et al., 2010; Wraith, 2006) and neuroimaging results usually find no neuroanatomical abnormalities (Koto, Horwitz, Suzuki, Tiffany, & Suzuki, 1978). However, recent neurocognitive findings (Davison et al., 2013) from eight individuals with MPS-IVa (aged 5 -17 years) suggested that mild/borderline cognitive impairments do exist. Age appropriate standardised tests (e.g. WASI, WISC) revealed full scale IQ scores either in the lower average range (80 – 90), borderline range (70 – 80), or extreme low range (<70) in four individuals. The remaining four individuals had normal Full Scale IQ (85 – 115). Attention problems were reported by the majority of parents using the Child Behavioural Checklist.

Mild cognitive impairments were supported by MRS findings in the same study. In three individuals with cognitive impairments there was a correlation between white matter metabolite concentrations (N-acetylaspartate) and cognitive indices. In addition, MRI findings revealed neuroanatomical abnormalities in more than half the MPS-IVa individuals. These included mild asymmetry of the lateral ventricles, prominent perivascular spaces and high signal white matter areas of the right frontal lobe. Unlike the MRS findings, there was no correlation with cognitive indices, but this could be because the behavioural measures lacked sensitivity. A formal assessment of attention is needed, especially in the context of the attentional difficulties reported by parents.

Tyrosinemia Type III (T3, OMIM 276710)

Tyrosine is an amino acid that is catabolised into fumarate and acetoacetate, both of which are important for gluconeogenesis and ketogenesis. Dysfunction at different points in this enzymatic pathway will lead to one of three identified hypertyrosinemia disorders: tyrosinemia types I to III, which result in the accumulation of plasma tyrosine levels and increased urine excretion of tyrosine (Chakrapani, Gissen, & Mckiernan, 2012). Tyrosinemia type I is caused by a deficiency in fumarylacetoacetate hydrolase, the final enzyme in the tyrosine catabolic pathway. It leads to an

accumulation of highly toxic fumaryl- and maleylacetoacetate in the liver. This produces organ dysfunction and carcinogenesis, with hepatocellular carcinoma a frequent cause of death in childhood. Tyrosinemia type II, results from a defect in tyrosine transaminase, the first enzyme in the tyrosine pathway. It affects the eyes and skin, but also mental development. Tyrosinemia type III (T3), the disease we are concerned with here, is caused by the deficiency of 4-hydroxyphenylpyruvate dioxygenase, the second enzyme in the pathway. It causes extreme accumulation and increased excretion of tyrosine, but not the increase in the toxic metabolites associated with the later type I enzyme defect (Chakrapani et al., 2012; Scott, 2006). Treatment consists of a low-protein diet and administration of ascorbic acid to control tyrosine levels (Scott, 2006).

T3 is the rarest of the three tyrosinemias and the effects of elevated concentrations of tyrosine in the central nervous system are not well established. Only 15 cases have been reported in the literature to date (Ellaway et al., 2001; Heylen et al., 2012; Szymanska et al., 2015). However, neurological and intellectual difficulties are commonly reported. Ellaway et al.'s (2001) review is the best description of cognitive functioning in 13 T3 individuals. Eight had neurological symptoms, such as developmental delay or mental retardation, attention deficit and behavioural disturbances, acute ataxia, tremor, hypotonia and absent deep tendon reflexes. The most common long-term complication was intellectual impairment (75% of individuals; Ellaway et al., 2001). Ellaway et al. speculated that the intellectual impairments resulted from the neurotoxic effects of elevated tyrosine levels. Their results, however, were based on aggregated scores (e.g. e.g. Stanford-Binet, WISC, WASI), so a profile of specific cognitive impairments remains to be determined.

Tyrosinemia type I is treated with 2-nitro-4-trifluromethylbenzoyl (NTBC), which makes treated type I an indirect window on T3. Treatment stops tyrosine metabolism earlier and prevents the accumulation of toxic metabolites that cause liver cancer, but it also gives T1 individuals the same biochemical profile as T3, raising tyrosine concentrations. T1 is a more common form of tyrosinemia and, in contrast to T3, more detailed reports of neurocognitive function exist (Bendadi et al., 2014; Thimm et al., 2011; Van Ginkel et al., 2016) Different studies have documented lower IQ scores and specific language production and comprehension impairments in NTBC-treated T1 individuals (Bendadi et al., 2014; Thimm et al., 2012).

The most detailed assessment of cognitive function in T1 was recently reported by Van Ginkel et al. (2016), measuring nineteen NTBC-treated individuals with both age-appropriate psychometric tests (e.g. WISC and WASI) and additional measures of executive functioning. Average IQ was lower (median 85: range 55 – 111) and there were specific working memory deficits (in reaction time and error rates). No differences in inhibition were found. The cognitive deficits from the T1/NTBC studies could be due to elevated levels of neurotoxic tyrosine, but they could also be a direct effect of NTBC (Van Ginkel et al., 2016). Results from T3 can help to clarify this.

Taken together, these findings (Bendadi et al., 2014; Ellaway et al., 2001; Thimm et al., 2012; Van Ginkel et al., 2016), give us some reason to predict cognitive impacts in tyrosinemia type III. As with Morquio Syndrome, we will characterise the presence or absence of deficits across cognitive domains (language, attention and ocular-motor control) to show whether there is a homogeneous pattern, or selective deficits in particular areas.

Patterns associated with functional deficits

We probed the cognitive domains of attention, language and oculomotor control to see if deficits selectively affect different domains in different diseases. Beyond this, some deficits predict characteristic patterns across domains. We distinguish deficits in processing speed, selective attention, sustained attention.

A deficit of *processing speed* should be particularly evident in a simple RT task, where cognitive demands are minimized, allowing speed to be a primary influence on outcomes, but effects should be noted in all speed dependent tasks, including conjunction and feature search and saccade onset times. Slower processing speed should produce exaggerated effects of difficulty in tasks where the number of operations is manipulated explicitly, like conjunction search. For the same reason, when dividing reaction times into quintiles, slower reaction times should show larger differences from controls than faster times, although there could be some difference even for faster times.

A deficit of *selective attention* should impact conjunction search specifically, leading to increased effects for larger displays, without changing feature search times, simple reaction time or saccade onsets. Larger differences should be noted in slower quintiles of the RT distribution for conjunction search only.

A deficit of *sustained attention* should lead to shorter times on target in the fixation task and more intrusive saccades. It should also create a subset of longer reaction times in search tasks (due to distraction). Faster reaction times, however, should not be different from controls.

Methods

Participants

Morquio Syndrome (MPS IVa) and Tyrosinemia Type III (T3)

Thirteen individuals with MPS-IVa (8 male; mean age: 9.59 years, range: 5.27 – 14.39 years) and eleven with tyrosinemia type III (3 male; mean age: 12.53 years, range: 4.40 – 19.58 years) were recruited at Birmingham Children's Hospital, UK (demographics in Appendix 1). Diagnosis of MPS-IVa was confirmed via genetic testing. All MPS-IVa individuals had a severe phenotype of the disease, but none had corneal clouding and none was ventilated. Tyrosinemia Type III was confirmed via genetic testing after elevated tyrosine levels were detected during newborn screening. All individuals were native English speakers. A number were bilingual (see Appendix 1). Both groups of children with disease were tested in 2012-13 at the Wellcome Trust Clinical Research Facility at the Birmingham Children's Hospital. Consent was obtained from parents/guardians prior to testing.

Typically Developing Controls (TD)

Individuals with metabolic disease were compared to typically developing controls (TD), using a developmental trajectory approach (e.g. Thomas et al., 2009). We tested a large sample of controls across a range of ages between 6 and 20 to establish the trajectory of typical development in each cognitive domain (Attention tasks N = 104, Language tasks N = 104, Oculomotor tasks N = 265). Prior to testing, informed consent was taken from undergraduate student participants or from parents of child participants.

Apparatus & Procedure

Our battery of test took approximately 1 – 1.5 hours to complete, depending on ability and need for breaks.

Attention and oculomotor tasks were created using Experiment Builder (SR Research Ltd., Mississauga, Ontario, Canada). Participants viewed stimuli from 60 cm on a 36 x 27 cm CRT-monitor

with a resolution of 1024 by 768 pixels and a refresh rate of 60Hz, producing a viewing area of 33.5° x 25.5° (width x height) of visual angle. Reaction time (RT, measured in milliseconds, msec) for manual responses were recorded using a Cedrus button box (http://cedrus.com/).

Eye movements were recorded via an EyeLink® 1000 Tower Mount (SR Research Ltd., Ontario, Canada). Eye position was measured using corneal reflection via an infrared camera. Head movement was minimized using a forehead and chin rest. Eye movements were calibrated to an accuracy of at least 1^{o} using a nine-point calibration array. Drift correction was employed before each trial. Participants were recalibrated if central fixation was inaccurately displayed. Saccade detection was based on default Eyelink 1000 settings (19-sample window): samples were classified as part of a saccade if eye velocity exceeded 22^{o}/second and eye position changed more than 0.3^{o}, otherwise samples were classified as fixations. Prior to analysis, trials from all tasks were visually inspected to ensure that participants were engaged in the task and artefacts had been removed.

Attention tasks

Attention tasks included simple reaction time and visual search. We familiarised participants with the visual search targets (red ladybirds) by introducing them in the simple reaction time task, which was always presented first.

Simple reaction time measured speed to respond to appearance of a visual target. Participants completed 20 trials where a red ladybird (5.5° x 7.5°) was presented in a box in one of the four quadrants around the centre (upper left, upper right, lower left, lower right; box size = 10° x 9°; 12° diagonal distance from screen centre). Quadrants were systematically probed in this task and in the eye-movement tasks because regions of space can be affected by attention or eye-movement deficits (e.g. neglect, usually affecting attention to the left side of space, or vertical supranuclear gaze palsy, affecting vertical eye-movements). The target appeared in each quadrant 5 times, with the order of the target locations randomised between participants. Each trial began with a centrally presented fixation cross (1° x 1°). After a variable delay of 1500–4000 ms the red ladybird target stimulus was presented in one of the four quadrants. Participants were instructed to respond with their preferred hand, using a single button press, as quickly as possible after the target appeared. The stimulus remained on the screen until

the participant made a response or for 3000 ms. There was an inter-trial interval of 1000 ms prior to the presentation of the fixation cross for the following trial.

Visual search is a prototypical task of selective attention that has been studied extensively and used to motivate a number of influential models (Bundesen, Habekost, & Kyllingsbæk, 2005; Moran, Zehetleitner, Muller, & Usher, 2013; Schwarz, 1993; Schwarz & Miller, 2016; A. Treisman & Gelade, 1980; A. Treisman & Sato, 1990; Verghese, 2001; Ward & McClelland, 1989; Wolfe, 2007). Our visual search task consisted of 3 feature search blocks and 3 conjunction search blocks. In all blocks participants indicated, with a yes/no button press, whether the red ladybird character was present. Feature search asked the participants to search for the red ladybird among a set of green ladybird distracters. The target was defined by colour only. This means that the target typically appears to "pop out" regardless of the number of distractors. Instead, conjunction search asked participants to search for a red ladybird among green ladybirds and red beetle distractors. No single feature dimension defines the target (it is red AND a ladybird) and this typically means that items need to be searched more systematically, leading to longer reaction times for larger displays. Each block contained 12 trials with 3 display sizes (4, 8, and 12 items). The target was present in half the trials in each block. Block order was randomised between participants. Search displays were created by dividing the screen into a 4 x 4 grid (each grid location = 8.5° x 6.5°). Search items (4° x 5° of visual angle) were randomly assigned to the 16 grid locations. Each item's position was jittered by a random amount to make displays less regular (x maximum ±2° and y maximum ±1°). For target-present conjunction trials one red beetle distractor was replaced with the target to ensure the number of red and green elements were equal.

Each trial began with a centrally presented fixation cross (1000 ms), followed by the stimulus. Participants used a Cedrus button box to make a response (left button = target absent; right button = target present). Stimuli disappeared after a response or if no response was registered within 10 seconds. A blank screen was displayed (1000 ms) prior to the following trial. At the beginning of each block participants were informed of the block type (feature or conjunction). Targets and distractors were shown on screen and verbal instructions given by the experimenter. There were four practice trials with feedback, at the beginning of the experiment, to familiarise participants with the task.

We used response times for correct responses to calculate a mean visual search reaction time (or intercept). We also measured search efficiency, defined as the slope of RT change with increasing display size.

Language tasks

Measures of verbal production (Boston Naming Task or BNT; Kaplan, 2000) and comprehension (British Picture Vocabulary Scale or BPVS; Dunn et al., 1997) were used to assess children's language abilities.

The BNT is a 60-item picture naming test where items increase in difficulty, becoming increasingly infrequent and unfamiliar (e.g. trellis). If a participant was unable to name an item, a semantic cue was offered. A phonetic cue was provided if the semantic cue failed to produce a correct response. An item was counted as correct if the participant was able to name it with or without a semantic cue, and incorrect if the participant failed to produce the name or did so only after a phonetic cue. The test is stopped if participants make more than 5 consecutive errors.

The BPVS is a receptive vocabulary test. Participants listen to a word and select a matching picture from four alternative line drawings. There is no spoken response. There were 14 sets of 12 items, which, like the BNT, increase in difficulty because the targets are increasingly uncommon items. The test was stopped if there were eight or more errors on a single set.

Oculomotor tasks

Oculomotor function was tested using a fixation and pro-saccade task (Fischer & Weber, 2010; Klein, 2001; Munoz & Everling, 2004; Salman et al., 2006).

Fixation Task

Participants were asked to fixate a central target, look to the target when it moved to one of four possible non-central positions and then maintain their gaze until the target disappeared. There were 20 trials. An elephant face target (1.5^{o} in size) appeared in the centre of the screen and then moved to a location 10^{o} to the left, right, above or below central fixation. Possible target locations were marked with small circles (0.5^{o} in size) to indicate where the target might appear. Trials began with the presentation of the target centrally for 1000 ms. The target disappeared and immediately reappeared (no gap or overlap) randomly at one of the four surrounding locations. The target remained at this location

for 5000ms and participants were asked to maintain fixation on the target until it disappeared. A 3 x 3° box surrounding each target was defined as a region of interest (ROI) and fixations within the ROI were counted as fixations on the target. Dwell time was defined as the length of time participants maintained eye position within the target's ROI (*FixDwell*). The frequency of intrusive saccades away from the target (greater than 2°) that moved the participant's gaze outside the target ROI were also measured (*FixSacc*).

Pro-saccade Task

Participants viewed 48 trials where a target elephant face (1.5^{o} in size) randomly appeared at one of eight positions around a central starting position. The four locations from the fixation task were used (see above), along with four additional locations, closer to the centre (eccentricity = 5^{o}). Each trial began with the stimulus displayed at the centre of the screen for a random amount of time between 1000 and 2000ms. The stimulus then disappeared and reappeared at one of the possible target locations for 1000ms without any gap or overlap. Participants were asked to look as quickly and accurately as possible directly at the target. There was no instruction to maintain fixation.

To be included in the analysis, saccades had to start at the centre of the screen and the eye movement had to be toward the peripheral target. We measured the onset of the first movement toward the target and its peak velocity. To simplify the presentation only data from targets at 10^{o}, where the distance is greater, are presented here, but data from closer targets did not change the pattern.

Data Analysis

Across tasks, reaction times that were greater than 3 SDs away from individual participant means were defined as outliers and removed. Individuals with metabolic disease and typically developing (TD) control children were compared using a developmental trajectory approach (Thomas et al., 2009). Developmental change was modelled as a linear or quadratic function of age. We were interested in differences in the *rate* of developmental change (differences in the slope of trajectories) and differences in absolute levels of performance (offsets between disease group and control trajectories). Data from our experimental tasks was fit using linear mixed effects models (LME, using the R package *lme4*) with fixed factors for *Group*, *Age* and, where appropriate, *Condition* (e.g. target location). A main effect of *Group* indicated an effect of disease across all ages, and a main effect of *Age*

indicated age-related developmental change. Interactions between *Group* and *Age* indicated that disease modified the rate of developmental change. Akaike's Information Criterion (AIC) was used to compare models (Burnham & Anderson, 2002).[1]

The set of models we compare are familiar from more traditional analyses (e.g. ANOVA). For example, in an experiment with factors of *Age*, *Group* and *Condition,* we start from a model whose most complex term is the three-way interaction *Age X Group X Condition* (this model will also contain all lower-level two-way interactions and main effects). We compare this to the set of simpler models that have one or more effects removed. If a simpler model accounts for data as well as a complex model, the missing effects were not important. Our final model includes only effects that are needed to account for the data. The highest order model for each analysis, the one reduced models are derived from, is listed as the *generating model* along with models results in Appendix A. Random effects structures for each analysis are also listed in Appendix A. The list of model results includes the models closest to the minimum AIC model. Models not listed were worse.

We report results from a large sample of controls and a much smaller sample of individuals with metabolic disease, as is inevitable in the study of rare diseases. This means that models that do not involve the *Group* term are largely determined by the data from normally developing control participants. For each task, we initially model only data from typically developing controls to characterise normal development (e.g. to ask whether the effect of Age is absent, linear or quadratic).

[1]Model selection using AIC is different from p-values, but not difficult to understand. AIC is preferred for model comparison because, unlike p-values, AIC balances fit and the number of model parameters when choosing models. In brief, better models produce smaller AIC values, but the absolute AIC values are not interpretable. Instead, the change in AIC (ΔAIC) between models is meaningful and captures the weight of evidence for each model (rather than being subject to a cut-off, like p-values). Evidence for a model starts to be clear if the ΔAIC exceeds 2. If ΔAIC between the "best" model and alternative models is less than 2 then the two models are substantially equivalent. When ΔAIC is between 2 and 10 there is decreasing support for an alternative model. A model with a ΔAIC > 10 has essentially no support. For models where the ΔAIC is less than 2, it is reasonable to favour the least complex model (i.e. model with fewest parameters/variables). Favoured models contain terms that are important in accounting for data. This is parallel to significant effects in an analysis using hypothesis testing. For example, if a highly-rated model has a term for *Group* but no interaction, this is parallel to a significant main effect of *Group* and a non-significant interaction.

Comparisons can be assisted by calculating Akaike weights (AICw; Burnham & Anderson, 2002). AIC*w* expresses the relative probability that a model is the best *in a particular set*, considering only the models from that set. It measures the weight of evidence for the models being compared. When values are relatively equal across two or more models, they are all relatively good models of the data. If one model has a high value and the others are low, there is a model that is clearly better.

Then, using the full dataset, we check interactions between terms in the best model for control data and *Group* (i.e. results from control and disease groups concentrate on evaluating the influence of the term for *Group*).

Because of possible heterogeneity among individuals within each disease group we also compared individuals with disease to typically developing controls (TD). The question was whether each individual was inside or outside the range of typically-developing values. To define the range of typically developing values, 95% prediction intervals were defined at each age, setting a cutoff above and below the TD means. Ninety-five percent of individual control values should fall between these points. Cutoffs for each age were smoothed to a boundary that applied across ages by fitting a curve, separately, to the set of upper boundary points and then lower boundary points. This minimises the impact of idiosyncratic estimates, smoothing the boundary position across ages by borrowing information from adjacent age groups. After defining limits, individuals with disease were compared to controls using the equivalent of z-scores, but using the upper or lower endpoints of the smoothed prediction intervals, rather than the noisier age-specific cuoffs. We label these z_{pi}.[2]

Results

Table 1 summarises domains where Morquio (MPS-IVa) and Tyrosinemia (T3) individuals had difficulties (shaded in grey) and is based on the results described in more detail below. The number of individuals who performed worse than the control mean, worse that 1 SD from the control mean and worse than 2 SD from the control mean are listed. MPS-IVa individuals displayed clear deficits in sustained attention and non-decision aspects of selective attention tasks, e.g. the amount of time fixation could be maintained and reaction time intercepts in visual search. T3 individuals exhibited clear deficits in language tasks (*BNT* and *BPVS*). They also had problems with sustained attention that affected fixation and visual search and modest slowing in simple reaction time.

Table 1 about here

[2] Specifically, for values above the mean, $SD_{pi} = (PI_{upper} - control\ mean)/1.96$ and $z_{pi} = Mean_{patient}/SD_{pi}$, where PI_{upper}= upper boundary of the smoothed prediction interval; *control mean* = the mean predicted by smoothing control mean values; and $Mean_{patient}$ = individual patient's mean value. Values below the mean were calculated in the same way except the 95% prediction interval boundary (PI_{lower}) used the lower boundary since boundaries were not necessarily symmetric.

Attention - Simple Reaction Time Task

Typically developing Controls (TD)

Reaction times were log transformed because, as is often the case, the raw RT distribution was right-skewed and a log transform produced a more balanced distribution. A model with a quadratic term for age was used because the developmental trajectory was curved, not linear (Figure 1). The best model included an effect of age, distinct trajectories for left and right locations and an interaction between location and age. Reaction times were slower to right targets. The left/right difference in younger children was 51 msec and then decreased with age. The majority of developmental change occurred during the first years of development, with little change after age 12 (~300 msec over the range from 6 to 12 years).

Figure 1 about here

Morquio syndrome (MPS-IVa)

RT means for MPS-IVa individuals are displayed in Figure 1a. They were within the confidence limits of healthy development in all cases ($z_{pi} > 2 = 0/12$; black points inside dashed black line, Figure 1). The group mean was shifted towards slower reaction times, with eleven of 12 individuals slower than the control mean (six expected by chance).

In the analysis of group differences, the model with just an *Age² X Condition* interaction was essentially equivalent to the best model (ΔAIC = 0.2; best model: *Condition X Age²* + *Group X Age²*), which means evidence of a group difference was weak (AIC model selection results are summarised in Appendix 2).

Tyrosinemia III

Individual T3 RT means are displayed in Figure 1b. They were evenly distributed across the control range. Two individuals had response times that were reliably slower than controls ($z_{pi} > 2$) when performance was averaged across target locations (patient 7 – 8.03 years, patient 9 – 17.64 years). The two youngest individuals (patient 1 - 4.40 years, patient 2 - 5.16 years) were on the border ($z_{pi} = -1.92$ and $z_{pi} = -1.90$ respectively).

In group comparisons, the best model included a *Condition* X Age^2 X *Group* interaction (AIC model selection results are summarised in Appendix 3). The effect of condition was not strong. There were, however, strong effects of *Age* and *Group*. A model with just main effects of *Age* and *Group* was nearly as good as the best model (ΔAIC = 1.97) and models without these terms had essentially no support (ΔAIC=11.12 and ΔAIC=17.19). The 3-way interaction occurs because T3 individuals are slower than the control mean at the earliest ages, closer to the mean in the middle of the age range and slow again when they are older, but by much less, and then the differences are somewhat different in the different quadrants. These interaction effects are, however, subtle and the small T3 sample makes them potentially unstable.

In sum, orienting and reacting to an abrupt onset was not slowed in MPS-IVa. Most individuals were slower than the mean, but scores were clustered quite closely in the region of the control mean with no individuals outside the control range, so a group difference was not supported. There was some evidence of slowing in T3 and models with a group difference were preferred. RT slowing was concentrated on the youngest and oldest participants. Since the effect is carried by relatively small numbers at the extremes of the age range we should be cautious in our interpretation. Participants between 6 and 15 years old were less clearly different. The lack of differences in MPS-IVa and the relatively modest differences isolated to young and old participants for T3 will influence our interpretation of other reaction time tasks where differences are more marked, although we must keep in mind that speed of processing could be affected in T3.

Visual Search

Typically developing Controls (TD)

Mean reaction time is plotted in the top panels of Figure 2 and search efficiency in the bottom panels. TD controls increased their overall speed with age (top panels) in both conjunction and feature search. The biggest change occurred during the early years, with changes slowing after about 10 years. Feature search efficiencies were close to zero at all ages (although there was some decline in variability). Accordingly, the best model of feature search included an effect of age but no effect of display size. This means the search for a feature difference happened in parallel across all locations in

the display no matter how many items there were to search, which is the normal pattern (A. M. Treisman & Gelade, 1980). The best model of conjunction search, in contrast, included an interaction between display size and age. Reaction times always increased when there were more distractors, but the time to search each item declined steadily from around 70 to around 32 milliseconds (bottom right panel) over the range between 6 and 19 years. There was also a large drop in variability. This indicates that finding the target in conjunction search requires some degree of item-by-item processing.

Accuracy supported the same pattern. Accuracy in feature search increased across ages (from 94% at age 6 to 98% at age 19). There was a difference in accuracy on present vs absent trials (90% vs 98% at age 6) that narrowed with age (99% vs 98% at age 19) and models supported an interaction between present/absent and age. There was no effect of display size. The best model of conjunction search accuracy included an interaction between display size and target present/absent and an interaction between age and present/absent. The interaction with age was weak because a model with just a main effect of age was nearly equivalent (ΔAIC = 1.23). The interaction with display sizes occurs because more targets are missed in larger displays. Accuracy increased with age (85% at age 6 vs 97% at age 19). These patterns do not modify the interpretation of the reaction time results (where larger displays take longer).

Figure 2 about here

Morquio Syndrome (MPS-IVa)

The biggest difference between individuals with MPS-IVa and controls occurred in the simpler task of feature search. z_{pi} was greater than two in six individuals and between one and two in one other individual. Eleven of 12 individuals were slower than the control mean. Six would be expected by chance. In conjunction search, z_{pi} was greater than two in three individuals and between one and two in four others. Ten of 12 individuals were slower than the control mean.

There were no abnormal increases time to search larger displays (i.e. display size effects were normal). In feature search, additional items took longer in three individuals (Patient 1, z_{pi} = 2.46; Patient

8, z_{pi} = 5.99; patient 11, z_{pi} = 11.25), and, in conjunction search, in two (Patient 8, z_{pi} = 2.57; Patient 11, z_{pi} = 3.55).

Group differences between MPS-IVa individuals and controls were analysed separately for feature and conjunction search. The best model of feature search did not include an effect of *Display Size* and there were no two or three-way interactions. There was a main effect of *Age* (younger participants were slower) and a main effect of age (MPS-IVa individuals were 445ms slower than controls). Models without terms for *Group* and *Age* were poor.

The lack of interaction between *Group* and *Age* shows that MPS-IVa individuals make progress at a normal rate, neither catching up with controls or falling behind, despite the slower overall times. We should be somewhat cautious about this result, however, because there is a small number of individuals at each age.

In the analysis of accuracy, *Group* was added to the *Age* by *Present/absent* interaction from the control model. Instead of the increasing accuracy seen in controls, MPS-IVa individuals initially show a bias to "present" and later a bias to "absent" with little consistent change to percent correct. This is not a theoretically important pattern in the present context. Averaging over present/absent conditions, the best model included an interaction between *Age* and *Group* that resulted from increasing overall accuracy in controls, but flat or slightly decreasing accuracy in MPS-IVa individuals. MPS-IVa individuals and controls had equivalent accuracy when averaging across ages (both 94%). Accuracy does not change the interpretation of reaction times. Accuracy is high in both MPS-IVa individuals and controls and there was no indication that longer RTs were associated with higher accuracy in the MPS-IVa group (R = -0.46, p=.13; a negative R value results when higher accuracy is associated with *lower* RT).

The top four models of conjunction search were all very similar in their ability to account for reaction times (Display size X Age + Group X Age; Display size + Group X Age; Display size X Age + Group; Display Size + Age + Group; max change in AIC = 0.8). Good models all included terms for *Display Size*, *Age* and *Group*. Interactions between *Group* and *Age* and between *Display Size* and *Age* were weak because the simpler model with only main effects was a close equivalent (ΔAIC=0.8).

It is important to note the lack of interaction between *Display Size* and *Group*. MPS-IVa individuals processed each item as quickly as controls, even when efficiencies were above zero and each item required additional processing time (~42 ms per item in both MPS-IVa and controls). The *Group* effect, instead, resulted from a constant amount of slowing at all display sizes (~378 ms for both large and small displays). Models without main effects of *Display Size*, *Age* or *Group* were poor.

In sum, MPS-IVa individuals were clearly slower in visual search. This difference, however, was constant across display sizes. Item-by-item processing was not slowed. The upward shift in reaction times without display size effects could have at least two origins. One possibility is that MPS-IVa individuals occasionally lose concentration or are distracted. If so, reaction times at the fast end of the range should be similar to controls, but there should be more difference in slower reaction times. Lapses could be distributed across small and large displays. A second possibility is that a stage that is not sensitive to attentional demands, is affected (a non-decision stage, in the terminology of search models). This could be time to initiate search or time to select and program a response.

To explore these alternatives, we binned reaction times into quintiles for each participant and compared quintiles across groups using the vincentizing method (Ratcliff, 1979; see Romani, MacDonald, De Felice, & Palermo, 2017 for this kind of analysis in a population of individuals with PKU). Figure 3 shows reaction times and standardised differences between means for individuals with MPS-IVa and controls in each quintile. To standardise, we divided the difference between MPS-IVa and control means by the standard error of the control mean. Because reaction time distributions have an extended tail associated with the longest reaction times, the variance in the final bin is higher than the other bins, and interpretation of quintile differences needs to take this increased variance into account. We will concentrate on two aspects of the difference curves. We will ask, first, if MPS-IVa quintile means are slower than controls (difference scores > 0). We will also ask if differences are increase across quintiles, particularly quintiles 2, 3 and 4, showing increasing differences with longer reaction times.

Figure 3 about here

With one exception, there were relatively stable differences across the full range of quintiles in MPS-IVa, not increasing differences. In feature search at ages 5-8, conjunction search at ages 8-10.5 and both feature and conjunction search at ages 10.5-14 reaction times were always slower, even in the fastest bins ($t>2.3$; $p<.03$) but differences did not increase. Feature search at ages 8-10.5 was the only condition with increasing differences (conjunction search at ages 5-8 showed no differences, all $t< 1.8$, all $p>0.08$), The overall pattern does not support distraction/failures of attentional control or the accumulating effects of general slowing. Instead, initiation or response processes are possible loci.

Tyrosinemia III (T3)

In feature search, four of eleven T3 individuals had $z_{pi} > 2$ for mean RT. An additional two had z_{pi} between 1 and 2 (Figure 4). In conjunction search, five of eleven individuals had $z_{pi} > 2$ and two had z_{pi} between 1 and 2. Only one individual had a search efficiency z_{pi} greater than two and three had z_{pi} between 1 and 2.

Figure 4 about here

Group differences were analysed separately for feature and conjunction search. The best model for feature search included an *Age* by *Group* interaction. There were more individuals further from controls at older ages. There was a ~76 ms difference between groups when participants were under 12 years old, with as many T3 individuals below the control mean as above it. Individuals older than 15 years were, on average, ~187 ms slower than controls, and all four T3 individuals were above the mean. There was no effect of *Display Size*, indicating that feature search was efficient for both groups.

The best model for accuracy in feature search included an interaction between *Present/absent, Age* and *Group,* adding *Group* to the interaction from the control model. As with MPS-IVa, there was a different pattern of *Present/absent* effects compared to controls. Accuracy was relatively constant, but *Present* trials showed some decrease. Overall, T3 individuals were less accurate than controls (89% vs 93%). There was no evidence of a speed/accuracy trade-off that would modify interpretation of the reaction time results.

In conjunction search, the best model of reaction times included main effects of *Display Size*, *Group* and *Age*, but no interactions to suggest group differences in display size or age effects. The group difference was an average of 71 msec when participants were below 12 years old and 133 ms when participants were above 15 years old.

Models of accuracy in conjunction search included a *Group* by *Display Size* by *Age* by *Present/absent* interaction. The *present/absent* variations are unlikely to be theoretically important. When present/absent conditions were collapsed and individuals put in two age groups to increase N (less than or greater than 12 years old), models required a *Group* by *Display size* by *Age* interaction. Accuracy was lower than controls (80% vs 90%) and the display size effect was larger in younger individuals only (accuracy dropped by 12% when young individuals viewed larger displays, but only by 1-3% in older T3 individuals, young controls and old controls).

We examined the source of slower RTs using the quintile method described above. In contrast to MPS-IVa, T3/control differences increased across quintiles (Figure 5). At ages 4.5-6, differences in both search tasks increased in bins 1-3 and then levelled out. At ages 6-8.5 there were no differences from controls. At ages 15-19, differences increased across all bins in both search tasks. In the youngest and oldest T3 individuals, reaction times were different even in the first bin (all $t>2.3$; $p<.03$), showing that T3 individuals in these groups were always slower, but the differences increased for longer reaction times. Different patterns at different ages should be interpreted with caution because the number of individuals at each age is small.

Figure 5 about here

In sum differences in visual search were sometimes observed more in accuracy and sometimes more in reaction times, but there was a relatively consistent difference. At the level of individuals there was considerable variability. Seven of 11 individuals recorded reaction times that were slower than the control mean in feature search and seven of 11 in conjunction search (not always the same individuals). This was not a marked difference because the value expected by chance is 5.5 (recall that in MPS-IVa nearly all individuals were above the control mean). There was some indication that performance got

worse with age. This was, however, based on small numbers, so we should be cautious about how strongly we interpret the effect. RT quintiles showed increasing differences for longer reaction times. This was not consistent with increased processing for more difficult displays because 1) the pattern was the same in both feature and conjunction search (difficulty does not increase with display size in feature search) and 2) conjunction search efficiencies (which measure increasing difficulty) were normal. It is possible that occasional long reaction times were based on distraction, but this would not account for the differences in the first quintiles. There may also be slowing which affects all quintiles, as in MPS-IVa, during initiation or response stages. We will return to the issue of distraction when we discuss the fixation task.

Language

Typically developing Controls (TD)

Control participants showed improvement with age in both the BPVS (ΔAIC = 163) and the BNT (ΔAIC = 58; Figure 6).

Figure 6 about here

Morquio Syndrome (MPS-IVa)

Patient 12 did not complete the Boston Naming Test (*BNT*) because of time constraints. *BPVS* and *BNT* scores are shown in Figure 6a. On the Boston Naming Test, one individual scored below the control mean (patient 6; $z_{pi} = -3.17$) and one scored above (patient 7; $z_{pi} = 2.74$). Three other individuals had z_{pi} between -1 and -2. The best model for the Boston Naming Test included only a term for changes with age. There was no evidence that MPS-IVa individuals performed worse than controls.

No *BPVS* scores were further than two standard deviations below the mean, but 7/12 individuals were between one and two standard deviations below (Figure 7). The best model for *BPVS* scores included an interaction between *Group* and *Age*, but evidence for the interaction was weak because the model without the interaction was nearly equivalent (ΔAIC = 0.7).

In sum, MPS-IVa individuals showed no difference from controls in productive vocabulary and only minor differences in receptive vocabulary with no individuals or very few outside the control range.

Tyrosinemia III (T3)

Just over half of the T3 individuals were more than two standard deviations below the mean on the *BNT* (6/10 individuals). The best model for *BNT* results included an interaction between *Age* and *Group*. The difference between T3 individuals and controls widened with age (Figure 6b). The difference at the earliest ages was 18, but widened to 56 in the oldest individuals.

On the *BPVS*, the majority of T3 individuals were more than two standard deviations below the mean (8/11 individuals). The best model included an interaction between *Age* and *Group*. Once again, the difference between T3 individuals and TD controls got larger with age (Figure 8). At the earliest ages (6 years) the estimated difference was 32 points. By 19, the estimated difference had grown to 63 points. Models without terms for either *Age* or *Group* were poor.

Unlike MPS-IVa, where the differences from controls were small, T3 individuals show clear problems with language tasks that become more marked with age.

Oculomotor Tasks

Oculomotor tests included a fixation task and a saccade task. Morquio patients 5 and 7 did not want to attempt the saccade task and patients 2, 3, 4 and 11 did not complete all saccade trials due to fatigue. T3 patient 2 was unable to complete the oculomotor tasks due to postural difficulties which prevented eye tracking.

Fixation task – Dwell time

Typically Developing Controls

Dwell time is the total time spent fixating the target during a trial (summing initial fixation and re-fixation times). Dwell time in controls was described by a quadratic model with an interaction between *Condition* and Age^2 (Figure 7). Systematic modelling `revealed some advantage for top targets, but this was subtle.

Figure 7 about here

Morquio Syndrome (MPS-IVa)

The majority of MPS-IVs individuals exhibited fixation dwell time deficits (Figure 7a; $z_{pi} < -2$ in 8/11 individuals). The best model of MPS-IVa and control data included all 2-way interactions. The *Group X Condition* interaction reflected differences between quadrants that were larger for MPS-IVa individuals than controls (control maximum difference, top vs bottom = 160 msec; MPS-IVa maximum difference, top vs bottom, 906 msec). The *Group* X Age^2 interaction reflected larger differences between groups at younger ages. The model without this interaction was nearly equivalent, however, so the difference was weak (ΔAIC=2.1). The strongest difference was that MPS-IVa individuals fixated the target for considerably less time than controls (difference = -1083 msec), and a model without the *Group* term was poor (*Condition X* Age^2 model, ΔAIC = 55).

Tyrosinemia Type III (T3)

Results are presented in Figure 7b. As a group, T3 dwell times were shifted below the control mean (Average $z_{pi} < 0$ in 8/10 individuals), with four individuals who were different enough to be completely outside the control range. The best model of T3 and control data included interactions for *Condition* X Age^2 and *Condition* X *Group*. The *Condition* X *Group* interaction reflected differences between quadrants that were larger in T3 individuals (maximum difference of 158 msec in controls compared to 528 msec in T3). There was a group difference (model without Group, ΔAIC = 17), but it was smaller and carried by fewer individuals than in MPS-IV.

Fixation Task – Intrusive Saccades

Typically Developing Controls

The number of intrusive saccades, where fixation moved away from the target during the period when participants were supposed maintain fixation, was defined by a quadratic function (Figure 7). where the majority of age-related change occurred before 10 years. A model which included a *Condition* X Age^2 interaction was clearly better than a model that did not (ΔAIC = 8). There were some minor differences due to location (fewer intrusive saccades for top targets) but these were subtle.

Morquio syndrome (MPS-IVa)

Seven MPS-IVa individuals made more intrusive saccades than controls (7/11 $z_{pi} > 2$; Figure 7a). These seven also had shorter dwell times. Dwell time would be expected to be shorter when there are frequent intrusive saccades. One individual, however, was within the control range for intrusive saccades but not dwell time (patient 7). They held their first fixation on the target for a shorter time but did not go back and forth between the target and other locations. Examples of saccade and fixation deficits in two individuals are shown in Figure 8.

Figure 8 about here

The best model of group differences included a main effect of *Group* and an interaction between Age^2 and *Condition*, showing that intrusive saccade suppression was poorer in MPS-IVa compared to controls.

Tyrosinemia III (T3)

Unlike MPS-IVa, intrusive saccade frequencies were normal in the majority of T3 individuals (Figure 7b; 8/10 $z_{pi} < 2$). Two older individuals (patient 9 and 11) had more substantial problems across target locations. Importantly, patient 9 also exhibited fixation time deficits, and it is clear from the eye movement traces that fixation on the target was disrupted by intrusive saccades towards the screen centre. Group differences were best explained a model with *Group* X Age^2 and *Condition* X Age^2 interactions, but a model with without an effect of *Group* was essentially equivalent (ΔAIC = .05).

Pro-saccade Task – Saccade Onset time

Typically Developing Controls

Figure 9 plots saccade onset times for TD controls. Saccade onset time decreased with age and the majority of change occurred between 5 and 12 years. The data were log transformed because residuals increased at higher values. There was a clear preference for a model that included an *Age* X *Condition* interaction. A model with separate trajectories for top and bottom targets, and a shared trajectory for horizontal targets was best. The location that took longest to react to at 6 years old was the

bottom target (max bottom, 307 ms vs min right, 247 ms). At older ages the differences narrow and the order changes (max top, 186 ms vs min right, 161 ms at 19 years).

Figure 9 about here

MPS-IVa (Morquio syndrome)

Individual MPS-IVa saccade onset values are shown in Figure 9a. All MPS-IVa individuals were within the control range. The best model included an interaction between *Group* and *Condition* and *Age* and *Condition*. Group differences varied some by target location, but there was not an interpretable pattern and there was no evidence of a main effect of group independent of the interaction with condition (The model with just an *Age* X *Condition* interaction was better than the model that included an *Age* X *Condition* interaction and a main effect of *Group*, ΔAIC = 1.39; means collapsing across location differed by less than 1 msec).

Tyrosinemia type III

Individual T3 saccade onset values are shown in Figure 9bf. Averaged across target locations, only one individual (patient 9, 17.98 years) had a $z_{pi} > 2$. This individual also had fixation time and intrusive saccade deficits. Models of saccade onset did not support a group difference.

In sum, neither metabolic disease group was systematically slower to initiate saccades.

Pro-saccade Task – Saccade Peak Velocity

There was no evidence of differences in saccade velocity when either disease group was compared to controls. The preferred model for MPS-IVa did not include a term for *Group*. The T3 model without a *Group* term was essentially equivalent to the best model (ΔAIC=0.6).

Discussion

We examined cognitive functioning in two rare inherited metabolic diseases – Morquio syndrome (MPS-IVa) and tyrosinemia type III. We compared individuals with metabolic disease to controls using formal tools for model selection and a developmental trajectory approach. This provided

a richer way of evaluating developmental differences than traditional age-matched comparisons. We were able to separate different types of developmental effects, including changes to the *rate* of development and constant decrements in performance in the presence of a normal rate of development. We included both analysis of group performance and individual comparisons (based on prediction intervals). These contributed different, but complementary information.

In both diseases we found cognitive effects, but with different functional profiles, showing that metabolic disease creates specific cognitive effects, and does not just produce homogenous and general decrements in performance. Sustained attention was affected in MPS-IVa and language, especially, but also sustained attention in T3. These are potential candidates for sensitive markers of disease progression and they document selective functional impacts that can be considered alongside descriptions of biological changes. A summary of effects across cognitive domains is shown in Table 1.

MPS-IVa (Morquio Syndrome)

Few studies have investigated cognitive performance in MPS-IVa (Morquio syndrome), primarily because clinical observations report normal intellectual function (Dvorak-Ewell et al., 2010; Wraith, 2006). A recent study, however, reported mild cognitive changes (Davison et al., 2013) and highlighted difficulties with attention. We did more detailed cognitive testing and we also found attention deficits in several tasks. Fixation dwell time was lower and disrupted by intrusive saccades. MPS-IVa individuals had difficulty sustaining attention after the natural engagement created by an abrupt visual onset had ended. Visual search times were slower, but without a decrease in search efficiency. Item-by-item processing was normal. Analysis of the reaction time distributions showed that both fast and slow search times were delayed and differences did not increase for slower reaction times. This was *not* consistent with occasional lapses of attention, which should introduce larger differences for slower reaction times. There was no evidence of general slowing since simple reaction time, saccade onset times and saccade velocities normal or only weakly affected and differences did not increase with difficulty (slower RTs were not more delayed and search efficiencies were normal). Instead, search results point to delays in a stage that is unrelated to attentional demands, like task preparation or response decision, which would slow both fast and slow responses. We did not find a

natural connection that links results from the fixation task and visual search. Each task suggested a different deficit. If there was a connection between them this is not currently clear.

An offset in search times, with slower times overall, but not slower item-by-item processing (i.e. unchanged effects of display size) has been reported under several conditions in the literature, most importantly, under a working memory load and in aging. When a non-spatial working memory task must be completed at the same time as visual search, search slows but search efficiency is unchanged (Oh & Kim, 2004; Woodman, Vogel, & Luck, 2001). Older adults also show slower search times, without increased display-size effects (Gorman & Fisher, 1998; Monge et al., 2017; see also Ratcliff, Thapar, & McKoon, 2006 for a similar increment in a discrimination task that is attributed to "non-decision" aspects of the task). Reduced capacity could be a factor in common between aging and dual tasking and may also explain our results, but note that this cannot be the same as general slowing since in MPS-IVa both simple RT and saccade onset times were either unaffected (saccade onset time) or only weakly affected (simple RT).

The fixation task may index the difficulties with sustained attention reported by parents in the Davison et al. (2013) study. Difficulties concentrating could be the everyday consequence of the same problem that reduces fixation times during formal testing. Indeed, maintaining fixation involves mechanisms that are not isolated within the system for eye-movement control. The areas that are responsible for maintaining fixation (Krauzlis, Goffart, & Hafed, 2017), including areas in dorso-lateral prefrontal cortex, the frontal eye fields and the superior colliculus, are areas that have connections to other modalities: auditory or proprioceptive, in the case of the frontal eye fields (Medendorp, Buchholz, Van Der Werf, & Leoné, 2011); or movement control, in the case of the superior colliculus (Krauzlis et al., 2017). The frontal eye fields, for example, are thought to "encode 'supramodal' representations to guide attention and behaviour" (Medendorp et al., 2011). Problems in these areas could plausibly the source of general attentional issues.

In sum, our results highlight difficulties in MPS-IVa with sustained attention and with task preparation or response decisions. They extend the preliminary results reported by Davison et al. (2012) and they show that the differences in attentional tasks occur without commensurate changes in other cognitive tasks (e.g. language tasks; item-by-item processing in selective attention) or general

slowing (as reflected in normal or near normal simple reaction time, saccade onset and saccade velocity).

Tyrosinemia III (T3)

In Tyrosinemia III, there were clear deficits in both language production (*BNT*) and comprehension (*BPVS*) and differences from controls appeared to widen with age.

There were also differences in average search times, simple RT and dwell time. Item-by-item processing in visual search was not affected (the same pattern as MPS-IVa), but the differences between T3 and controls were more marked for longer reaction times, consistent with occasional attentional failures (different from MPS-IVa). Dwell times in the fixation task also pointed to problems sustaining attention. There was some evidence of slowing in simple RT. Slowing, like lapses in attention, predicts differences that increase with difficulty, consistent with the conjunction search results. Slowing, however, does not provide a good account other aspects of the pattern because: 1) Display size effects in conjunction search were normal and slowing predicts larger differences for more difficult conditions; 2) Simple feature search produced the same pattern as conjunction search--increasing differences with slower RTs--but larger displays are not more difficult in feature search, showing that slowing occurs even when there is no change in difficulty; 3) There was no evidence of slowing in saccade onsets or saccade velocity. Search RT differences were larger for slower RTs, but even the fastest T3 search RTs were slower than controls. This means that lapses in attention do not explain the whole T3 pattern and, like MPS-IVa, there must be problems that are unrelated to attentional demands.

To our knowledge, this study represents the most detailed cognitive assessment of T3 to date and highlights, especially, language impairments, but also problems with sustained attentional and age-related decline. Our results provide formal tests to confirm and extend the case-based findings of Ellaway et al. (2001), who described mild to moderate intellectual impairments in a sample of 12 individuals, but without specific information about cognitive testing.

The prevalence of language deficits is a frequently-reported cognitive feature of T1 individuals treated with NTBC (Bendadi et al., 2014; Thimm et al., 2011). Since NTBC treatment blocks the biochemical pathway at the same place it is blocked in T3, treated T1 resembles T3, except that T1 individuals are subject to both high levels of tyrosine and potential collateral effects of NTBC.

Our results suggest that the cognitive impairments in T1 are related to the high levels of tyrosine rather than NTBC. Several lines of evidence converge on this. Our T3 individuals have raised tyrosine but no NTBC and they have similar impairments to treated T1 individuals. Learning impairments are present in the mouse model of T1 after NTBC treatment, but not in healthy mice who are given NTBC, suggesting that raised tryosine is the source of impairment, not NTBC (Hillgartner et al., 2016). Finally, Ellaway et al. (2001) reported that T3 individuals who were put on a low tyrosine diet later had more severe symptoms than individuals who began a diet earlier and Ellaway et al. speculated that elevated tyrosine levels in cerebrospinal fluid (CSF) were particularly damaging during infancy. We did not analyze biochemical data, but some of the variance in the cognitive performance in our sample could result from individual differences in tyrosine levels that, in turn, could be based on variations in early exposure or dietary compliance. Relating tyrosine concentrations at different ages to cognitive performance will be an important step toward understanding how and when tyrosine affects the brain and when dietary control is critical for effective treatment.

The behavioural consequences for T1 mice treated with NTBC are worth noting in relationship to the language deficits we found. Hillgartner et al (2016) found that NTBC treated T1 mice were slower to learn in a maze task. Both the BNT and the BPVS are tests of vocabulary. That is, they test the facility with which words are processed *or stored* in the mental dictionary. Both tests would be sensitive to problems with word learning. It could be word *learning*, rather than a deficit affecting phonological or lexical representations directly, that is the source of differences in our T3 individuals.

The cognitive neuropsychology of metabolic disease

Our results show that neurodegeneration from metabolic disease is not a uniform process. We did not find homogeneous declines across diseases and tasks. Patterns from were different. There was a diversity of preserved and impaired capacities across language, sustained attention, selective attention and processing speed even if the precise underlying mechanisms are difficult to pinpoint at this stage. Since the different metabolic diseases affect different chemical systems in the brain, this shows that neuro-cognitive systems are not just differentiated by anatomical location. Brain networks critical for different cognitive functions are differentially sensitive to specific cellular support systems. We can begin to ask "why?" Are there systems, for example, that are particularly sensitive to neural

transmission speed, network synchronization, specific neurotransmitters, growth or membrane changes—all properties that would be directly influenced by the biochemical environment? What is the *functional* aspect of the system that makes these characteristics critical? This is a relatively unexplored dimension, within neuropsychology, that is orthogonal to the understanding of anatomical loci and networks. These questions will be critical, however, to understanding the brain as an integrated biochemical and anatomical system and they will be critical to our understanding of disease.

For MPS-IVa, what we can say about the relationship between metabolic processes and cognition is limited because it has typically been described as a condition that doesn't have cognitive impacts. As a result, there has been little effort to explain how a deficiency in *N*-acetylgalactosamine-6-sulfatase (GALNS) might affect the brain. Davison et al. (2013) were probably the first to offer some possibilities. GALNS breaks down karatan sulfate and chondroitin-6-sufate, and these substances help coordinate neuroaxonal connection formation (Miller, Sheppard, & Pearlman, 1997). In addition, MPS IVa may affect calcium signalling, and connections between calcium signalling, mitochondria and neurodegeneration have been made for Huntington's disease, amyotrophic lateral sclerosis, Alzheimer's disease and Parkinson's disease (Marambaud, Dreses-Werringloer, & Vingtdeux, 2009). Finally, calcium signalling plays an important role in long term potentiation and excitotoxicity, two additional routes to cognitive effects (Marambaud et al., 2009).

Tyrosinemia type III, allows more specific hypotheses. As we have noted, T3 involves a defect in the enzyme which is needed to break down tyrosine, and tyrosine is a precursor of both dopamine and norepinephrine. This creates several avenues for cognitive effects, as outlined by Hillgartner et al. (2016) in their discussion of NTBC-treated T1: 1) increased tyrosine; 2) increased dopamine (produced by conversion from tyrosine); 3) decreased large amino acids (because they are outcompeted by tyrosine, including tryptophan, the precursor of serotonin); or 4) decreased serotonin (due to lack of precursor). Hillgartner et al., however, go on to hypothesize that cognitive problems in NTBC-treated T1 are unrelated to any of these primary effects, and are caused, instead, by two by-products of defective tyrosine break-down: one that is neurotoxic (succinylacetone) and the other that is both neurotoxic and causes demyelination (δ-ALA). This is possible if, as Hillgartner et al. speculate, NTBC is poorly transported into the brain, and tyrosine catabolism in the brain proceeds past the point

where NTBC treatment stops it in the rest of the body. The problem with this hypothesis is that it cannot explain cognitive effects in T3. The enzyme that is defective in T3 acts at the same point that NTBC acts, but it will be defective in all cells, including those in the brain. The toxic by-products that Hillgartner et al., identify are not produced in T3, but our results show that there is still cognitive impairment, pointing to some combination of the primary biochemical effects as the cause.

Cognitive neuropsychology of development

The study of cognitive development in populations with inherited metabolic disease presents both challenges and opportunities (see also, for example, Pitchford & Funnell, 1999). Development involves changes over time and also potential dependencies among changing capacities. Theories of development must specify the time course of intrinsic or environmentally conditioned change, the environmental inputs that are required for development, and potential dependencies among capacities. The pattern of dependencies that support successful development is not guaranteed to persist when the system is mature, which means that development cannot be transparently inferred from either the mature system or insults to that system.

The complexity of the relationship between adult and developmental models can be illustrated, in terms that will be familiar to readers of *Cognitive Neuropsychology,* in the domain of reading. It is relatively clear, from cognitive neuropsychological results, that sub-lexical and lexical mechanisms for reading are largely independent components in the mature system and all major reading models include them (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999; Zorzi, 2010). The relative independence of these systems in adulthood, however, does not mean that they are independent during learning. This is apparent from developmental dyslexia (e.g. Hulme & Snowling, 2016) or more directly, in the context of the current discussion, from acquired dyslexia after a stroke in childhood (Pitchford & Funnell, 1999). As they have in other areas of neuropsychology, the consequences of brain damage can reveal properties of components and dependencies in a *developmental* (rather than a static) model that are not available in other ways. Although there are many studies of developmental cognitive difficulties, the approach that involves characterising

components and dependencies in a detailed *developmental* cognitive model, of the sort that readers of *Cognitive Neuropsychology* would find familiar, is relatively underdeveloped.[3]

Our results are too preliminary to show this promise except in outline, but they do show that developmental effects from metabolic disease can be selective enough to allow tests of functional models and they illustrate both some methodological issues and some approaches that offer solutions.

One of the basic problems is that development is a moving target. The statistical models of developmental trajectories that we have introduced here provide some solutions. They allowed us to compare individuals in disease groups and controls across a range of ages when performance is changing and to separate offsets, which create differences across all ages, from changes to developmental rates. Age did not substantially modify the group differences in visual search or fixation tasks for MPS-IVa. There were differences at all ages. Age, instead, did modify T3 language results, where there appeared to be age-related decline.

Capturing developmental trajectories becomes an even more powerful tool when longitudinal data are available. Trajectories from individuals with disease can be compared to control trajectories. Variance in *rates* of development will show whether there is, for example, consistent decline, or heterogeneity in the impact of disease over time. If individuals need to be tracked clinically, longitudinal trajectories could identify individuals who are not following the normal course of development before their performance is different from controls at any single point in time, providing a more sensitive test of disease progression.

The issue of sensitivity and the contrast between what is apparent in individual data and group data is already clear with our samples. Effects in some individuals that do not occur in all individuals in the group are common. Data from a group, however, can also show that there are deficits even when the individual data do not. Conjunction search results for MPS-IVa were a case in point. Only 3 individuals were outside the control range, but the population, as a whole, was clearly shifted toward slower RTs, with only two individuals faster than the control mean. There is a very low probability that

[3]This is not to deny that lexical and sub-lexical mechanisms can display dissociations also during development (see summary in Castles, Bates, & Coltheart, 2006), but these dissociations do not exclude the possibility that there are also learning *dependencies* between components that could be revealed by developmental studies in a way that is not possible in adults (as Castles, Bates & Coltheart, 2006, note, p. 881, and as the methods of Pitchford & Funnel, 1999, illustrate).

so many scores would be slower the control mean in an unimpaired sample. This is clinically important because it highlights the limits of individual samples for some questions (i.e. is there impairment from disease in this group or not?) and also reinforces the point that techniques like longitudinal sampling may be needed to detect decline at the earliest possible moment in individuals.

The statistical approach we have adopted is a powerful method to deal with developmental data and with relatively small samples. There are, however, limits that small samples, which are inevitable in rare diseases, impose. We cannot, for example, say very much about whether a group trajectory derived from cross-sectional data is a good representation of how diseases progress in *individual*s, and we have avoided fitting a trajectory to the disease groups. In a small sample, especially with relatively high variability, the precise shape, inclination or location of the trajectory can be highly unstable and shows, again, the need for longitudinal data. When the changes over time are clearer, as in the T3 language data, the statistical models give stronger support to the interaction between group and age. Even when this is the case, however, the precise location or shape of the trajectory may be relatively uncertain.

Conclusion

Our study introduces new data and methods that illustrate the promise of cognitive assessment for inherited metabolic diseases. Two cohorts showed that specific measures were sensitive in each disease (e.g. sustained attention in MPS-IVa or language in T3) and could be developed further into new tools to help track disease progression or quantify treatment benefits. In addition, we have shown how cognitive assessment can help to reveal the biochemical basis of cognitive effects. Finally, our results make several extensions possible. Cognitive performance can be related to changes in structure, activity or biochemistry as measured by MRI, fMRI, MRS or EEG to triangulate the biological basis of cognitive changes. Our analysis of different aspects of performance show, however, that a detailed *functional* description of cognitive processes will remain a central component of this project. Extending the methods described here to longitudinal data is a promising next step for understanding the cognitive impacts of inherited metabolic diseases.

References

Bax, M. C., & Colville, G. a. (1995). Behaviour in mucopolysaccharide disorders. *Archives of Disease in Childhood*, *73*(1), 77–81. https://doi.org/10.1136/adc.73.1.77

Bendadi, F., De Koning, T. J., Visser, G., Prinsen, H. C. M. T., De Sain, M. G. M., Verhoeven-Duif, N., … Van Hasselt, P. M. (2014). Impaired cognitive functioning in patients with tyrosinemia type i receiving nitisinone. *Journal of Pediatrics*, *164*, 398–401. https://doi.org/10.1016/j.jpeds.2013.10.001

Biernacka, M., Jakubowska-Winecka, A., & Tylki-Szymanska, A. (2010). The development of cognitive functions in children with Hurler phenotype mucopolysaccharidosis type I on enzyme replacement therapy with laronidase. *Pediatric Endocrinology, Diabetes and Metabolism*, (1), 249–254.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. *Psychological Review*, *112*(2), 291–328. https://doi.org/10.1037/0033-295X.112.2.291

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (Vol. 2). New York: Springer.

Castles, A., Bates, T., & Coltheart, M. (2006). John Marshall and the developmental dyslexias. *Aphasiology*, *20*(9), 871–892. https://doi.org/10.1080/02687030600738952

Chakrapani, A., Gissen, P., & Mckiernan, P. (2012). Disorders of Tyrosinemia Metabolism. In J. . Saudubray, van den Berghe, & J. H. Walter (Eds.), *Inborn Metabolic Diseases: Diagnosis and Treatment* (Fifth Edit, pp. 233–243). Springer. https://doi.org/10.1007/978-3-642-15720-2

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037//0033-295X.108.1.204

Davison, J., Kearney, S., & Horton, J. (2013). Intellectual and neurological functioning in Morquio syndrome (MPS IVa). *Journal of Inherited Metabolic Disease*, *32*(2), 323–328.

De Laet, C., Munoz, V. T., Jaeken, J., François, B., Carton, D., Sokal, E. M., … Goyens, P. J. (2011). Neuropsychological outcome of NTBC-treated patients with tyrosinaemia type 1. *Developmental Medicine and Child Neurology*, *53*(10), 962–4. https://doi.org/10.1111/j.1469-8749.2011.04048.x

Dvorak-Ewell, M., Wendt, D., Hague, C., Christianson, T., Koppaka, V., Crippen, D., … Vellard, M. (2010). Enzyme replacement in a human model of mucopolysaccharidosis IVA in vitro and its biodistribution in the cartilage of wild type mice. *PloS One*, *5*(8), e12194. https://doi.org/10.1371/journal.pone.0012194

Ellaway, C. J., Holme, E., Standing, S., Preece, M. a, Green, a, Ploechl, E., … Leonard, J. V. (2001). Outcome of tyrosinaemia type III. *Journal of Inherited Metabolic Disease*, *24*(8), 824–32.

Fischer, B., & Weber, H. (2010). Express saccades and visual attention. *Behavioral and Brain Sciences*, *16*(03), 553. https://doi.org/10.1017/S0140525X00031575

Gorman, M. F., & Fisher, D. L. (1998). Visual Search Tasks: Slowing of Strategic and Nonstrategic Processes in the Nonlexical Domain. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53B*(3), P189–P200. https://doi.org/10.1093/geronb/53B.3.P189

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. https://doi.org/10.1037//0033-295X.106.3.491

Hendriksz, C. J., Harmatz, P., Beck, M., Jones, S., Wood, T., Lachman, R., … Tomatsu, S. (2013). Review of clinical presentation and diagnosis of mucopolysaccharidosis IVA. *Molecular Genetics and Metabolism*. https://doi.org/10.1016/j.ymgme.2013.04.002

Heylen, E., Scherer, G., Vincent, M.-F., Marie, S., Fischer, J., & Nassogne, M.-C. (2012). Tyrosinemia Type III detected via neonatal screening: Management and outcome. *Molecular Genetics and Metabolism*, *107*(3), 605–607. https://doi.org/10.1016/j.ymgme.2012.09.002

Hillgartner, M. A., Coker, S. B., Koenig, A. E., Moore, M. E., Barnby, E., & MacGregor, G. G. (2016). Tyrosinemia type I and not treatment with NTBC causes slower learning and altered behavior in mice. *Journal of Inherited Metabolic Disease*, *39*(5), 673–682. https://doi.org/10.1007/s10545-016-9949-6

Hulme, C., & Snowling, M. J. (2016). Reading disorders and dyslexia: *Current Opinion in Pediatrics*, *28*(6), 731–735. https://doi.org/10.1097/MOP.0000000000000411

Klein, C. (2001). Developmental functions for saccadic eye movement parameters derived from pro-and antisaccade tasks. *Experimental Brain Research*, 1–17. https://doi.org/10.1007/s002210100711

Koto, A., Horwitz, A. L., Suzuki, K., Tiffany, C. W., & Suzuki, K. (1978). Morquio syndrome: Neuropathology and biochemistry. *Annals of Neurology*, *4*(1), 26–36. https://doi.org/10.1002/ana.410040106

Krauzlis, R. J., Goffart, L., & Hafed, Z. M. (2017). Neuronal control of fixation and fixational eye movements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160205. https://doi.org/10.1098/rstb.2016.0205

Marambaud, P., Dreses-Werringloer, U., & Vingtdeux, V. (2009). Calcium signaling in neurodegeneration. *Molecular Neurodegeneration*, *4*(1), 20. https://doi.org/10.1186/1750-1326-4-20

Martin, H. R., Poe, M. D., Reinhartsen, D., Pretzel, R. E., Roush, J., Rosenberg, A., … Escolar, M. L. (2008). Methods for assessing neurodevelopment in lysosomal storage diseases and related disorders: a multidisciplinary perspective. *Acta Paediatrica. Supplementum*, *97*(457), 69–75. https://doi.org/10.1111/j.1651-2227.2008.00651.x

Masurel-Paulet, a, Poggi-Bach, J., Rolland, M.-O., Bernard, O., Guffon, N., Dobbelaere, D., … Touati, G. (2008). NTBC treatment in tyrosinaemia type I: long-term outcome in French patients. *Journal of Inherited Metabolic Disease*, *31*(1), 81–7. https://doi.org/10.1007/s10545-008-0793-1

Medendorp, W. P., Buchholz, V. N., Van Der Werf, J., & Leoné, F. T. M. (2011). Parietofrontal circuits in goal-oriented behaviour: Sensorimotor integration in parietofrontal cortex. *European Journal of Neuroscience*, *33*(11), 2017–2027. https://doi.org/10.1111/j.1460-9568.2011.07701.x

Miller, B., Sheppard, A. M., & Pearlman, A. L. (1997). Developmental expression of keratan sulfate-like immunoreactivity distinguishes thalamic nuclei and cortical domains. *The Journal of Comparative Neurology*, *380*(4), 533–552. https://doi.org/10.1002/(SICI)1096-9861(19970421)380:4<533::AID-CNE9>3.0.CO;2-2

Monge, Z. A., Geib, B. R., Siciliano, R. E., Packard, L. E., Tallman, C. W., & Madden, D. J. (2017). Functional modular architecture underlying attentional control in aging. *NeuroImage*, *155*, 257–270. https://doi.org/10.1016/j.neuroimage.2017.05.002

Moran, R., Zehetleitner, M., Muller, H. J., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, *13*(8), 24–24. https://doi.org/10.1167/13.8.24

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews. Neuroscience*, *5*(3), 218–28. https://doi.org/10.1038/nrn1345

Neufeld, E. F., & Muenzer, J. (2001). The Mucopolysaccharidoses. In C. Scriver, A. Beaudel, W. Sly, & D. Valle (Eds.), *The Metabolic and Molecular Basis of Inherited Disease* (8th Edn, pp. 3421–3452). New York: McGraw-Hill.

Oh, S.-H., & Kim, M.-S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, *11*(2), 275–281. https://doi.org/10.3758/BF03196570

Pitchford, N. J., & Funnell, E. (1999). An acquired form of developmental phonological dyslexia. *Cognitive Neuropsychology*, *16*(6), 573–587. https://doi.org/10.1080/026432999380735

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, *21*(2), 353–371. https://doi.org/10.1037/0882-7974.21.2.353

Romani, C., MacDonald, A., De Felice, S., & Palermo, L. (2017). Speed of processing and executive functions in adults with phenylketonuria: Quick in finding the word, but not the ladybird. *Cognitive Neuropsychology*, 1–28. https://doi.org/10.1080/02643294.2017.1320278

Salman, M. S., Sharpe, J. a, Eizenman, M., Lillakas, L., Westall, C., To, T., … Steinbach, M. J. (2006). Saccades in children. *Vision Research*, *46*(8–9), 1432–9. https://doi.org/10.1016/j.visres.2005.06.011

Schwarz, W. (1993). A diffusion model of early visual search: Theoretical analysis and experimental results. *Psychological Research*, *55*(3), 200–207. https://doi.org/10.1007/BF00419607

Schwarz, W., & Miller, J. (2016). GSDT: An integrative model of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1654–1675. https://doi.org/10.1037/xhp0000247

Scott, C. R. (2006). The Genetic Tyrosinemias. *American Journal of Medical Genetics. Part C*, *142C*, 121–126. https://doi.org/10.1002/ajmg.c

Shapiro, E., Bjoraker, K., Guler, E., Delaney, K., Charnas, L., Tolar, J., … Thomas, K. (2009). 130. Neuropsychological outcomes in Hurler syndrome: 7–20 years post-hematopoietic stem cell transplantion. *Molecular Genetics and Metabolism*, *96*(2), S40–S41. https://doi.org/10.1016/j.ymgme.2008.11.131

Szymanska, E., Sredzinska, M., Ciara, E., Piekutowska-Abramczuk, D., Ploski, R., Rokicki, D., & Tylki-Szymanska, A. (2015). Tyrosinemia type III in an asymptomatic girl. *Molecular Genetics and Metabolism Reports*, *5*, 48–50. https://doi.org/10.1016/j.ymgmr.2015.10.004

Thimm, E., Herebian, D., Assmann, B., Klee, D., Mayatepek, E., & Spiekerkoetter, U. (2011). Increase of CSF tyrosine and impaired serotonin turnover in tyrosinemia type I. *Molecular Genetics and Metabolism*, *102*, 122–125. https://doi.org/10.1016/j.ymgme.2010.11.003

Thimm, E., Richter-Werkle, R., Kamp, G., Molke, B., Herebian, D., Klee, D., … Spiekerkoetter, U. (2012). Neurocognitive outcome in patients with hypertyrosinemia type I after long-term treatment with NTBC. *Journal of Inherited Metabolic Disease*, *35*(2), 263–8. https://doi.org/10.1007/s10545-011-9394-5

Thomas, M., Annaz, D., & Ansari, D. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language and Hearing Research*, *52*(2), 336–358.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

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

Van Ginkel, W. G., Jahja, R., Huijbregts, S. C. J., Daly, A., Macdonald, A., De Laet, C., … Van Spronsen, F. J. (2016). Neurocognitive outcome in tyrosinemia type 1 patients compared to healthy controls. *Orphanet Journal of Rare Diseases*, *11*, 1–9. https://doi.org/10.1186/s13023-016-0472-5

Verghese, P. (2001). Visual Search and Attention. *Neuron*, *31*(4), 523–535. https://doi.org/10.1016/S0896-6273(01)00392-0

Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 664–672. https://doi.org/10.1037/0096-1523.15.4.664

Wolfe, J. M. (2007). Guided Search 4.0. In W. D. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99–119). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195189193.003.0008

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, *11*(2), 269–274. https://doi.org/10.3758/BF03196569

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual Search Remains Efficient when Visual Working Memory is Full. *Psychological Science*, *12*(3), 219–224. https://doi.org/10.1111/1467-9280.00339

Wraith, J. E. (2006). Mucopolysaccharidoses and Oligosaccharidoses. In *Inborn Errors of Metabolism: Diagnosis and Treatment* (4th Edition, pp. 495–507). Springer.

Zorzi, M. (2010). The connectionist dual process (CDP) approach to modelling reading aloud. *European Journal of Cognitive Psychology*, *22*(5), 836–860. https://doi.org/10.1080/09541440903435621

Table 1: Summary of cognitive deficits across cognitive domains.

Task	MPS-IVa				Tyrosenimia III			
	Number of patients with scores worse than control mean	Number of patients with $z_{pi} > 1$	Number of patients with $z_{pi} > 2$	Evidence of control/ patient difference?	Number of patients with scores worse than control mean	Number of patients with $z_{pi} > 1$	Number of patients with $z_{pi} > 2$	Evidence of control/ patient difference?
Simple reaction time	11/12	3/12	0/12	weak	7/11	4/11	3/11	**yes**
Visual search								
Feature search								
Overall search time	11/12	7/12	6/12	**yes**	7/11	6/11	4/11	**older patients**
Effect of display size	7/12	5/12	3/12	no	1/11	1/11	0/11	no
Increasing differences across quintiles?				no				**yes**
Conjunction search								
Overall search time	10/12	7/12	3/12	**yes**	7/11	7/11	5/11	**yes**
Effect of display size	6/12	2/12	2/12	no	4/11	3/11	1/11	no
Increasing differences across quintiles?				no				**yes**
Language								
Boston Naming Test	7/11	4/11	1/11	no	9/10	8/10	6/10	**yes**
BPVS	9/11	7/11	0/11	weak	11/11	10/11	7/11	**yes**
Oculomotor tasks								
Fixation task								
Dwell time	11/11	10/11	8/11	**yes**	9/10	6/10	2/10	**yes**
Intrusive saccades	9/11	7/11	5/11	**yes**	3/11	2/11	1/11	no
Prosaccade task								
Onset time	5/9	1/9	0/9	no	5/9	2/9	1/9	no
Saccade velocity	6/9	1/9	0/9	no	1/9	0/9	0/9	no

A) MPS-IVa

B) Tyrosinemia III

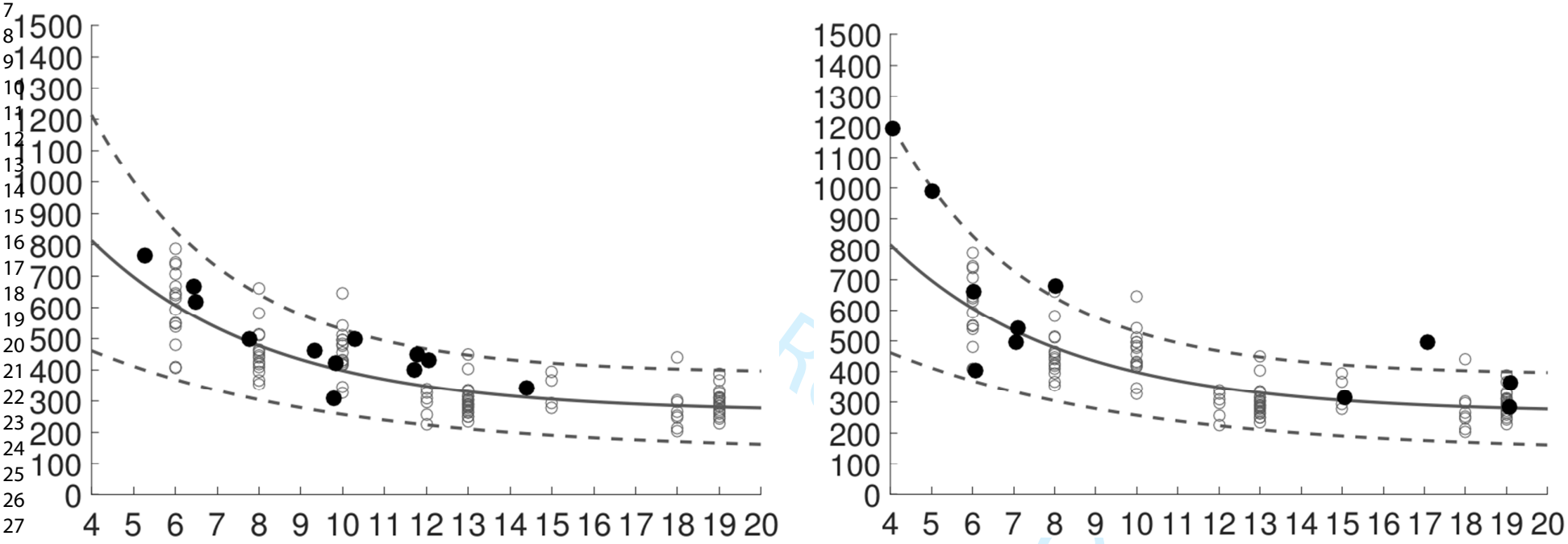


Figure 1. Mean simple reaction time for (A) MPS-IVa and (B) Tyrosinemia III individuals (filled circles) and typically developing controls (open circles). The developmental trajectory for controls is plotted as a solid line with 95% prediction intervals (dashed lines). Reaction times are collapsed across target locations.

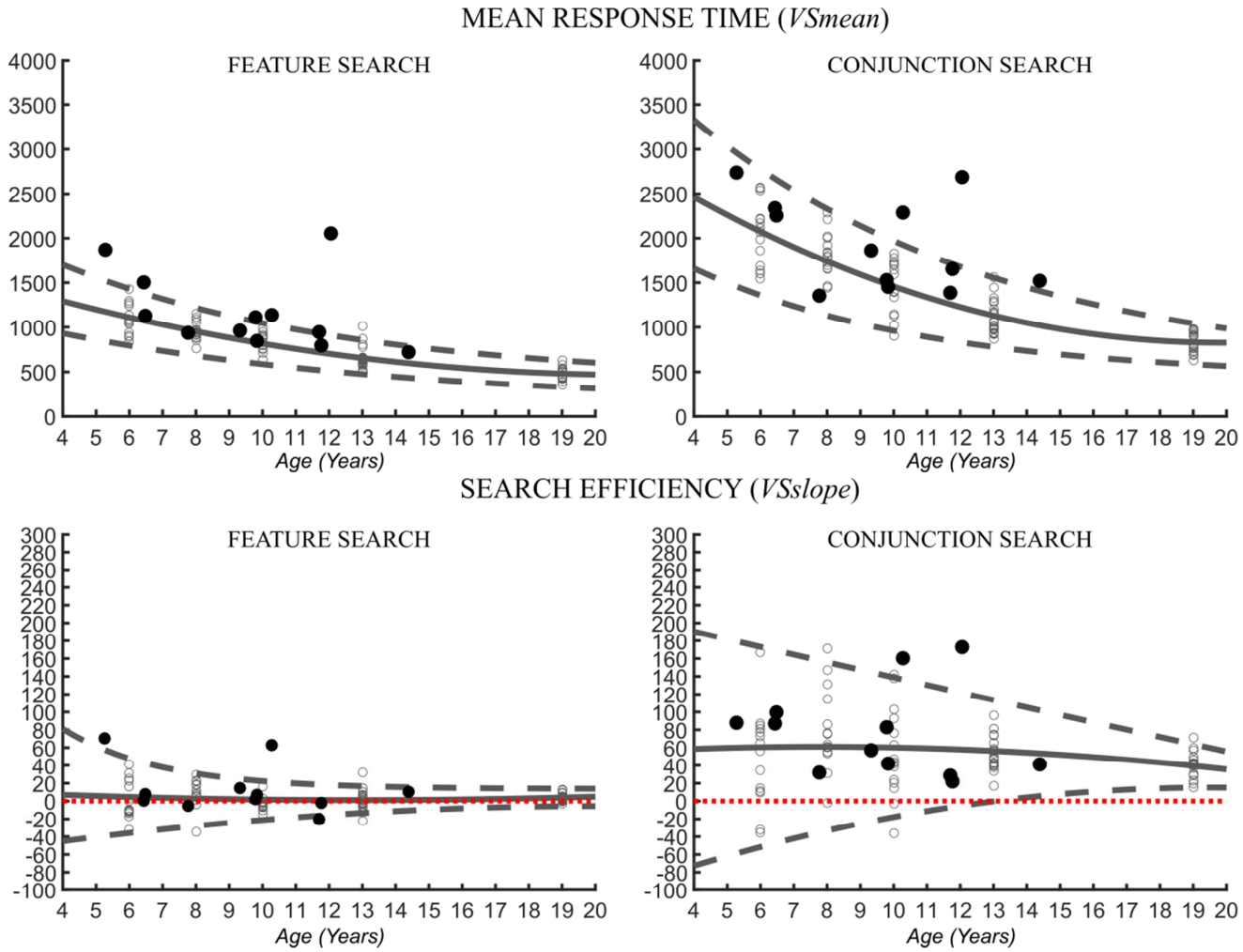


Figure 2. Overall mean RT and search efficiencies (display size slopes) in a visual search task for MPS-IVa individuals (filled circles) and controls (open circles). Feature search results are in the left panels and conjunction search on the right. The developmental trajectory for controls are is plotted as a solid line with 95% prediction intervals (dashed lines). The red dotted line in the efficiency plots represents efficient search (no increase in reaction time for larger displays).

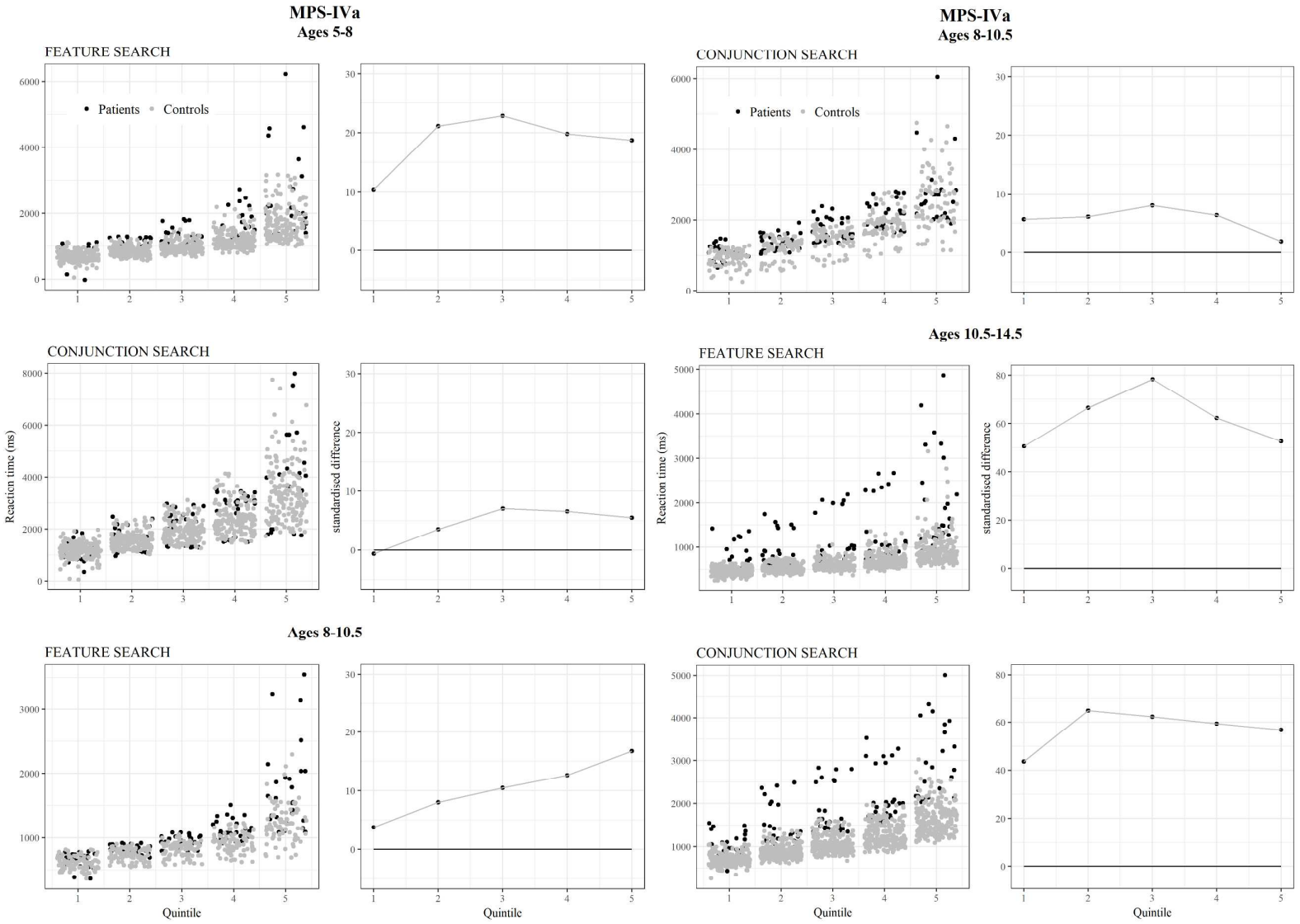


Figure 3. MPS-IVa and Control reaction times from feature and conjunction search. Reaction times are displayed in quintile bins along with the difference between quintile means for patients and controls (control value = 0). Differences between means were standardised by the standard error from the control bin. Note change in scale for final two difference plots.

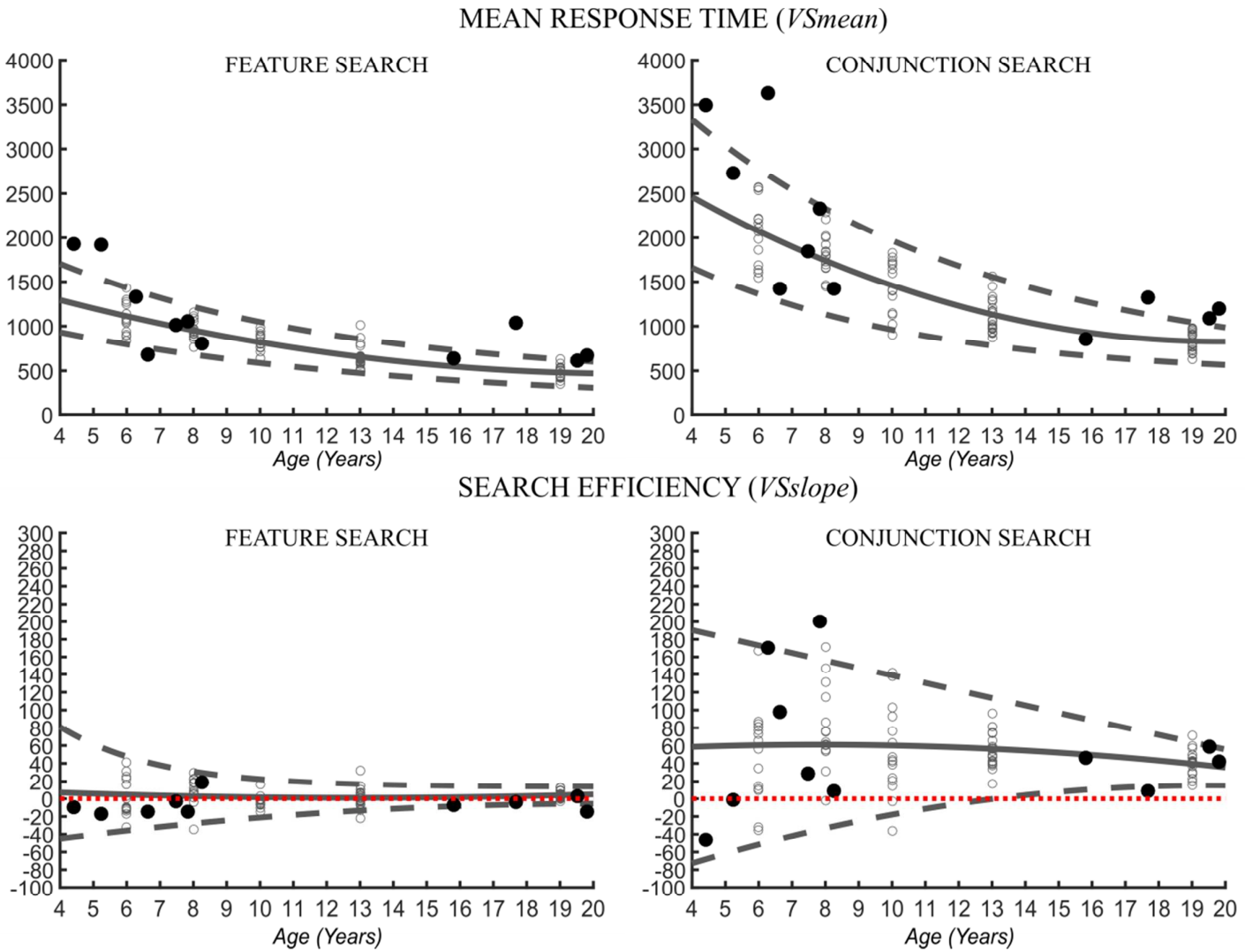


Figure 4. Overall mean RT and search efficiencies (display size slopes) in a visual search task for T3 individuals (filled circles) and controls (open circles). Feature search results are in the left panels and conjunction search on the right. The developmental trajectory for controls are is plotted as a solid line with 95% prediction intervals (dashed lines). The red dotted line in the efficiency plots represents efficient search (no increase in reaction time for larger displays).

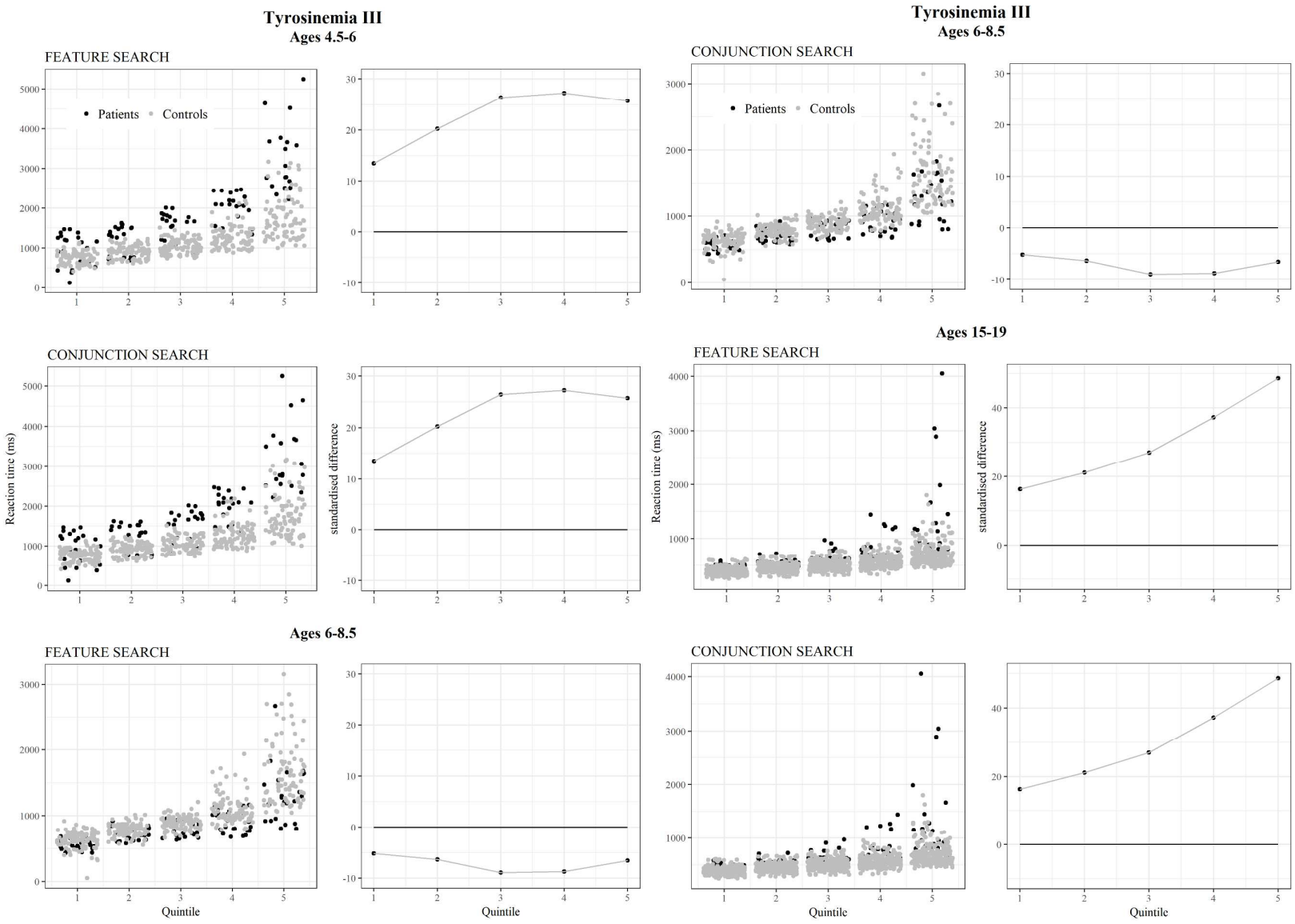
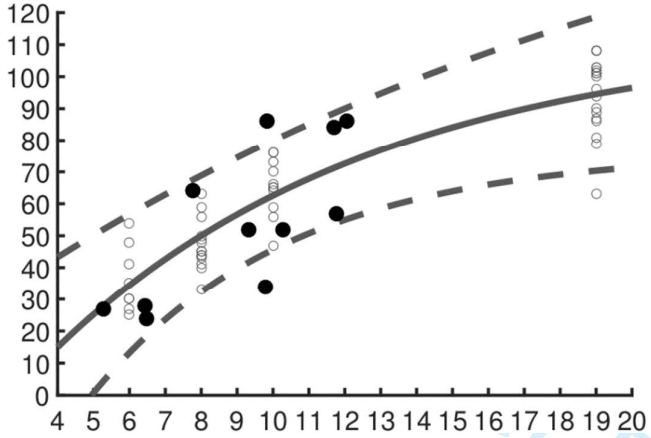


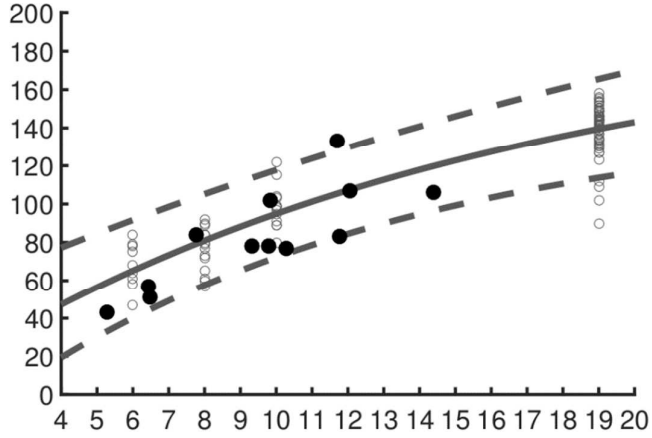
Figure 5. Tyrosinemia III and Control reaction times from feature and conjunction search. Reaction times are displayed in quintile bins along with the difference between quintile means for patients and controls (control value = 0). Differences between means were standardised by the standard error from the control bin. Note change in scale for final two difference plots.

A) MPS-IVa

BNT

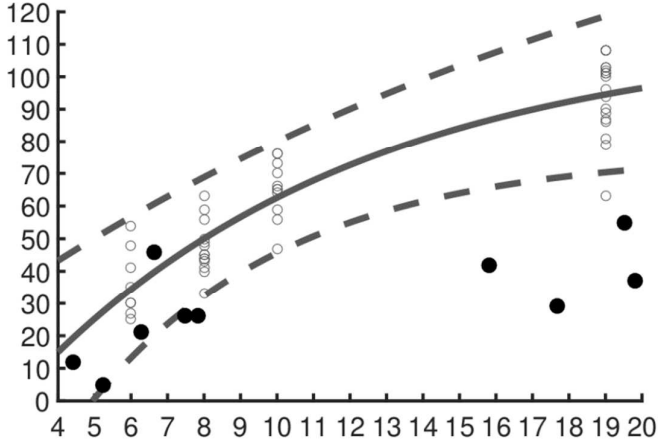


BPVS



B) Tyrosinemia III

BNT



BPVS

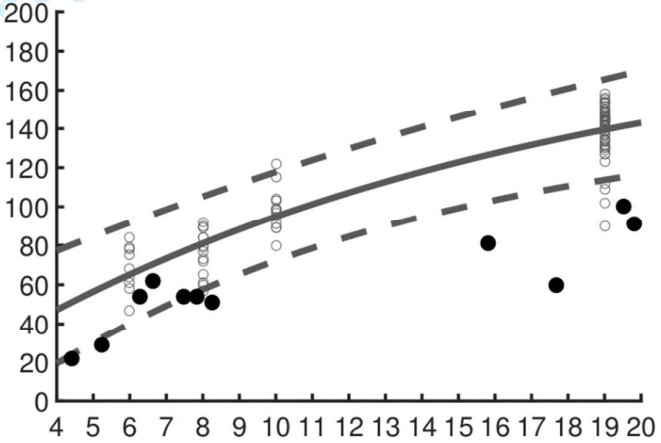


Figure 6. Boston Naming Test (BNT; production) and British Picture Vocabulary Scale (BPVS; comprehension) scores for (A) MPS-IVa and (B) Tyrosinemia III individuals and controls. Patients are filled circles and typically-developing controls are open circles. Typically developing trajectories (solid line) and 95% prediction intervals are included (dashed lines).

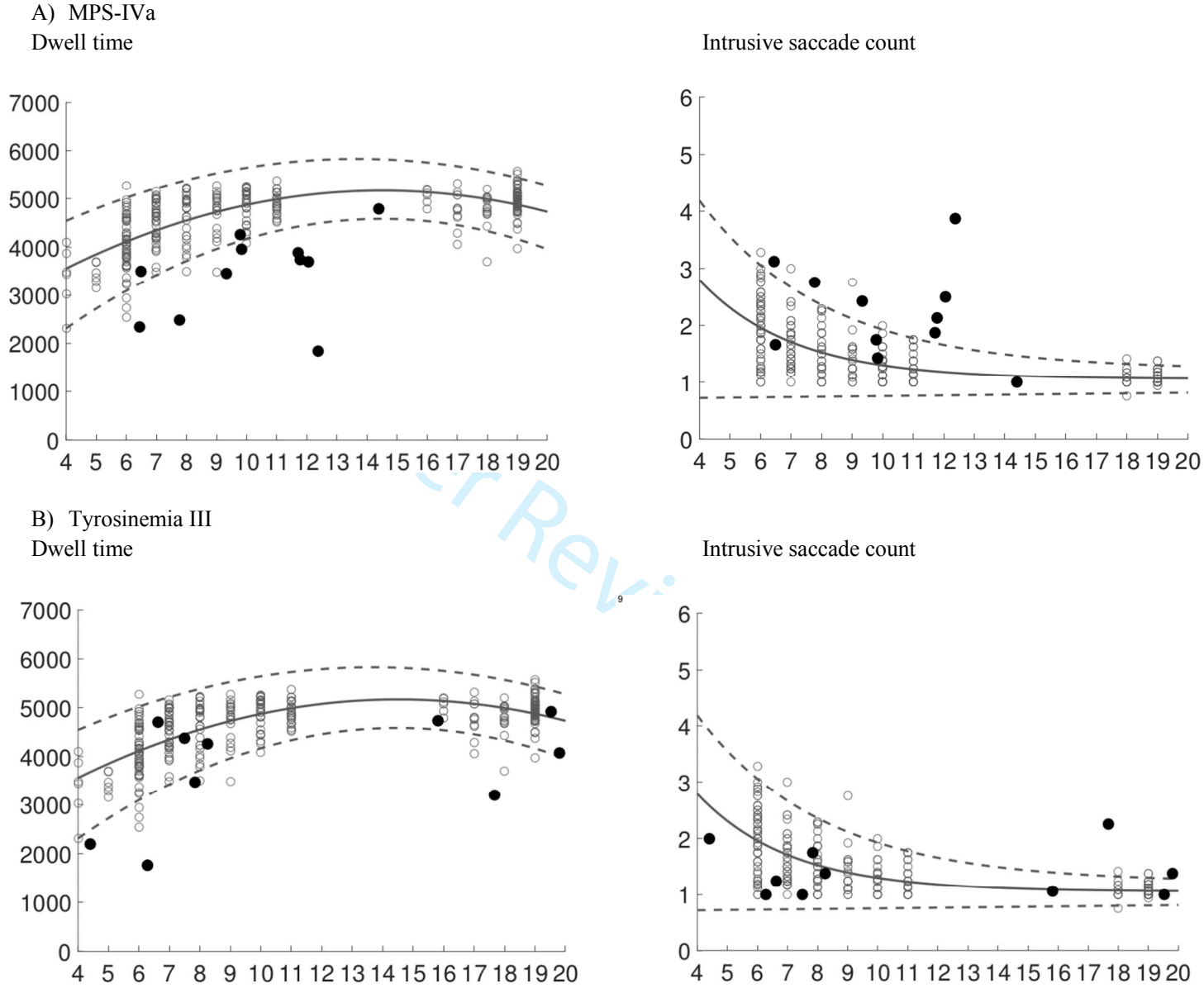


Figure 7. Fixation dwell time and intrusive saccade counts for (A) MPS-IVa and (B) Tyrosinemia III individuals (filled circles) and TD controls (open circles). Typically developing trajectories (solid line) and 95% prediction intervals (dotted lines) are included. Measures are collapsed across target locations.

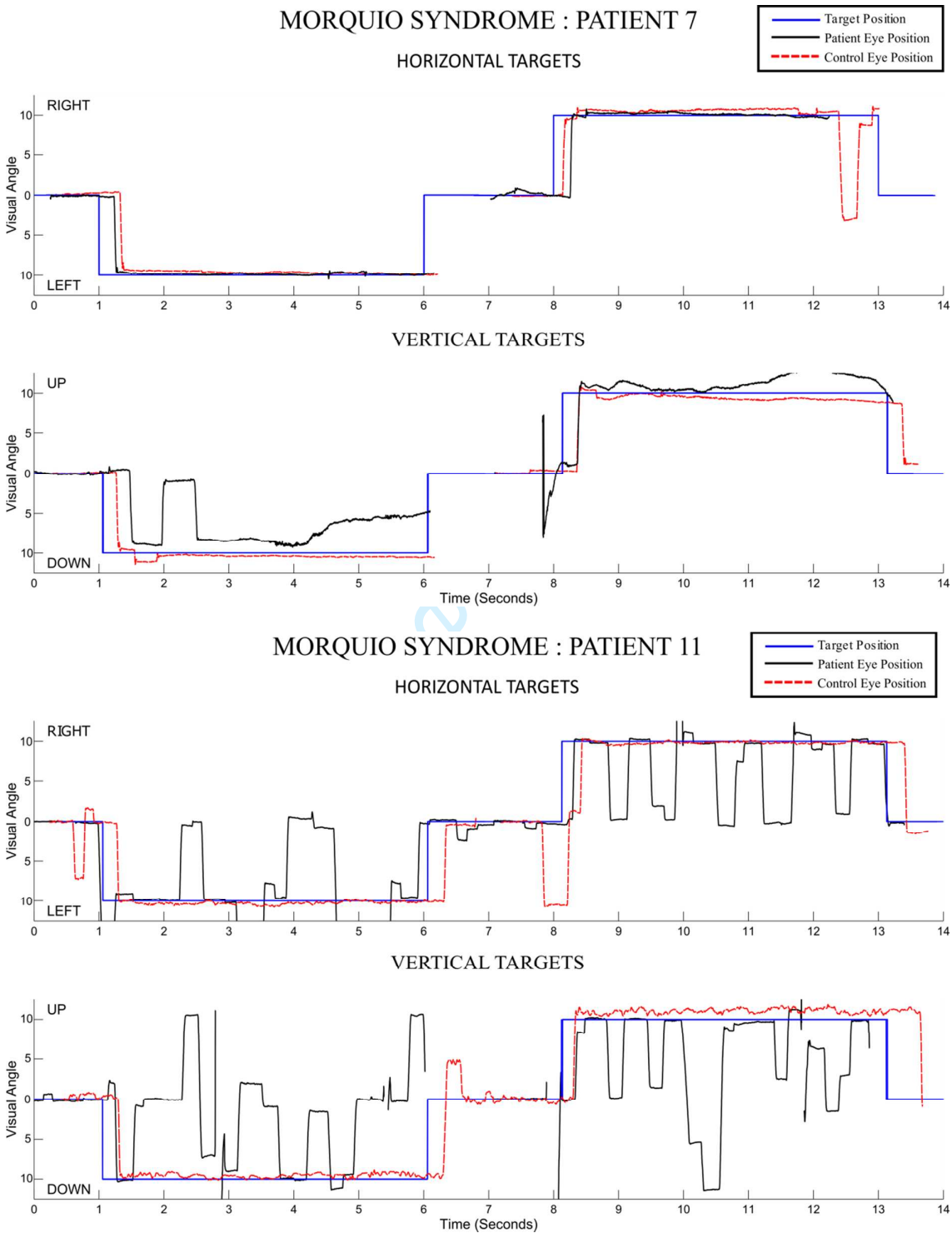


Figure 8. Example eye-movement records showing fixation time and intrusive saccades for MPS-IVa patients 7 and 11. The visual stimulus (blue line) is presented along with eye position of the patients (black line) and age-matched TD controls (red dashed line). Both patients showed fixation duration deficits buts only patient 11 displayed an elevated intrusive saccade count. Patient 7 disengages before the end of the trial.

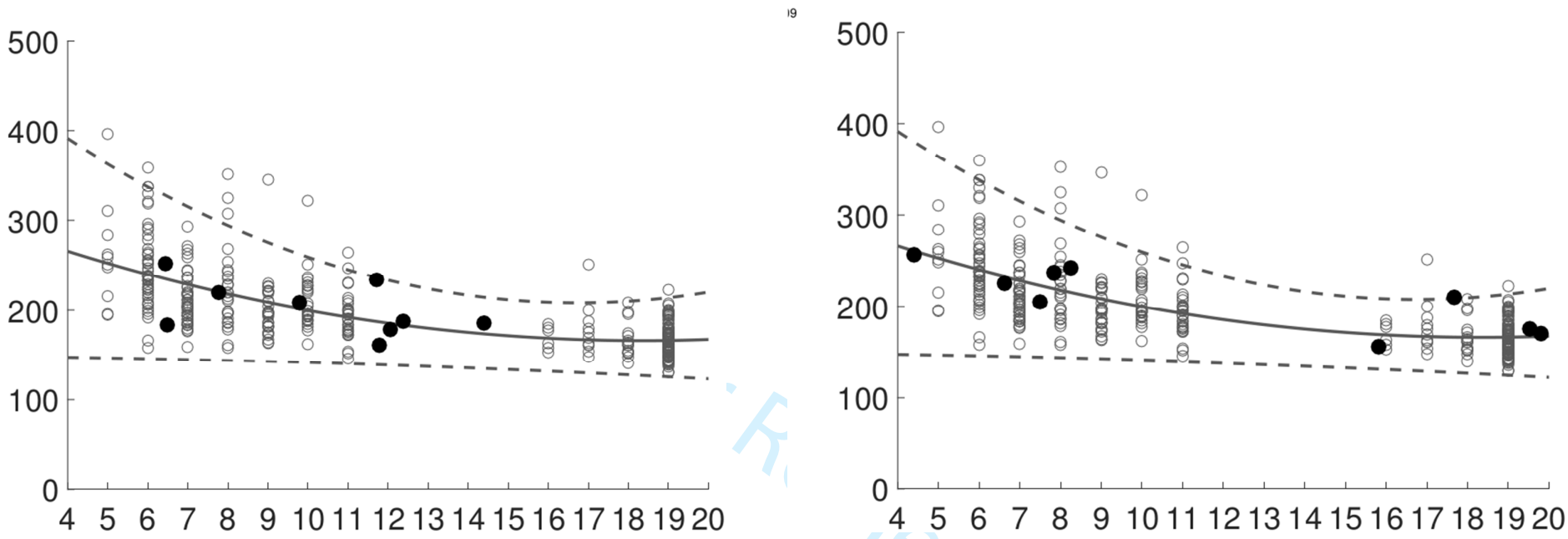


Figure 9. Average saccadic onset time for (A) MPS-IVa and (B) tyrosinemia III individuals (black dots) and TD controls (outlined grey dots). Typically developing trajectories (solid line) and 95% prediction intervals (dotted lines) are included. Onset times are collapsed across target locations.

Appendices

Appendix 1. MPS-IVa and Tyrosinemia III patient demographics

Group	PID	Gender	CA (Years)	BPVS MA (Years)	Linguistic Profile
MPS-Iva	1	M	5.27	4.02	BL – English / Pashto
MPS-Iva	2	F	6.44	5.00	BL – English / Pashto
MPS-IVa	3	F	6.48	5.07	BL – Pashto / English
MPS-IVa	4	M	7.77	8.02	BL – English / Pashto
MPS-IVa	5	F	9.33	8.07	BL – Pashto / English
MPS-IVa	6	F	9.79	10.08	ML – English
MPS-IVa	7	M	9.84	7.07	BL – Manderin / English
MPS-IVa	8	M	10.28	7.06	BL – Pashto / English
MPS-IVa	9	M	11.71	15.10	ML - English
MPS-IVa	10	M	11.78	7.10	ML - English
MPS-IVa	11	F	12.05	11.04	BL – English / Pashto
MPS-IVa	12	M	14.39	11.03	BL – English / Pashto
T3	1	F	4.42	5.32	BL – English / Pashto
T3	2	M	5.23	3.08	BL – Pashto / English
T3	3	F	6.28	5.32	BL – English / Punjabi
T3	4	M	6.62	9	ML - English
T3	5	F	7.48	5.92	BL – English / Pashto
T3	6	F	7.83	5.00	BL – Pashto / English
T3	7	M	8.25	2.92	BL – English / Pashto
T3	8	M	15.81	5.32	BL – English / Pashto
T3	9	M	17.68	7.92	BL – Pashto / English
T3	10	F	19.52	6.08	BL – English / Punjabi
T3	11	F	19.81	10.32	BL – English / Punjabi

Note: ML, Monolingual, BL, Bilingual

Appendix 2. MPS-IVa model results.

Model	AIC	ΔAIC	Akaike Weight
Simple RT			
Generating model: *Condition X Age^2 X Group*			
Random effects structure: *(1|Participant) + (1|Age)*			
Condition X Age^2 + Age^2 X Group	-458.7	0.0	0.23
Condition X Age^2 + Group	-458.7	0.0	0.22
Condition X Age^2	-458.5	0.2	0.20
Condition X Age^2 + Age^2 X Group + Condition X Group	-455.8	2.8	0.05
Condition X Age^2 + Condition X Group	-455.8	2.9	0.05
Age^2 X Group	-455.4	3.3	0.04
Feature search			
Generating model: *Display Size X Age X Group*			
Random effects structure: *(1+Display Size|Participant)*			
Age + Group	3182.2	0.0	0.24
Display Size + Age + Group	3182.5	0.3	0.20
Group X Age	3183.6	1.4	0.12
Display Size X Age + Group	3183.7	1.5	0.11
Age X Group + Display Size	3183.9	1.7	0.10
Display Size X Group + Age	3184.5	2.3	0.07
Display Size X Age + Group X Age	3185.1	2.9	0.06
Display Size X Age + Display Size X Group	3185.7	3.5	0.04
Conjunction search			
Generating model: *Display Size X Age X Group*			
Random effects structure: *(1 + Display Size | Participant)*			
Display Size X Age + Group X Age	5539.5	0.0	0.19
Display Size + Age X Group	5539.5	0.0	0.19
Display Size X Age + Group	5540.2	0.8	0.13
Display Size + Age + Group	5540.3	0.8	0.13
Display Size X Age + Group X Age + Display Size X Group	5541.3	1.8	0.08
Display Size X Age X Group	5543.1	3.6	0.03
Boston Naming Test			
Generating model: *Age X Group*			
Age	560.5	0.0	1.00
Age X Group	829.3	268.8	0.00
British Picture Vocabulary Scale			
Generating model: *Age X Group*			
Age X Group	829.3	0.0	0.53
Age + Group	830.0	0.7	0.38
Age	833.0	3.6	0.09
Group	1004.8	175.5	0.00

Fixation dwell time

Generating model: *Condition X Age^2 X Group*
Random effects structure: *(1|Participant)*

Condition X Age^2 + Age^2 X Group + Condition X Group	16713.0	0.0	0.52
Condition X Age^2 X Group	16714.1	1.2	0.29
Condition X Age^2 + Condition X Group	16715.0	2.1	0.18
Condition X Age^2 + Age^2 X Group	16721.7	8.8	0.01

Intrusive saccades

Generating model: *Condition X Age^2 X Group*
Random effects structure: *(1|Participant)*

Condition X Age^2 + Group	1734.9	0.0	0.29
Condition X Age^2 + Condition X Group	1735.1	0.3	0.26
Condition X Age^2 + Age^2 X Group	1736.0	1.1	0.17
Condition X Age^2 + Age^2 X Group + Condition X Group	1736.1	1.2	0.16
Condition X Age^2 X Group	1738.2	3.3	0.06

Saccade onset

Generating model: *Condition X Age X Group*
Random effects structure: *(1|Participant)*

Condition X Age + Condition X Group	-977.0	0.0	0.58
Condition X Age + Age X Group + Condition X Group	-975.0	2.0	0.21
Condition X Age	-972.8	4.2	0.07

Saccade velocity

Generating model: *Condition X Age X Group*
Random effects structure: *(1|Participant)*

Condition X Age	15446.6	0.0	0.63
Condition X Age + Group	15448.5	2.0	0.24
Condition X Age + Age X Group	15450.2	3.7	0.10

Appendix 3. Tyrosinemia III model results.

Model	AIC	ΔAIC	Akaike Weight
Simple RT			
Generating model: *Condition X Age^2 X Group*			
Random effects structure: *(1|Participant)*			
Condition X Age^2 X Group	-422.1	0.00	0.45
Age^2 + Group	-420.1	1.97	0.17
Age^2 + Group + Condition	-419.5	2.60	0.12
Age^2 X Group	-418.5	3.63	0.07
Feature search			
Generating model: *Display Size X Age X Group*			
Random effects structure: *(1 + Display Size|Participant)*			
Age X Group	3232.2	0.0	0.29
Group X Age + Display Size X Group	3233.1	0.9	0.18
Display Size + Group X Age	3233.2	1.1	0.17
Display Size X Age + Group X Age + Display Size X Group	3233.7	1.6	0.13
Display Size X Age + Group X Age	3233.8	1.6	0.13
Display Size X Age X Group	3235.4	3.2	0.06
Conjunction search			
Generating model: *Display Size X Age X Group*			
Random effects structure: *(1 + Display Size|Participant)*			
Display Size + Age + Group	5565.7	0.00	0.18
Display Size + Age X Group	5565.8	0.07	0.17
Display Size X Age + Group	5566.5	0.81	0.12
Display Size X Age + Group X Age	5566.6	0.87	0.11
Display Size X Group + Age	5567.2	1.47	0.09
Group X Age + Display Size X Group	5567.2	1.52	0.08
Display Size + Age	5568.0	2.29	0.06
Display Size X Age + Group X Age + Display Size X Group	5568.0	2.31	0.06
Display Size X Age X Group	5568.7	2.97	0.04
Display Size X Age	5568.8	3.08	0.04
Boston Naming Test			
Generating model: *Age X Group*			
Age X Group	532.6	0.0	0.93
Age + Group	537.8	5.2	0.07
Age	566.4	33.8	0.00
Group	593.9	61.4	0.00
British Picture Vocabulary Scale			
Generating model: *Age X Group*			
Age X Group	807.3	0.0	0.93

Age + Group	812.5	5.2	0.07
Age	876.6	69.3	0.00
Group	975.7	168.4	0.00
Fixation dwell time			
Generating model: *Condition X Age^2 X Group*			
Random effects structure: *(1|Participant)*			
Condition X Age^2 + Condition X Group	16009.9	0.0	0.34
Condition X Age^2 X Group	16010.2	0.3	0.29
Condition X Age^2 + Group	16011.2	1.3	0.18
Condition X Age^2 + Age^2 X Group + Condition X Group	16011.9	2.0	0.13
Condition X Age^2 + Age^2 X Group	16013.2	3.3	0.07
Intrusive saccades			
Generating model: *Condition X Age^2 X Group*			
Random effects structure: *(1|Participant)*			
Condition X Age^2 + Age^2 X Group	1645.5	0.0	0.29
Condition X Age^2	1645.6	0.0	0.28
Condition X Age^2 X Group	1646.9	1.4	0.15
Condition X Age^2 + Group	1647.4	1.9	0.11
Age^2 + Condition	1649.2	3.7	0.05
Saccade onset			
Generating model: *Condition X Age X Group*			
Random effects structure: *(1|Participant)*			
Condition X Age	-975.4	0.0	0.51
Condition X Age + Group	-974.1	1.3	0.26
Condition X Age + Age X Group	-972.1	3.3	0.10
Saccade velocity			
Generating model: *Condition X Age X Group*			
Random effects structure: *(1|Participant)*			
Condition X Age + Condition X Group	15455.0	0.0	0.36
Condition X Age	15455.8	0.8	0.24
Condition X Age + Group	15456.2	1.2	0.19
Condition X Age + Age X Group + Condition X Group	15457.0	2.0	0.13
Condition X Age + Age X Group	15458.2	3.2	0.07