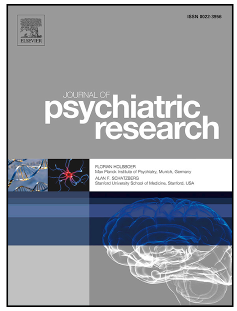
Journal Pre-proof



On the relationship between mathematics and visuospatial processing in Turner syndrome

Joseph M. Baker, Megan Klabunde, Booil Jo, Tamar Green, Allan L. Reiss

PII: S0022-3956(19)30538-2

DOI: https://doi.org/10.1016/j.jpsychires.2019.11.004

Reference: PIAT 3769

To appear in: *Journal of Psychiatric Research*

Received Date: 8 May 2019

Revised Date: 8 November 2019

Accepted Date: 12 November 2019

On the relationship between mathematics and visuospatial processing in Turner syndrome

Introduction

Advanced mathematics is a uniquely human ability that emerges as a function of basic numerical and visuospatial competencies (Dehaene, 2011; Dehaene et al., 1999). Evidence from studies of infant behavior highlight humans ability to discern numerical information from visuospatial displays within hours after birth (Izard et al., 2009). Throughout development, many of the core concepts that underlie mathematics (e.g., ordinality, cardinality) develop through repeated interactions with low-level non-verbal numerical magnitudes and visuospatial stimuli (Ansari, 2008). Within typically developing populations, associations between numeracy and visuospatial reasoning are evident in infancy, are predictive of mathematical aptitude years prior to the onset of language (Lauer and Lourenco, 2016), and persist throughout development (Geer et al., 2019; Möhring et al., 2018; Rittle-Johnson et al., 2019)

. The onset of language and other key developmental milestones ultimately assist in the development of mathematical abilities by attributing a symbolic representation (e.g., Arabic digits) to non-verbal numerical quantities (Dehaene, 2011). Despite language's role in mathematics, evidence suggests that one's ability to discern low-level numerical information (e.g., numerical magnitude) from visuospatial displays strongly predicts mathematical aptitude throughout life (Halberda et al., 2008; Libertus et al., 2012).

Disruption of basic low-level numerical abilities is associated with inhibited mathematical capability throughout development (Halberda et al., 2008; 2012; Libertus et al., 2012; 2011; Mazzocco et al., 2011). For example, students with learning disabilities specific to mathematics (e.g., dyscalculia) demonstrate significantly poorer approximate numerical discrimination performance than typically achieving students, and this relationship persists even after controlling for domain-general math abilities (Mazzocco et al., 2011). Moreover, in this sample, children's low-level numerical ability failed to differentiate low achieving from typically achieving students, suggesting that deficits to core numerical processing is relegated specifically to individuals with math learning disability. Furthermore, dyscalculia has also been associated with poor visuospatial processing skills, suggesting that deficits in visuospatial reasoning may be a domain general cause of math learning disability (Hong et al., 2009). Taken together, these results and others (Irwin, 1996; Simon et al., 2008) highlight the overlapping relationship between visuospatial perception and one's mathematical abilities (Ansari, 2008; Baker and Reiss, 2015; Dehaene, 2011; Halberda et al., 2012; Irwin, 1996; Izard et al., 2009; Libertus et al., 2011).

The specific nature of this relationship has important implications for patient populations, such as children and adults with Turner syndrome. Turner syndrome (TS) is a relatively common (1 in 1,900) genetic abnormality that occurs in females and results from the complete or partial loss of one sex chromosome (Baker and Reiss, 2015; Davenport et al., 2007). While overall IQ commonly remains intact in this population, individuals with TS are at a greater risk for difficulties with mathematics and visuospatial reasoning compared to their peers (Baker and Reiss, 2015; Hong et al., 2009; Rovet, 1993; Simon et al., 2008). Interestingly, deficits within these domains may be further isolated to individual subsystems of math and visuospatial cognition. For example, meta-analysis of all studies investigating math processing in individuals with TS highlights a pronounced performance deficit in math problems that require explicit calculation (e.g., addition, subtraction, multiplication, division) compared to non-calculation math questions (e.g., bisection, counting, digit comparison, reading numbers, transcoding numbers) (Baker and Reiss, 2015). Similarly, children with TS have been shown to underperform their peers on some tests of visuospatial reasoning (e.g., visuo-perceptual), but not others (e.g., tactile-spatial), thus supporting theories of modularity in the development of spatial skill (Temple and Carney, 1995). Furthermore, despite coincident difficulties in both cognitive domains, not all visuospatial processing deficits common within individuals with TS are related to their math difficulties, indicating that such shortfalls may arise from distinct neuropsychological deficits (Hong and Reiss, 2014; Mazzocco, 2001). For instance, visuospatial tasks that focus on the spatial location (i.e., where) aspects of a display may be more related to the math deficits seen in TS as compared to tasks that focus on the object (i.e., what) aspects of a display (Mazzocco et al., 2006). Taken together, these findings and others (Simon et al., 2008) indicate that not all visuospatial domains are alike insofar as their relationship to mathematics in TS.

The results from these studies indicate that associations between visuospatial processing and math performance varies across visuospatial subdomains and is mediated by the chromosomal abnormalities related to TS. These findings have important implications for our understanding of the relationship between mathematics and visuospatial processing in general, as well as the role that TS may have on individual elements of this interaction. Specifically, identifying the visuospatial domains that are most closely related to math performance could further clarify how these ubiquitous processes interact. Furthermore, these findings may also elucidate the neurocognitive precipitants for math deficits in individuals with TS. This, in turn, may lead to novel interventions for improving math performance in TS (Butterworth and Kovas, 2013).

For this study, we conducted a longitudinal analysis of the relationship between mathematics and visuospatial processing in age-matched adolescent girls with and without TS (45,X0 monosomy). Specifically, we hypothesized that subdomains of visuospatial processing, assessed through a battery of standardized neuropsychological evaluations, differ in their relationships to mathematics. Based on previous research, we expect such differences to emerge in visuospatial domains that target visuomotor skills and spatial orientation (Kesler et al., 2006; Mazzocco et al., 2006; Temple and Carney, 1995). Moreover, we also hypothesized that the participant's group affiliation (i.e., TS or control) will influence this relationship and thus highlight the subdomains of visuospatial processing that are uniquely impacted by TS.

Methods

All data reported herein were collected as part of a longitudinal assessment of the interaction between genes, brain development, and behavior in adolescent girls with TS and their peers. The presence of monosomic TS was established in all participants via parent of origin analysis. In order to reduce the likelihood of spurious results due to inclusion of TS participants with partial chromosomal loss (i.e., mosaicism), only individuals with classical monosomy (45,X) were included in the TS group. A total of 112 participants (N_{Turner} = 54, mean age = 11.1yrs, range = 7.0 – 15.9yrs; $N_{Control}$ = 58, mean age = 11.5yrs, range = 6.3 – 17.3yrs; see Table 1 for participant demographic breakdown across yearly visits to Stanford) were recruited for participation and underwent a battery of standardized neuropsychological assessments once a year for four years. The majority of individuals in both groups were white (TS = 88.9%, control = 66.7%), with a similar distribution of black participants (TS = 3.2%, control = 2.7%). The control group contained a greater proportion of Asian participants (TS = 0%, control = 12.7%), as well as participants who declined to state their race (TS = 0%, control = 18.2%). The inter-quartile range of household income was slightly higher in the control (IQR = $100,000 - $200,000 per year) compared to TS group (IQR = $75,000 - $150,000). No comparisons achieved statistical significance after controlling for inflated Type I errors due to multiple testing. English was the primary language spoken in the home for all participants.

All experimental activities were carried out in accordance with the latest version of the Declaration of Helsinki, and were approved by the Stanford University Institutional Review Board. Prior to participation, informed consent was obtained from the parent or caregiver of each participant, and informed assent was obtained from each participant. Trained assessors administered all tests in the standardized battery while under the direct supervision of a licensed clinical psychologist. The test battery included a broad assortment of standardized neuropsychological, cognitive, and academic achievement tests, which included but were not limited to measures of mathematics and visuospatial reasoning. In order to assess a broad range of visuospatial domains, we employed the following assessments:

- Wechsler Intelligence Scale for Children 4th Edition (WISC-IV). Five primary index scores (visual spatial index, verbal comprehension index, fluid reasoning index, working memory index, and processing speed index) yielded 3 composite scores to measure cognitive abilities relevant to assessment and identification of specific learning disabilities.
- The Wide Range Assessment of Visual Motor Abilities (WRAVMA). Assesses three domains (visual motor, visual-spatial, and fine motor) to evaluate visual-motor skills of children and adolescents.
- The developmental NEuroPSYchological Assessment (NEPSY). Six functional domains made up of 32 subtests and four delayed tasks that are designed to assess cognitive abilities related to disorders that are typically diagnosed in childhood and that are required for success in an academic environment.

A complete list and description of all subtests used for the current study may be found in Table 2 (see Table 3 for performance means and standard errors). The primary outcome variable for each model reported below was participant's performance on the Wide Range Achievement Test 4th Edition (WRAT-4) Math Computation subtest. This commonly used standardized test of math achievement consists of 40 items of graded difficulty that is administered in a 15-minute session, and included simple whole number arithmetic, problems with fractions and decimals, long division, percentages, and algebra with exponents and two unknowns. The test yielded a composite score and age-appropriate norms that yield interpretation about grade-appropriate attainment of mathematical achievement (Wilkinson et al., 2006). In order to assess the relationship between participant's performance on the WRAT-4 and each visuospatial measure[1], we employed linear mixed effects (LME) modeling. All models were constructed and evaluated in R (Team, 2017) using the 'lme4' package (Bates et al., 2015). LME is optimally suited for modeling longitudinal data, as it accommodates both multilevel fixed and random-effects to account for within-subject nested data, and does not require that time-points be matched between participants (Raudenbush and Bryk, 2002).

We started with a basic model containing a random intercept allowing for variation across *participants* in terms of their baseline, separately for the TS and Control groups. Next, we added a random slope term to allow for individual variation in terms of how they change over time. The two model specifications were then compared in each group using a likelihood ratio test to determine whether an additional random slope term is necessary. Comparison of these models

[1] Pearson's correlation revealed significant collinearity between each visuospatial predictor (see Table 4). While expected, this has implications for our models, as inclusion of multiple collinear independent variables in the same model will influence the predictive power of our visuospatial predictors. In order to eliminate inclusion of collinear predicting variables in our models, each model contained only one visuospatial predictor.

was done to establish the optimal growth model onto which *group* and each *visuospatial predictor* would be regressed. Similarly, we also compared models with and without allowing for nonlinear trends by adding a quadratic growth term. After selection of the optimal growth model, a total of 8 models were ran, each of which containing a fixed effect for group and one visuospatial predictor. The significance of each predictor and their interaction was then assessed.

Results

Prior to modeling, we visually inspected the shape and distribution of each data set. All data were largely normally distributed (see Figures 1a & 2a-h) for the TS and control groups and showed lower performance for the TS group across each measure. We then produced scatter plots of each data set that showed group-level means and standard error for each measure across each year of the study (see Figures 1b & 2i-p). Similar to the density plots collapsed across years, these plots show a clear distinction in group performance on each task. Furthermore, these plots show a similar longitudinal trajectory of scores for both groups, which is apparent in the largely parallel group lines in each plot.

Based on comparisons of models with and without random intercept, and with and without a quadratic growth term using likelihood ratio test, our results indicate that adding a random slope or quadratic term does not improve the fit. Given this finding, we consistently used random intercept models imposing a linear trend over time. Table 5 provides the predictor coefficients and significance for each model. The estimated difference in WRAT-4 performance across all testing points was significantly lower for the TS compared to the control groups within all models (see Table 5 'Group Effect'). Depending on the model, the average difference in WRAT-4 performance between the TS and control groups is estimated to vary between 14.14 and 29.32 points. The 'Time Effect' variable was not significant within any model, indicating that participant's testing session (1-4) did not account for significant variance in WRAT-4 performance.

Only four visuospatial variables (i.e., WISC IV Block Design, WISC IV Symbol Search, NEPSY Arrows, and NEPSY Picture Puzzle) significantly predicted WRAT-4 performance in our sample. Furthermore, of these visuospatial predictors, only WISC IV Block Design and NEPSY Picture Puzzle were significantly influenced by group affiliation. That is, while WISC IV Symbol Search and NEPSY Arrows significantly predicted WRAT-4 performance, they did so equally for the Turner and control groups. Conversely, the significant interaction between group and both WISC IV Block Design and NEPSY Picture Puzzle tests indicate that their association with WRAT-4 performance was significantly inter-dependent on group affiliation. These outcomes indicate that TS influences performance on the WISC IV Block Design and NEPSY Picture Puzzle tasks.

Discussion

Our results show a clear distinction between TS and control groups on math performance that is persistent over time. Specifically, girls with TS showed poorer math performance compared to their age-matched peers (see Figure 1a), and this deficit was consistent across all 4 years of our study (see Figure 1b). Notably, our random intercept model closely predicted all participants' math performance for each year of the study (see Figure 1b). These findings are important, as they indicate high levels of agreement between observed and estimated WRAT-4 performance outcomes. Furthermore, our results indicate that girls with TS demonstrate persistent deficits relative to their peers in each of the visuospatial domains that we tested. While visually apparent in Figures 1 and 2, these differences were confirmed statistically for WRAT-4 performance by significant Group effects in each model, and for each visuospatial measure via exploratory analyses not reported above[2]. Both groups are largely homogenous and normally distributed across scores on each measure (see Figure 1a & 2 a-h), suggesting that subgroups characterized by distinct neurocognitive profiles are not readily apparent within either group included in our study. However, such groups may emerge in studies containing a larger sample size, and should be further explored. Finally, performance differences across groups in our study were consistent over time, which is visually apparent in the largely parallel Group lines across years of study in Figures 1b & 2i-p.

Taken together, these results may have important implications for describing how TS impacts math and visuospatial processing throughout development. For example, our results suggest that neurocognitive deficits associated with TS emerge early in development, prior to our assessment at approximately 6 years of age. Indeed, it currently remains unknown if girls with TS are born with comparable neurocognitive abilities as their peers, which then diverge early in development, or whether such deficits are present at birth. This distinction is important, however, as it may help elucidate the neurobiological underpinnings of poor mathematics in TS. That is, the former argument may indicate that the neurobiological infrastructure that is known to support basic visuospatial and numerical processing (Dehaene, 2011) is disrupted in infants with TS, leading to atypical behavioral performance on tests of visuospatial processing and numeracy from birth. Alternatively, these abilities may be intact in infants with TS, leading to similar behavioral performance as neurotypical peers in infancy that diverged developmentally prior to our assessment. In either case, our results indicate that girls with TS follow a similar developmental trajectory as their peers throughout adolescence despite decreased overall performance, and support previous claims that impaired mathematical and visuospatial abilities in TS persist throughout development (Baker and Reiss, 2015; Hong and Reiss, 2014; Kesler et al., 2006; Mazzocco et al., 2006; Simon et al., 2008).

Closer inspection of the relationship between visuospatial processing and math performance in our sample revealed inconsistencies across visuospatial domains. For instance, only 4 of the 8 visuospatial predictors included in our study (see Table 4) were significantly related to math performance in our sample. The WISC IV Symbol Search and NEPSY Arrows tasks were equally related to WRAT-4 performance within the TS and control groups, suggesting that the visuospatial domains targeted by these tests generally relate to math performance equally across groups. Conversely, our results highlight a significant interaction between group affiliation and performance on the WISC IV Block Design and NEPSY Picture Puzzle tasks, indicating that the significant relationship between mathematics and the visuospatial domain targeted by these tasks differs significantly between adolescents with TS and their peers.

A common feature of the Block Design and Picture Puzzle tasks is their reliance on figure-ground discriminations. When presented with visual stimuli, humans use cues such as object size, color, and shape to aid in the perception of depth from two-dimensional retinal images. Ultimately, this useful ability allows us to identify individual shapes embedded within a design. In Ruben's classic example, the outline of a vase is apparent within the empty space between two profiles (Figure 3) (Hershenson, 1999); Viewers commonly report a black vase atop a white background, or two white profiles above a black background, and often switch between the two.

[2] Exploratory t-test analyses indicated significant performance differences between TS and control groups on each visuospatial measure when all assessment timepoints are combined. All exploratory analyses were corrected for inflated Type I error using the FDR correction procedure.

As a mathematical analogy, part-whole understanding refers to a person's ability to split equations into individual components. For example, '1 + 2 = X' requires an understanding that the two parts (i.e., 1 and 2) may be added together to find a missing whole (i.e., X). When posed with '1 + X = 3', the property of equality allows us to solve for 'X' by subtracting '1' from both sides of the equation. However, in order for a student to apply these and other rules, they must first be able to identify and switch between elements of the equation. We hypothesize that disrupted figure-ground perceptual abilities related to TS may negatively affect the fundamental ability to identify and parse the elements of math equations. That is, girls with TS may have difficulty discriminating the *part* and *whole* elements of an equation, leading to poor overall math performance compared to their peers. As this perceptual deficit is associated specifically with TS, this may account for the significant Group x Predictor interaction that we observed for the Block Design and Picture Puzzle tests.

Interestingly, disruption of the relationship between number and visuospatial properties in TS may also be indicative of a specific deficit in a core cognitive function related to the perception of number. Specifically, *number sense* – the ability to perceive and discriminate nonverbal quantities on the basis of number – is thought to be a shared evolutionary ability that forms the cognitive foundation for our uniquely human numerical abilities such as mathematics (Dehaene, 2011). Disruption of number sense in humans may manifest behaviorally in a manner similar to other forms of math learning disabilities such as dyscalculia (Price et al., 2007). However, given the differing underlying cause of math deficits between TS and other groups with similar behavioral phenotypes, the optimal course of treatment may differ. That is, instead of arising from deficits in domain general processes such as poor executive processing, attention, or working memory, which affect many educational domains including math, women with TS may struggle with mathematics because of a specific deficit in their ability to perceive and process numerical information from a visuospatial display. Future studies should investigate this hypothesis in greater detail.

While our results and others indicate that many aspects of visuospatial cognition are related to math performance, the degree to which other neurocognitive deficits common to TS inhibit mathematics in this group is intriguing. For instance, girls and women with TS often demonstrate deficits, relative to their peers, in tests of attention, executive function, working memory, cognitive flexibility, and abstract reasoning (Hart et al., 2006; Lepage et al., 2011). Within typically developing populations, these domain general abilities have been associated with specific subsets of mathematics. For instance, Hassinger-Das and colleagues report that attention problems were significantly related to students' calculation performance, whereas executive functioning issues were significantly related to performance on applied math problems (Hassinger-Das et al., 2014). Within the TS population, associations between math and domain general skills are not consistently reported (Hong et al., 2009), suggesting that numerical processing deficits may be more associated with numeracy-specific deficits in this group compared to their executive functioning abilities (Mazzocco and Hanich, 2010). However, a consistent finding within the TS population is a dramatic decline in math performance when a timing restriction is in effect (Baker and Reiss, 2015). These findings indicate that processing speed related to mathematics and number-related material may be uniquely affected in girls with TS, which in turn has a significantly negative influence on their math performance. Future research is needed to fully elucidate the relationship between domain general processes on mathematics in the TS population.

Our findings may have significant implications for optimal approaches to math education for girls with TS. For instance, understanding part-whole relationships is a key ability required to solve mathematical problems (Irwin, 1996), and may be a prime target of focus when teaching math to girls with TS. Specifically, extended efforts to improve the ability to identify and parse elements of an equation, especially early in math training, may help address math learning deficits that are specific to TS. These efforts may be reinforced by introducing figure-ground perceptual training into the educational plan of children with TS, with the goal of improving basic perceptual abilities prior to the introduction of complex math concepts. Furthermore, training basic numerical competencies (Park and Brannon, 2014; 2013; Thompson and Opfer, 2008; Valle-Lisboa et al., 2016) by improving visuospatial processing in collaboration with enhancement of other cognitive abilities such as executive functioning, attention, and working memory may prove particularly beneficial for the TS population. Importantly, in each instance, training may be given in the form of a computer- or tablet-based application that embeds each element in a game-like application (Halberda et al., 2012; Park and Brannon, 2014; Valle-Lisboa et al., 2016).

We experienced a large amount of participant attrition across the duration of our study (see Table 1). A large component of our longitudinal study that was not presented here is the use of structural and functional MRI to assess the neural development of the girls in our cohort. As such, participation at each time point was contingent upon participant's adherence to our MRI protocol, and multiple participants were lost to common life occurrences that restricted this adherence (e.g., braces). Other common reasons for exiting our study included moving away from the region, financial restrictions, or loss of interest. Notably, our analysis approach (i.e., linear mixed effects modeling) is robust to missing data (Raudenbush and Bryk, 2002), and thus minimized the impact of attrition on our results. However, as mentioned above, a larger or more generalizable participant sample may highlight important characteristics that were missed in our results. Moreover, as psychosocial functioning in TS is known to change over time (Dołęga et al., 2014), it may be important to interrogate the developmental relationship between social factors (e.g., psychosexual development, personal and family resources, socio-economic status, etc.), mathematics, and visuospatial reasoning.

References:1

Ansari, D., 2008. Effects of development and enculturation on number representation in the brain. Nature Reviews Neuroscience 9, 278–291. doi:10.1038/nrn2334

Baker, J.M., Reiss, A.L., 2015. A meta-analysis of math performance in Turner syndrome. Developmental Medicine & Child Neurology.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48. doi:10.18637/jss.v067.i01

Butterworth, B., Kovas, Y., 2013. Understanding neurocognitive developmental disorders can improve education for all. Science 340, 300–305.

Davenport, M.L., Hooper, S.R., Zeger, M., 2007. Turner syndrome in childhood. Neurogenetic developmental disorders: variation of manifestation in childhood 03–45.

Dehaene, S., 2011. The number sense: How the mind creates mathematics.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., Tsivkin, S., 1999. Sources of mathematical thinking: Behavioral and brain-imaging evidence. Science 284, 970–974.

Dołęga, Z., Jeż, W., Irzyniec, T., 2014. The cohort effect in studies related to differences in psychosocial functioning of women with Turner syndrome. Endokrynologia Polska 65, 287–294. doi:10.5603/EP.2014.0039

Geer, E.A., Quinn, J.M., Ganley, C.M., 2019. Relations between spatial skills and math performance in elementary school children: A longitudinal investigation. Developmental Psychology 55, 637.

Halberda, J., Ly, R., Wilmer, J.B., Naiman, D.Q., Germine, L., 2012. Number sense across the lifespan as revealed by a massive Internet-based sample. Proc Natl Acad Sci USA 109, 11116–11120.

Halberda, J., Mazzocco, M.M., Feigenson, L., 2008. Individual differences in non-verbal number acuity correlate with maths achievement. Nature 455, 665–668.

Hart, S.J., Davenport, M.L., Hooper, S.R., Belger, A., 2006. Visuospatial executive function in Turner syndrome: functional MRI and neurocognitive findings. Brain 129, 1125–1136. doi:https://doi.org/10.1093/brain/awl046

Hassinger-Das, B., Jordan, N.C., Glutting, J., Irwin, C., Dyson, N., 2014. Domain-general mediators of the relation between kindergarten number sense and first-grade mathematics achievement. J Exp Child Psychol 118, 78–92. doi:10.1016/j.jecp.2013.09.008

Hershenson, M., 1999. Visual Space Perception. MIT Press.

Hong, D., Scaletta Kent, J., Kesler, S., 2009. Cognitive profile of Turner syndrome. Developmental Disabilities Research Reviews 15, 270–278. doi:10.1002/ddrr.79

Hong, D.S., Reiss, A.L., 2014. Cognitive and neurological aspects of sex chromosome aneuploidies. The Lancet Neurology 13, 306–318. doi:10.1016/S1474-4422(13)70302-8

Irwin, K.C., 1996. Children's Understanding of the Principles of Covariation and Compensation in Part-Whole Relationships. Journal for Research in Mathematics Education 27, 25. doi:10.2307/749195

Izard, V., Sann, C., Spelke, E.S., Streri, A., 2009. Newborn infants perceive abstract numbers. Proc Natl Acad Sci USA 106, 10382–10385. doi:10.1073/pnas.0812142106

Kesler, S.R., Menon, V., Reiss, A.L., 2006. Neurofunctional Differences Associated with Arithmetic Processing in Turner Syndrome. Cereb. Cortex 16, 849–856. doi:10.1093/cercor/bhj028

Lauer, J.E., Lourenco, S.F., 2016. Spatial Processing in Infancy Predicts Both Spatial and Mathematical Aptitude in Childhood:. Psychological Science 27, 1291–1298. doi:10.1177/0956797616655977

Lepage, J.-F., Dunkin, B., Hong, D.S., Reiss, A.L., 2011. Contribution of Executive Functions to Visuospatial Difficulties in Prepubertal Girls With Turner Syndrome. DEVELOPMENTAL NEUROPSYCHOLOGY 36, 988–1002. doi:10.1080/87565641.2011.584356

Libertus, M.E., Feigenson, L., Halberda, J., 2011. Preschool acuity of the approximate number system correlates with school math ability. Developmental science 14, 1292–1300. doi:10.1111/j.1467-7687.2011.01080.x

Libertus, M.E., Odic, D., Halberda, J., 2012. Intuitive sense of number correlates with math scores on college-entrance examination. Acta Psychologica 141, 373–379. doi:10.1016/j.actpsy.2012.09.009

Mazzocco, M.M.M., 2001. Math Learning Disability and Math LD Subtypes. J Learn Disabil 34, 520–533. doi:10.1177/002221940103400605

Mazzocco, M.M.M., Feigenson, L., Halberda, J., 2011. Impaired Acuity of the Approximate Number System Underlies Mathematical Learning Disability (Dyscalculia). Child Development 82, 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x

Mazzocco, M.M.M., Hanich, L.B., 2010. Math achievement, numerical processing, and executive functions in girls with Turner syndrome: Do girls with Turner syndrome have math learning disability? Learning and Individual Differences 20, 70–81. doi:10.1016/j.lindif.2009.10.011

Mazzocco, M.M.M., Singh Bhatia, N., Lesniak-Karpiak, K., 2006. Visuospatial skills and their Association with Math performance in Girls with Fragile X or Turner Syndrome. Child Neuropsychology 12, 87–110. doi:10.1080/09297040500266951

Möhring, W., Frick, A., Newcombe, N.S., 2018. Spatial scaling, proportional thinking, and numerical understanding in 5- to 7-year-old children. Cognitive Development 45, 57–67. doi:10.1016/j.cogdev.2017.12.001

Park, J., Brannon, E.M., 2014. Improving arithmetic performance with number sense training: An investigation of underlying mechanism. Cognition 133, 188–200. doi:10.1016/j.cognition.2014.06.011

Park, J., Brannon, E.M., 2013. Training the Approximate Number System Improves Math Proficiency. Psychological Science 24, 2013–2019. doi:10.1177/0956797613482944

Price, G.R., Holloway, I., Räsänen, P., Vesterinen, M., Ansari, D., 2007. Impaired parietal magnitude processing in developmental dyscalculia. Current Biology 17, R1042–R1043. doi:10.1016/j.cub.2007.10.013

Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models. SAGE.

Rittle-Johnson, B., Zippert, E.L., Boice, K.L., 2019. The roles of patterning and spatial skills in early mathematics development. Early Childhood Research Quarterly 46, 166–178. doi:10.1016/j.ecresq.2018.03.006

Rovet, J.F., 1993. The Psychoeducational Characteristics of Children with Turner Syndrome. J Learn Disabil 26, 333–341. doi:10.1177/002221949302600506

Simon, T.J., Takarae, Y., DeBoer, T., McDonald-McGinn, D.M., Zackai, E.H., Ross, J.L., 2008. Overlapping numerical cognition impairments in children with chromosome 22q11.2 deletion or Turner syndromes. Neuropsychologia 46, 82–94. doi:10.1016/j.neuropsychologia.2007.08.016

Team, R.C., 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2017). R Foundation for Statistical Computing.

Temple, C.M., Carney, R.A., 1995. Patterns of Spatial Functioning in Turner's Syndrome. Cortex; a journal devoted to the study of the nervous system and behavior 31, 109–118. doi:10.1016/S0010-9452(13)80109-8

Thompson, C.A., Opfer, J.E., 2008. Costs and benefits of representational change: Effects of context on age and sex differences in symbolic magnitude estimation. Journal of Experimental Child Psychology 101, 20–51. doi:10.1016/j.jecp.2008.02.003

Valle-Lisboa, J., Cabana, Á., Eisinger, R., Mailhos, Á., Luzardo, M., Halberda, J., Maiche, A., 2016. Cognitive abilities that mediate SES's effect on elementary mathematics learning: The Uruguayan tablet-based intervention. Prospects 46, 301–315. doi:10.1007/s11125-017-9392-y

Wilkinson, G.S., Lutz, G.R., Assessment, F.P., 2006. Wide range achievement test (WRAT4). v-psyche.com

.

Figure 1. WRAT-4 score distributions

A. Distribution of WRAT scores for the Turner syndrome and control groups for all 4 years combined. While lower on average for the TS group, the distribution of scores were normal for both groups. B. Average WRAT scores across all four years. Each solid point represents the observed mean WRAT score for each time point, and the bars represent standard deviation. The transparent points represent our random intercept models' estimated WRAT outcome at each time point.

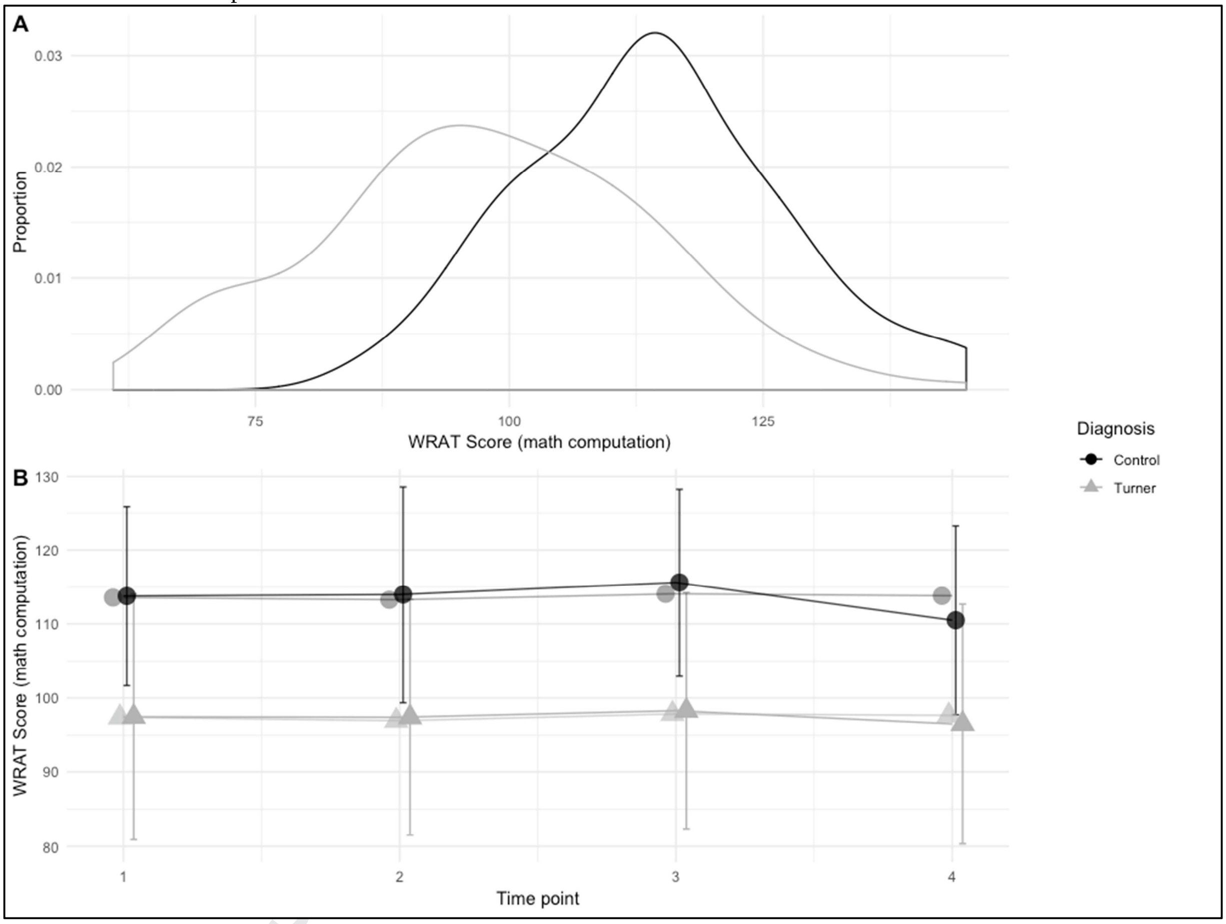


Figure 2. Visuospatial test score distributions

A-H. Density plots for each test. These plots show the distribution of scores for the Turner and control groups separately, and included data from all time points. The control group consistently outperformed the Turner group on each test, which is apparent in the right-shift along the x-axis (i.e., test score) for the blue (i.e., control) compared to red density curves. I-P. Group means and standard deviations for each test stratified across test years. The y-axis of each plot is held constant for each test. Consistently higher scores for the control group are seen here as a greater y-axis intercept for the blue compared to red lines. Furthermore, these plots show a general consistency in the distribution of scores across each year, which is apparent in the largely parallel group lines on each scatter plot. These results indicate that the developmental pattern of responding to each test is similar between members of the Turner and control groups. It is important to note that due to participant attrition the final year of study contained relatively few (n = 18, $n_{Turner} = 6$) participants.

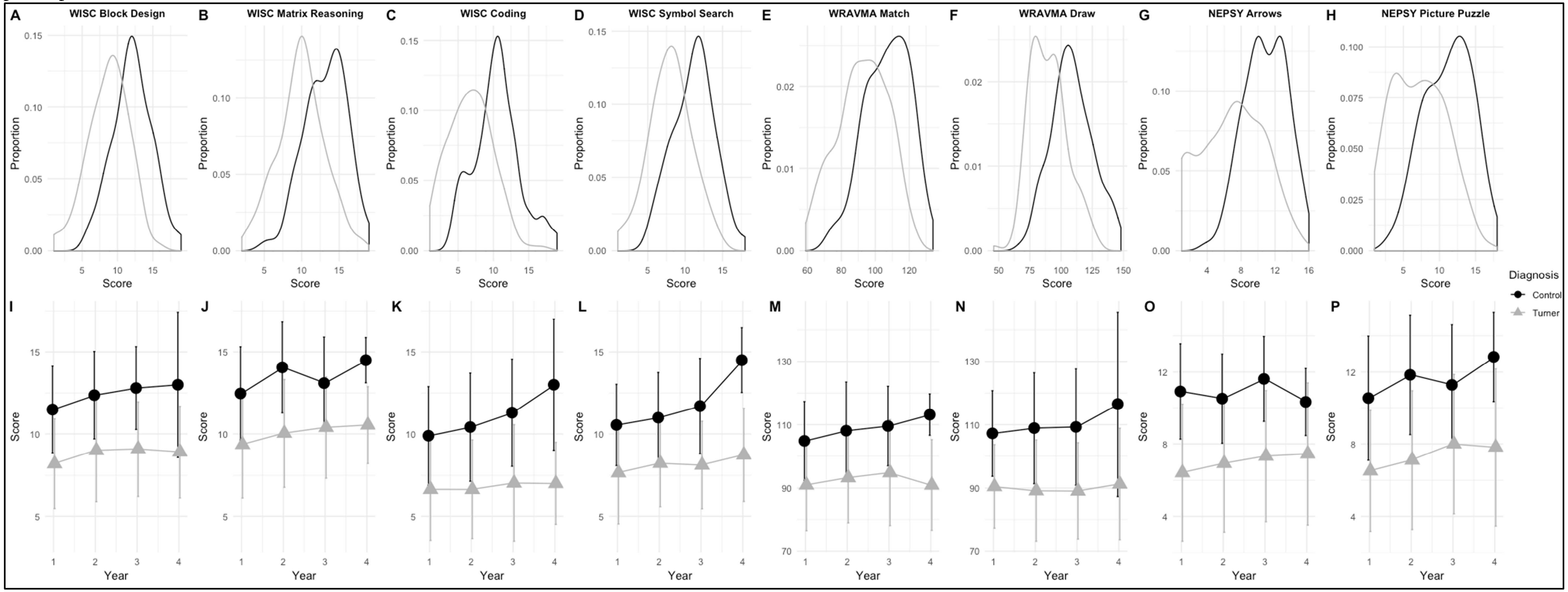


Figure 3. Rubin's vase
This visual illusion depicts the silhouette of a vase in black and the profiles of two inward-looking faces in white. The figure-ground distinction made by the brain during visual perception determines which image is seen.

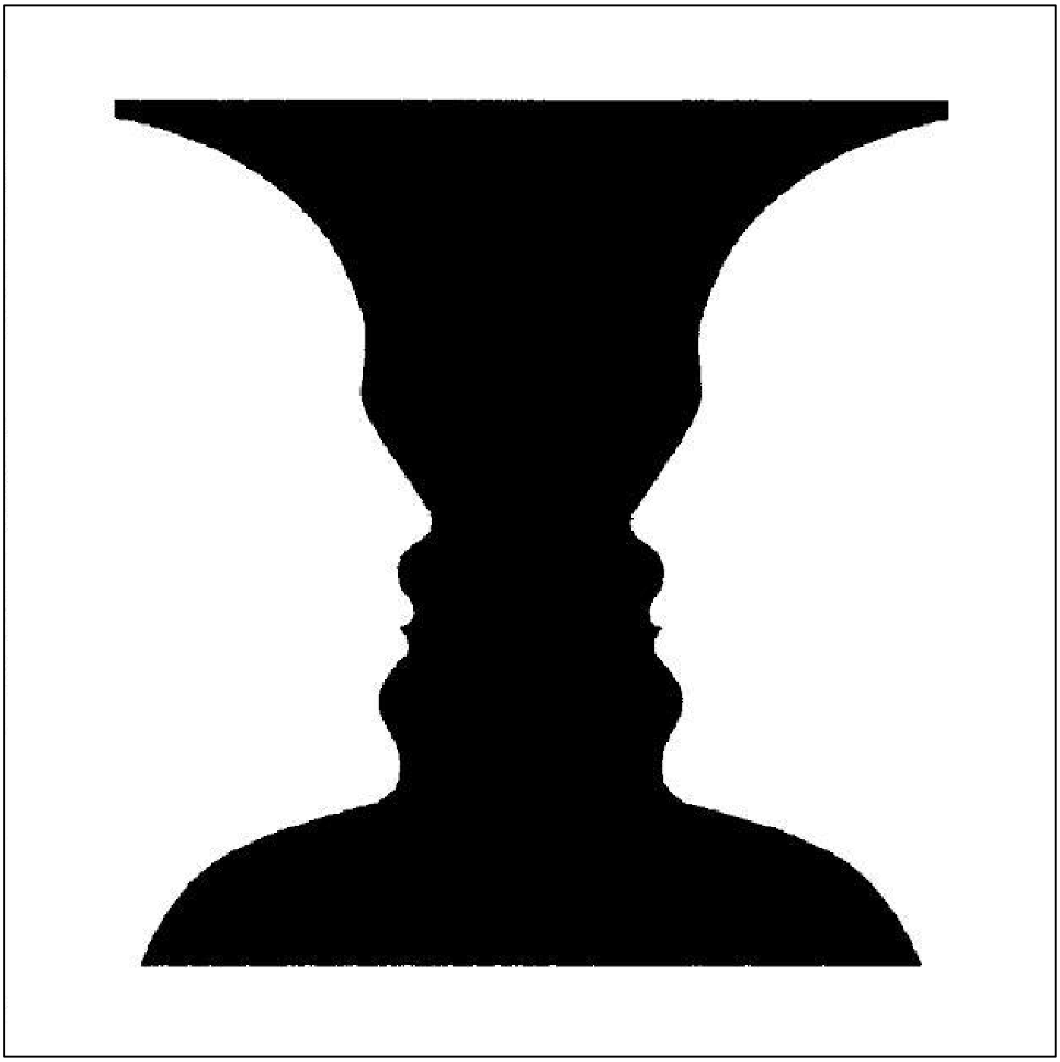


Table 1. Participant demographics across successive testing time points

Group	Visit 1			Visit 2			Visit 3			Visit 4		
	N	*Age*	*Height*	*N*	*Age*	*Height*	*N*	*Age*	*Height*	*N*	*Age*	*Height*
Turner	54	10.59 (2.38)	129.40 (11.95)	33	11.44 (2.62)	135.34 (12.41)	20	12.21 (2.38)	138.24 (11.31)	6	13.26 (2.60)	137.46 (7.55)
Control	58	10.34 (2.07)	141.44 (14.97)	45	11.42 (2.09)	147.87 (12.25)	38	12.17 (1.93)	150.85 (8.57)	12	12.81 (1.58)	156.05 (6.07)

Table 2. Standardized test battery items and description.
WRAT 4 = Wide Range Achievement Test 4, WISC IV = Wechsler Intelligence Scale for Children IV, WRAVMA = Wide Range Assessment of Visual Motor Abilities, NEPSY = Developmental Neuropsychological Assessment. Internal Consistency was measured with Cronbach Alpha and Test-Retest Stability was assessed by a Fisher's z transformation.

Tests	**Internal Consistency/ Stability Coefficients**	**Description**	**Domain**
WRAT 4 Math Computation	0.89/0.84	Measures an individuals' ability to perform basic mathematics computations through counting, identifying numbers, solving simple oral problems and calculating written mathematics problems.	Math
WISC IV Block Design	0.86/0.81	Measures nonverbal concept formation and requires perceptual organization, spatial visualization and abstract conceptualization.	Perceptual Reasoning
WISC IV Matrix Reasoning	0.89/0.85	Involves perceptual reasoning ability without a speed component. Perceptual matching, attention to detail, concentration, analogic reasoning and serial reasoning are required for successful performance.	Perceptual Reasoning
WISC IV Coding	0.85/0.81	Assesses the speed and accuracy of visual-motor coordination, speed of mental operation, attentional skills, visual acuity, visual scanning, tracking, short-term memory for new learning, cognitive flexibility, and handwriting speed. This test requires the child to copy symbols that are paired with other symbols.	Processing Speed
WISC IV Symbol Search	0.79/0.68	Involves perceptual discrimination, speed and accuracy, visual scanning attention and concentration, short-term memory, and cognitive flexibility. This test includes minor visual-motor coordination, and requires the child to look at a symbol and decide whether it is present in an array of other symbols.	Processing Speed
WRAVMA Matching	0.81/0.89	Provides a measure of spatial skill by presenting visuospatial tasks developmentally arranged in increasingly difficult order. The child marks which choice "goes best" with the item standard. The correct choice is heavily dependent on visuospatial skills such as perspective, orientation, rotation, and size discrimination.	Visuospatial
WRAVMA Drawing	0.75/0.81	Measures integrated visual-motor abilities and requires the child to use fine motor abilities to copy designs that are developmentally arranged in order of increasing difficulty.	Visuomotor
NEPSY Arrows	0.74/0.79	A non-motor subtest that measures the child's ability to judge the direction and orientation of lines. This test requires the child to choose the arrow(s) that point directly to the center of a target.	Visuospatial
NEPSY Picture Puzzle	0.89/0.83	Measures non-motor aspects of visual perception. This test requires visual integration, intact local processing, visual scanning, and an understanding of part-whole relationships of visual scenes. The child is asked to identify and match salient visual details outside of the picture with details in a picture.	Visuospatial

Table 3. Subtest performance means and standard errors. Values are aggregated across all time points

Measure	Subtest	Turner mean(se)	Control mean(se)
WRAT-4	Mathematics	97.4(15.9)	113.9(12.9)
WISC	WISC Block Design	8.7(2.8)	12.1(2.76)
	Matrix Reasoning	10.0(3.2)	13.4(2.9)
	Coding	6.8(3.1)	10.4(3.2)
	Symbol Search	8.1(2.8)	11.4(2.8)
WRAVMA	Match	92.9(15.2)	107.8(13.8)
	Draw	90.2(15.2)	109.4(17.4)
NEPSY	Arrows	7.0(3.7)	10.9(2.4)
	Picture Puzzle	7.3(3.8)	11.5(3.1)

Table 4. Correlation between visuospatial variables of interest

	WISC Block Design	WISC Matrix Reasoning	WISC IV Coding	WISC IV Symbol Search	WRAVMA Match	WRAVMA Draw	NEPSY Arrows
WISC IV Matrix Reasoning	0.578	-	-	-	-	-	-
WISC IV Coding	0.529	0.445	-	-	-	-	-
WISC IV Symbol Search	0.540	0.458	0.661	-	-	-	-
WRAVMA Match	0.598	0.649	0.436	0.463	-	-	-
WRAVMA Draw	0.538	0.425	0.495	0.430	0.471	-	-
NEPSY Arrows	0.649	0.579	0.488	0.517	0.629	0.504	-
NEPSY Picture Puzzle	0.667	0.617	0.504	0.497	0.592	0.473	0.636

All coefficients are significant ($p < 0.05$)

Table 5. LME predictor coefficients and significance

The 'group effect' provides an estimated difference in WRAT-4 performance between the TS and Control groups. This difference was significant within all models, and was estimated to vary between 14.14 and 29.32 points. WRAT-4 performance did not vary across time, and there were no significant interactions with the Time variable. The 'WISC IV Symbol Search' and 'NEPSY Arrows' measures significantly predicted WRAT-4 performance for both groups equally. The significant predictor x group interaction indicates that the predictive significance of the visuospatial predictor included in the model is influenced by group affiliation. Asterisks (*) and bold numbers indicate significance of $p < 0.05$.

	Predictor coefficients and significance						
	Group Effect	**Time Effect**	**Visuospatial Predictor**	**Predictor x Group**	**Time x Group**	**Time x Predictor**	**Time x Group x Predictor**
WISC IV Block Design	**-29.32***	-1.53	-0.03	**1.55***	2.33	0.15	-0.26
WISC IV Matrix Reasoning	**-14.14***	0.19	0.57	-0.02	-0.29	-0.01	0.01
WISC IV Coding	**-20.53***	-0.13	0.46	0.86	0.89	0.01	-0.12
WISC IV Symbol Search	**-14.58***	5.39	**1.71***	-0.91	-7.06	-0.47	0.69
WRAVMA Match	**-30.57***	-6.56	0.04	0.16	10.39	0.06	-0.10
WRAVMA Draw	**-25.22***	2.66	0.09	0.12	4.14	-0.02	-0.05
NEPSY Arrows	**-14.48***	4.00	**1.17***	0.03	-3.84	-0.32	0.29
NEPSY Picture Puzzle	**-24.55***	4.03	-0.14	**1.12***	-3.08	-0.30	0.15