From *BDNF* to Reading: Neural Activation and Phonological Processing as Multiple Mediators

Sara Mascheretti[1*], Meaghan V. Perdue[2,3*], Bei Feng[4], Chiara Andreola[1,5], Ginette Dionne[4], Kaja K. Jasińska[3,6], Kenneth R. Pugh[2,3], Elena L. Grigorenko[3,7,8,9] & Nicole Landi[2,3]

[1] Child Psychopathology Unit, Scientific Institute, IRCCS E. Medea, Bosisio Parini, LC, Italy

[2] Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA

[3] Haskins Laboratories, New Haven, CT, USA

[4] School of Psychology, Université Laval, Québec, Canada

[5] Universitè de Paris, Laboratoire de Psychologie de Développement et de l'Éducation de l'Enfant (LaPsyDÉ), Paris, France

[6] Department of Linguistics and Cognitive Science, University of Delaware, Newark, DE, USA

[7] Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX, USA

[8] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

[9] St. Petersburg State University, Russia

[*] These authors equally contributed to the manuscript.

Corresponding author:

Nicole Landi, Ph.D.

Bousfield A302

Department of Psychology

University of Connecticut

406 Babbidge Road, Unit 1020

Storrs, CT 06269-1020

nicole.landi@uconn.edu

Abstract

The *BDNF* gene is a prominent promoter of neuronal development, maturation and plasticity. Its Val^{66}Met polymorphism affects brain morphology and function within several areas and is associated with several cognitive functions and neurodevelopmental disorder susceptibility. Recently, it has been associated with reading, reading-related traits and altered neural activation in reading–related brain regions. However, it remains unknown if the intermediate phenotypes (IPs, such as brain activation and phonological skills) mediate the pathway from gene to reading or reading disability. By conducting a serial multiple mediation model in a sample of 94 children (age 5-13), our findings revealed no direct effects of genotype on reading. Instead, we found that genotype is associated with brain activation in reading-related and more domain general regions which in turn is associated with phonological processing which is associated with reading. These findings suggest that the *BDNF*-Val^{66}Met polymorphism is related to reading *via* phonological processing and functional activation. These results support brain imaging data and neurocognitive traits as viable IPs for complex behaviors.

Keywords: *BDNF*-Val^{66}Met polymorphism; Reading; Intermediate phenotypes; Serial multiple mediation model; Compensatory mechanisms.

1. Introduction

Reading is a complex task that requires the coordination of multiple cognitive and perceptual systems [1-2]. A substantial amount of research has established that individual variability in reading acquisition and reading skill is driven by neurobiological factors [3-4]. The neurocognitive organization of reading ability depends on rapidly integrating a vast circuit of brain areas over the course of reading development. This "reading circuit" is made up of neural systems that support language as well as visual and orthographic processes, working memory, attention, motor movements and higher-level comprehension and cognition [1,5-8]. After initial processing of print occurs in the visual word form area, a large left hemisphere circuit including the supramarginal gyrus (orthography to phonology mapping), the superior temporal gyrus (phonological processing), the inferior parietal lobule and the angular gyrus (lexical-semantic processing), and the inferior frontal gyrus (phonological and semantic processing, working memory), is engaged [7,9-10]. Moreover, subcortical regions implicated in long-term and working memory, procedural learning and rapid sequential auditory processing (thalamus, basal ganglia and hippocampus), have also been implicated in reading [11-13]. Given the complex structure of cognitive and perceptual brain systems involved in reading, reading skill is likely to be influenced by multiple genes, and by complex gene-environment and gene-gene interplay and interdependence [7-8,14-18]. Indeed, several genes have been linked to reading disability phenotypes specifically [7-8,14-18], and so called "generalist genes", which have been associated with cognition more broadly, are likely to impact reading ability *via* various cognitive and neurobiological processes [19-23].

One generalist gene that has recently been linked to reading skill is *BDNF* [22]. The brain-derived neurotrophic factor (*BDNF*) gene, located on chromosome 11p13, is a prominent player in neuronal development, maturation and plasticity of the central as well as the peripheral nervous systems in both the developing and adult brain [24]. The highest levels of BDNF protein expression occur in the prefrontal cortex and hippocampus, and *BDNF* has been implicated in the biology of psychiatric disorders as well as learning and memory [24]. The BDNF protein and *BDNF* gene have variable expression over early life during periods critical for language and cognitive development [25], and this expression differs by brain region [26] in a manner that is consistent with regional brain maturation [27]. Likewise, children's cognitive and linguistic abilities develop concurrently through the early grade-school years, guided by the maturation of neural sites and systems, all of which support them as they are learning to read [28].

Although several genetic variants have been identified within the *BDNF* gene, the exonic Val^{66}Met polymorphism (rs6265), which results in a valine (Val) to methionine (Met) substitution, has been the focus of a large number of genetic association studies. These studies find that this substitution

is associated with neurocognitive function and may be a risk factor for the development of neuropsychiatric disorders [29]. Specifically, the *BDNF* Met allele has been associated with impairments in memory, learning, visuospatial skills, and cognition [30]. Translating cognitive performance to brain structure and function, the Val^{66}Met polymorphism has been shown to affect morphology and function within several brain areas. In particular, volumetric reductions within the deep gray matter structures (i.e. hippocampus, amygdala, thalamus) and cortical gray matter (e.g. temporal inferior, middle and superior temporal gyri, fusiform gyrus, parahippocampal and left superior frontal gyri, frontal dorsolateral prefrontal cortex), and decreased integrity in white matter microstructure (e.g. splenium of the corpus callosum, inferior fronto-occipital fasciculus), have been observed among Met allele carriers [24]. Moreover, the Met allele has been associated with memory-related hippocampal activity [29,31-32].

With respect to research on developmental disabilities, the "generalist genes hypothesis" posits that genes such as *BDNF,* with known impacts on general cognition, may contribute significantly to specific skills such as reading [19-21]; however, only two studies to date have tested the effects of the Val66Met polymorphism upon language-related traits [22,33]. Simmons and colleagues [33] showed that subjects who are homozygous for the Met allele at the *BDNF*-rs6265 and carry susceptibility alleles within the 13q21 locus [34], have a greater risk for developing developmental language disorder. More recently, we found that Val/Val homozygotes outperformed Met allele carriers on assessments of reading comprehension and phonological memory (though not on other measures of reading or language). In the same study we found that Met allele carriers showed greater activation in reading–related regions (i.e. the bilateral fusiform gyrus, the left inferior frontal gyrus and left superior temporal gyrus) as well as in brain regions supporting domain-general cognitive processes that are important for reading (i.e. the hippocampus) during a word and pseudoword reading task. In addition, we observed that greater activation in these brain regions was correlated with better performance on a number of reading-related tasks [22]. Given that greater activation was observed for risk allele carriers, who also tended to have lower scores on some reading (and related) behaviors, we hypothesized that this activation might be compensatory in nature. More specifically, we suggested that development of reading proficiency may be facilitated by compensatory neural resources (i.e. increased activation in classic language areas and in regions supporting learning and memory more generally) in individuals who have a genetic predisposition for poorer memory performance. While these findings provided some preliminary evidence about the relationship between *BDNF* risk/non-risk allele carriers, reading, and developing brain regions that support reading, they also left some open questions. Specifically, while we hypothesized that the relationship between presence of the risk allele and reading skill was mediated by functional activation (possibly compensatory in nature), we did not explicitly test this

relationship. The current approach seeks to close this gap by determining whether this polymorphism relates to reading *via* brain activation. In addition, based on a significant amount of extant research that establishes the relationship between phonological processing and reading [6-7,35-38,39], and given the availability of substantial assessment data, we also consider phonological processing as a second intermediate phenotype (IP) mediator between genotype and reading skill.

Here we present findings from a serial multiple mediation model conducted to simultaneously test the direct and indirect effects from the Val^{66}Met polymorphism to reading skills *via* multiple IP mediators (i.e. brain activation and phonological processing) in a sample of 94 unrelated, children with typical reading ability. IPs reflect lower-level neurophysiological, biochemical, endocrinological, neuroanatomical, cognitive or neuropsychological processes [40-42], which are associated with a trait or disorder and might link specific genes to a phenotype [43-44]. Testing IPs as mediating variables has been proposed as an effective approach to unravel the complex pathways between genes and behavior [40,45-46]. Moreover, testing mediation effects of IPs is particularly salient in candidate gene studies of complex disorders, as this approach can improve our understanding of clinical heterogeneity, thus reshaping classical nosological systems, and opening new perspectives for targeted remediation treatments [45-46]. Instead of applying separate analyses of variance for testing specific main effects, running a serial multiple mediation model allowed us to formulate and test a global model for all variables based on relevant theoretical background, and to describe the structure of data in a simple, understandable and interpretable way. Based on our own previous findings and the larger literature, we hypothesized that the Met allele would be associated with decreased reading performance *via* its impact on neural activation in developing brain regions that support reading and phonological skills.

2. Materials and Methods

This study was approved by the Yale University Institutional Review Board. Written informed consent and verbal assent were obtained from parents and their participating children, respectively.

2.1 Participants

Ninety-four children between the ages of 5 and 13 (54 males, 40 females, mean age=8.4±1.3) were included in this study. With respect to ethnicity, the vast majority of the participants (86.17%) were Caucasian; of the remaining participants, 3.19% were of African-American ethnicity, 3.19% were of Hispanic ethnicity, 3.19% were of Asian ethnicity, and 4.26% were of mixed ethnicity. Eighty-one subjects have been included in a previous study investigating the *BDNF* Val^{66}Met polymorphism influences upon reading ability and patterns of neural activation [22]. The participants in this study are part of a larger longitudinal study investigating genetic links to structural and functional brain changes over a period in development corresponding to reading acquisition. Participants for this study were

included if they had an average full-scale IQ on the Wechsler Abbreviated Scale of Intelligence (standard score of 75 or above) [47], normal or corrected to normal vision and normal hearing, and reading abilities within the typical range. All children had no history of severe developmental or neuropsychological disorders. From the larger longitudinal study sample, participants who had completed the behavioral battery, fMRI task, and had provided a saliva sample were included.

2.2 Genotyping

During behavioral testing sessions with participants, we obtained biological samples using sterile Oragene™ saliva collection kits (DNA Genotek, Inc). DNA was extracted from the samples using the manufacturer's protocol. We used the Applied Biosystems Inc. (ABI) TaqMan protocol for SNP genotyping. Specifically, the Assays-on-Demand™ SNP Genotyping Product containing forward and reverse primers as well as the probe for the SNP of interest was utilized. In order to amplify the region of interest, a polymerase chain reaction (PCR) was carried out using MJ Research Tetrad Thermocycler on a 384-wellplate format. TaqMan reactions included 100 ng of genomic DNA, 2.5 µl of ABI Taqman1 Universal PCR Master Mix, 0.2 µl of ABI 40X Assays-on-Demand™ SNP Genotyping Assay Mix (assay IDC__11592758_10), 2.0 µl of sterile H2Oand 0.5 µl of Bovine Serum Albumin (BSA). The genotyping call rate was 92%; quality was controlled by regenotyping. The derived/minor allele frequency (here for the Met allele) was 0.15; the distribution of alleles did not violate Hardy-Weinberg equilibrium (p=0.077). We tested the effect of the presence/absence of the Met allele and the genotypes were classified into two-level variables, i.e. Met allele carriers (Val/Met and Met/Met; n=29, 30.9%; coded as '0'), and Val/Val homozygotes (n=65, 69.1%; coded as '1'). There were no significant differences between our two genotype groups in age (Val/Val: 8.49±1.38 *vs.* Met allele carriers: 8.11±1.18; $T_{(92)}$=1.275, p=0.205), sex (Val/Val: 36 males and 29 females *vs.* Met allele carriers: 18 males and 11 females; $\chi 2_{(1)}$ =0.367, p=0.545), or IQ (Val/Val: 113.44±15.72 *vs.* Met allele carriers: 112.14±15.87; $T_{(89)}$=0.363, p=0.717). See Supplementary Table 1.

2.3 fMRI task

The fMRI task has been described in detail elsewhere [22-23,48-51]. Briefly, it was a cue-target identity task that required a match/mismatch judgment on each trial *via* a button press. The task required participants to view pictures of common objects (e.g. a dress) while a single word or pseudoword was presented in print below the image or auditorily through MRI-compatible headphones. Participants were asked to press one button when the picture and word matched (match condition) or press a different button when the picture and word did not match (mismatch condition). The majority of trials (80%) were mismatches, and only data from mismatch trials were included in analyses so that brain responses were compared on a common "mismatch" decision. Six types of mismatch trials were presented: spoken and printed high-frequency (HF) monosyllabic real words (e.g., DREAM); spoken

and printed monosyllabic pseudowords (e.g., DREAK); printed HF monosyllabic words that are semantically related to the picture (e.g., SHIRT), and printed consonant strings (e.g., DRLST). Our baseline was a rest period during which children viewed a fixation cross. Stimulus presentation and response collection was controlled by a PC running E-prime 1.2 (Psychology Software Tools, Pittsburgh, PA, USA). In the current analysis, as in Jasinska et al. [22], we focused on activation to printed words and pseudowords only in order to isolate patterns of neural activation underlying reading, rather than lexical processing more broadly.

2.4 fMRI Acquisition

Brain images were acquired using a Siemens Sonata 1.5-Tesla MRI Scanner. Twenty axial-oblique anatomic images (TE 11 ms; TR 420 ms; FOV 20 x 20 cm; 6mm slice thickness, no gap; 256 x 256 x 1 NEX) parallel to the intercommissural line were acquired prior to functional imaging. A single-shot gradient echo, echo-planar pulse sequence (FA 80˚; TE 50 ms; TR 2000 ms; FOV 20 x 20 cm; 6mm slice thickness, no gap; 64 x 64 x 1 NEX) was used for acquisition of functional images at the twenty slice locations used for the anatomic images. Stimuli were presented at jittered interstimulus intervals of 4, 5, 6, and 7s durations, with occasional longer intervals (i.e., null trials). High-resolution anatomical images were acquired for 3D co-registration (sagittal MPRAGE acquisition, FA 8˚; TE 3.65 ms; TR 2000 ms; FOV 256 x 256 mm; 1 mm slice thickness, no gap; 256 x 256 x 1 NEX; 160 slices total). A maximum of 10 imaging runs was collected for each participant.

2.5 fMRI Data Analysis

The Analysis of Functional Neuroimages software package, AFNI (version 3.40) [52] was used for processing and statistical analysis of fMRI data [22]. The preprocessing pipeline included correction for slice acquisition time (3dTshift), motion correction (3dvolreg), and affine transformation (3dWarp) to a standardized reference space defined by the Montreal Neurological Institute (MNI) by mapping the participants' high-resolution anatomical scans to the 'Colin27' brain [53-54]. An 8 mm FWHM Gaussian filter was then applied for spatial smoothing (3dmerge). The hemodynamic response was estimated at the single subject level using a multiple regression analysis with six movement parameters treated as nuisance regressors. A generalized least squares time series fit with a restricted maximum likelihood estimation of the temporal auto-correlation structure (3dREMLfit) was used in the regression.

2.6 Regions-of-interest Selection

Regions-of-interest (ROI) were selected based on findings from a prior analysis in our lab in which a significant main effect of the *BDNF* Val^{66}Met genotype was observed in patterns of neural activation during reading [22]. These clusters were identified using a group (Val/Val versus Met allele carriers) ANCOVA in AFNI's 3dMVM program [55] that included age, gender, and IQ as covariates.

Cluster-wise correction for multiple comparisons was applied at a threshold of 0.05, corresponding to a cluster size of 309. Cluster sizes were calculated using AFNI's 3dClustSim program [22]. The six ROIs that showed a significant main effect of genotype group (Met allele carriers > Val/Val homozygotes) were: 1. bilateral precuneus extending into left inferior parietal lobule, 2. bilateral hippocampus/parahippocampal gyrus/fusiform gyrus/cerebellum, 3. left middle frontal gyrus/inferior frontal gyrus/thalamus, 4. right cingulate/middle frontal gyrus/superior frontal gyrus, 5. left cingulate/medial frontal gyrus/middle frontal gyrus/precentral gyrus, 6. right superior temporal gyrus/inferior parietal lobule/superior parietal lobule (Figure 1).

For the present study, the mean activation in each of the six ROIs was extracted for each participant using the 3dCalc tool in AFNI and entered into subsequent gene-brain-behavior mediation model.

2.7 Behavioral Assessment

Participants completed a battery of cognitive, language and reading assessments as well as educational and neuropsychological history evaluations. For the present study, we were interested in genetic and neural associations with reading and language, so we focused on the following assessments: letter-word identification, pseudoword reading ("Word Attack"), spelling, passage comprehension, oral comprehension, and picture vocabulary from the Woodcock-Johnson Test of Achievement III [56], and blending words, memory for digits, non-word repetition, and blending non-words from the Comprehensive Test of Phonological Processing (CTOPP) [57].

Descriptive statistics for the language and reading assessments are reported in Table 1. Note that missing test scores (< 10%)[a] were imputed in order to maintain the statistical power afforded by the full sample using the R package 'missForest', which employs a nonparametric random forest classification and has been effectively applied to characterize reading profiles in samples of children with missing data points [58-59]. This imputation is likely to have had little impact on the coefficient and variance component estimates and their precision in the actual data [59]. As mean bivariate correlations (r) were substantial among these assessments (r=0.481; data available upon request), we ran a principal component analysis to find the optimal weights for the variables to account for the maximum amount of variance in the dataset with the smallest number of underlying factors [60]. Using a promax rotation method, we obtained two factors with an eigenvalue >1.0, i.e. 'Reading' and 'Phonology', explaining 55.44% and 12.04% of the total variance, respectively (Kaiser-Meyer-Olkin measure of sample adequacy=0.874, Bartlett test of sphericity, X^2=733.434, df=45, p<0.001; Table 2). Standardized regression scores have been saved for each subject and entered as behavioral outcomes

[a] Letter-word identification=1.1%, Word Attack=1.1%, Spelling=1.1%, Passage comprehension=1.1%, Oral comprehension=0.0%, Picture vocabulary=0.0%, Blending words=5.3%, Memory for digits=5.3%, Non-word repetition=5.3%, Blending non-words=7.4%.

(reading) and mediators (phonology) in subsequent analyses. Table 3 shows the descriptive statistics of all study variables for the whole sample.

2.8 Statistical Analysis

Direct correlations (1) gene–brain activation ROIs, (2) gene–reading, (3) gene–phonology, (4) brain activation ROIs–reading, and (5) brain activation ROIs–phonology, were calculated using two-tailed bivariate Pearson correlations as implemented in IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp. Released, 2012).

Indirect effects were tested by using Structured Equation Modelling (SEM) as implemented in the Mplus 8.1 software package [61]. SEM simultaneously models all paths, giving more powerful, accurate and robust estimation of mediation effects than more traditional tests based on sequential regressions, especially when more than one mediator is implemented in the model. Given our a priori hypotheses, which are supported by prior findings linking phonological processing and reading ability [6-7,35-39], and linking our identified brain regions with reading ability [22], the serial multiple mediation model that specified *BDNF* Val^{66}Met –> brain activation –> Phonology –> Reading was probed (Figure 2). Given that some, but not all, ROIs could mediate the Val^{66}Met –> brain activation –> Phonology –> Reading relationship, we conducted separate multiple mediation models for each brain activation cluster. Indirect effects were examined using the 2000 bootstrap technique to assess non-normality in the product coefficient [62]. Confidence intervals (95% CIs) that do not contain zero indicated significant indirect effects [63]. This approach offers the best power, confidence interval placement, and overall control for Type I error [64]. Here, we report the full model findings. Note that, although not initially hypothesized, we also test for a bi-directional relationship between reading and phonology [39], via a serial multiple mediation model that specified BDNF Val66Met -> brain activation -> Reading -> Phonology. These results are reported in Supplementary Table 2.

Because raw scores were used for behavioral measures in our PCA, age was included in each model (Figure 2). Finally, as we used PCA factor scores for each subject, we centered and scaled the values for brain activation ROIs and age by subtracting the group mean from each value and dividing each value by the group standard deviation using the STANDARDIZE function in Mplus to convert the data to the same scale while maintaining the distribution of each variable.

3. Results

3.1 Bivariate associations between gene and brain activation ROIs, gene and reading, gene and phonology, brain activation ROIs and reading, and brain activation ROIs and phonology

3.1.1 Bivariate associations between gene and brain activation ROIs

The *BDNF* Val[66]Met polymorphism significantly correlated with brain activation in all clusters (Table 4). As Met allele carriers were coded as '0' and Val/Val homozygotes were coded as '1', brain activation in all clusters was lower for Val/Val homozygotes compared to the Met allele carriers. This is consistent with our previous report.

3.1.2 Bivariate associations between gene and reading, and gene and phonology

No significant correlations were found (Table 4).

3.1.3 Bivariate associations between brain activation ROIs and reading, and between brain activation ROIs and phonology

Brain activation in most clusters, all except for ROI 4, was significantly associated with 'Reading' (Table 4). Similarly, except for ROI 4 and ROI 6, activation in most clusters revealed a significant association with 'Phonology' (Table 4). The absence of significant correlations between ROI 4 and behavioral traits and between ROI 6 and 'Phonology', could be due to the leftward asymmetry of the "reading circuit".

3.2 Indirect effects - The serial multiple mediation model[b]

The mediation model for each brain activation cluster (ROIs 1-6) explained 54.3%, 55.3%, 55.4%, 53.8%, 53.2% and 53.9% of the variance in the Reading outcome, respectively. Using 2000 bootstrapping analyses and bias-corrected 95% CI, the significant indirect effects of X on Y *via* M1 and M2 (i.e. $a_1d_1b_2$) were the paths from the *BDNF* Val[66]Met polymorphism to Reading outcome *via* Phonology and activation in ROIs 1, 2, 3, and 4 (Table 5). Inspection of beta scores revealed that the indirect effect along this pathway is negative. Specifically, Val/Val homozygotes have lower activation relative to Met allele carriers in brain ROIs 1, 2, 3, and 4. Brain activation is positively associated with phonological processing, and phonological processing is positively related to Reading.

In addition, in all of the significant full mediation models, the indirect effect from the *BDNF* Val[66]Met polymorphism to Reading outcome *via* Phonology (i.e. a_2b_2) is significant (Table 5). Inspection of beta scores revealed that the indirect effect along this pathway is positive. That is, the Val/Val genotype is positively associated with phonological skills, which in turn are associated with reading skills, even though genotype was not directly associated with phonology (Table 4). Finally, the indirect effect from the *BDNF* Val[66]Met polymorphism to Reading outcome *via* ROIs 2 and 3 is

[b] To confirm that our results remain consistent when accounting for participants' ethnicity, we ran a follow-up analysis including only Caucasian participants (n=81). We ran a PCA in this sub-sample by using a promax rotation method, and we obtained two factors with an eigenvalue >1.0, i.e. 'Reading' and 'Phonology', explaining 55.59% and 11.89% of the total variance, respectively (Kaiser-Meyer-Olkin measure of sample adequacy=0.868, Bartlett test of sphericity, X^2=643.641, df=45, p<0.001). The PCs derived from the Caucasian sub-sample were highly correlated with the PCs derived from the total sample (Reading: r=1.000, p<0.001; Phonology: r=0.998, p<0.001). The results of the serial multiple mediation model were similar (Supplementary Table 3); the indirect effects from the *BDNF* Val[66]Met polymorphism to Reading outcome *via* Phonology and ROIs 1, 2, 3, and 4, were significant. Moreover, some additional significant mediation effects including just one of the IP have been found within the above-described full mediation models including ROIs 2 and 3. In particular, the indirect effects from the *BDNF* Val[66]Met polymorphism to Reading outcome *via* ROIs 2 and 3 (i.e. a_1b_1), and *via* Phonology (i.e. a_2b_2) remain significant (Supplementary Table 3).

significant (i.e. a_1b_1; Table 5). Inspection of beta scores revealed that the indirect effect along this pathway is negative. That is, Val/Val homozygotes have lower activation relative to the Met allele carriers in brain ROIs 2 and 3; brain activation is positively related to Reading.

4. Discussion

Informed by previous results demonstrating that the *BDNF* Val^{66}Met polymorphism is associated with reading-related skills and activation in reading-related brain regions [22], the current study simultaneously targeted a sequence of possible etiological factors from gene to reading skills in children. In particular, we examined the presence of direct effects of the Val^{66}Met polymorphism on reading, as well as the indirect pathways involving IPs (i.e. reading-related brain activation and phonological processing) as mediators of this association, by using a serial multiple mediation model. Results indicated that the Val^{66}Met polymorphism was related to reading only through the influences of neural activation in a number of reading-related and more domain general brain regions [65-66] and phonological skills. The brain regions where we observed greater neural activation in Met allele carriers relative to Val/Val homozygotes included a broad network of regions known to be important for reading in children. In particular, the fusiform gyrus is strongly associated with visual word processing; the left inferior frontal gyrus has been implicated in lexical, morphological and syntactic processing; the parietal lobe is involved in language processing and the left inferior parietal lobule is part of the temporo-parietal circuit involved in cross-modal integration; the right frontal regions are related to executive control and speech production; the hippocampal area supports (pseudo)word decoding; the cerebellum is believed to be crucial for the acquisition of fluent reading skills as it is richly connected with all the brain regions involved in reading acquisition [67-70]. Overall, the model explained about 55% of the variance in reading skills. In addition, phonology was a more proximal factor that linked reading skills to distal factor of brain activation in regions relevant to reading, and of genetic risk (i.e., the Met allele).

Contrary to our previous results in which the Met allele was associated with poorer performance on some reading and reading-related tasks (reading comprehension and phonological memory) [22], our findings did not show any direct correlation between genotype and indicators of reading performance (cf. Table 4). Although this could be due in part to statistical reasons[c], these findings strengthen the notion that the *BDNF* polymorphism is associated with reading behavior *via* intermediate factors (brain activation, phonological skills) some of which (i.e. brain activation) may represent putative compensatory resources which facilitate the development of reading proficiency in

[c] The sample of this study partially overlaps with that of the previous study as 13 subjects have been added, and behavioral tasks have been differently modelled as we ran a principal component analysis instead of considering each test independently. Moreover, as SEM simultaneously controls for all included variables, paths are residual paths indicating unique contributions above other independent variables.

genetically at-risk subjects. That is, despite the putative (and previously observed) genetic predisposition for poorer behavioral performance among Met carriers on some reading related tasks [22], this model suggests that any association is fully mediated by intermediate phenotypes. In our previous report, we suggested that the greater activation we observed for Met carriers in the ROIs considered in the current study, may have been compensatory given that these individuals tended to have lower scores on some reading measures. Evidence from functional imaging studies has revealed bilateral activation in typical readers [71-72], in left anterior regions, bihemispheric inferior frontal areas, and right posterior sites; all of which have been interpreted as compensating for the failure to develop the left posterior circuits adequately [9]. The present findings are partially consistent with this speculation, as increased activation in both reading related and more domain general ("compensatory") regions is positively associated with phonological skills, which in turn are associated with reading skills, even though genotype was not directly associated with reading or phonology. As such, it is plausible to hypothesize that risk (Met) allele carriers require greater levels of brain activation to achieve comparable levels of reading. This hypothesis is consistent with evidence from functional imaging studies which has proposed that additional recruitment of left anterior regions, bi-hemispheric inferior frontal areas, and right posterior sites, may support word reading in at-risk readers [9, 73-74].

Furthermore, our findings support the association between phonological skills and reading performance [6-7,35-39]. A large amount of evidence has now been accumulated to support the relationship between phonological awareness and reading ability across languages [6,36,38,75]. Phonological awareness is believed to be important for mapping speech sounds onto their homologous visual letters, which in turn underlies the attainment of fluent reading levels [6]. Importantly, a growing amount of data shows that variation in phonological awareness is an important predictor of reading in every language, though its influence was stronger in less consistent orthographies [7].

The present study therefore extends upon past studies by demonstrating indirect pathways linking the *BDNF*-Val^{66}Met polymorphism with reading skills *via* the effects of this genetic variant on brain activations in regions relevant to reading and phonological skills as well as in regions supporting domain-general cognitive processes. Our data support the hypothesis that the Val^{66}Met polymorphism may influence underlying brain and neurocognitive IPs (e.g. regional brain activity and phonological skills including phonological awareness and phonological working memory) that support reading. Further, the current findings also support imaging data and neurocognitive traits as viable IPs for complex neurobehavioral traits like reading as they are more tractable to genetic mapping than a primary phenotype, principally because they are presumed to be closer to the underlying biology [40-41,44-45,76]. According to the multiple deficit model underlying the liability of complex traits [77-78], the direct effect of genetic variation is limited and represents only the first step in a chain of events that

may ultimately lead to the behavioral phenotype [46]. For this reason, testing IPs as mediating variables has been proposed as an effective approach to unravel the complex pathways underlying the association between genetic and lower-level brain and neurocognitive underpinnings of behavior [40,45-46,79].

There are limitations of the current study. First, the cross-sectional nature of the study and the implemented statistical method do not allow for determination of causal influences among the measures over time. Longitudinal studies are therefore needed in order to address this issue. Second, although the sample size is smaller compared with classical molecular genetic studies, it is substantial for combined gene-brain-behavior approaches. Further, the present SEM approach which utilized Monte Carlo modelling for 1,000 samples [80], yielded better estimated post-hoc statistical power for some of the models than others (range = 0.630 for ROI 4 - 0.886 for ROI 3). This could be due to the slightly smaller effect sizes obtained with some ROIs (e.g. -0.104 for ROI 4 compared to -0.147 for ROI 3). Given that small effect sizes are characteristic of neuroimaging-genetic data, testing in larger sample sizes is desirable to detect small effects and limit Type II error. Regardless of this limitation, this finding supports the notion that using IPs for tracing effects of genetic variants on reading, is an effective alternative approach to unravel the complex pathways between a specific genetic variant and a behavioral phenotype [40,45-46] as they are more genetically tractable [44]. Moreover, using 95% CIs and resampling methods like the bootstrap for testing the mediated effects, we are able to capture 95% of the distribution, to assess non-normality in the product coefficient and to increase statistical power [63]. However, as literature on the *BDNF* Val^{66}Met polymorphism is now large and contains a number of inconsistent findings between and within academic subfields interested in the effects of this genetic variant [24], replications in independent, larger datasets are warranted.

4.1 Conclusions

This first-time investigation of the etiological sequence from the *BDNF* Val^{66}Met polymorphism to reading *via* brain activation and phonological skills contributes to the growing literature on the neurogenetic machinery of reading development. Moreover, by demonstrating potential sequential effects, whereby the Val^{66}Met polymorphism drives activity in developing brain areas that, in turn, contribute to phonological skills which are essential for reading competency, this study may open new perspectives for intervention. Specifically, one interpretation of our findings is that treatments which target deficits in specific IPs [45] are likely to be more effective for some groups of children, and that degree of response to such interventions may in part be determined by genetic factors. As such, our findings may one day be informative for identification of early profiles that presage specific treatment approaches.

Acknowledgements:

The authors thank all children who took part in this study. Sara Mascheretti was supported by the Italian Ministry of Health Grant RC2019 and RC2020. Meaghan Perdue was supported by NSF IGERT grant DGE-1144399 and NSF GRFP DGE-1747453. fMRI data collection and analysis, and collection of saliva were supported by NIH R01 HD48830 and R03 HD053409. Genotyping and genetic data processing were supported by NIH P50 HD052120. The authors declare no conflict of interest.

References

[1] Norton E, Wolf M. **Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities.** *Annual Review of Psychology*. 2012;63:427-452.

[2] Wylie J, Thomson J, Leppänen P, Ackerman R, Kanniainen L, Prieler T. Cognitive processes and digital reading. In: Verhoeven L, van den Broek P, eds. *Studies in written language and literacy - learning to read in a digital world (volume 17).* Amsterdam/Philadelphia: John Benjamins Publishing Company; 2018:57-90.

[3] Lyon G, Shaywitz S, Shaywitz B. **A definition of dyslexia**. *Annals of dyslexia*. 2003;53:1-14.

[4] Bailey S, Aboud K, Nguyen T, Cutting L. **Applying a network framework to the neurobiology of reading and dyslexia.** *Journal of neurodevelopmental disorders*. 2018;10(1):37.

[5] D'Mello A, Gabrieli J. **Cognitive neuroscience of dyslexia.** *Language, speech, and hearing services in schools*. 2018;49(4):798-809.

[6] Vellutino F, Fletcher J, Snowling M, Scanlon D. **Specific reading disability (dyslexia): What have we learned in the past four decades?**. *Journal of child psychology and psychiatry, and allied disciplines*. 2004;45(1):2-40.

[7] Peterson R, Pennington B. **Developmental dyslexia.** *Lancet (London, England)*. 2012;379(9830):1997-2007.

[8] Peterson R, Pennington B. **Developmental dyslexia.** *Annual review of clinical psychology*. 2015;11:283-307.

[9] Pugh K, Mencl W, Jenner A, et al. **Functional neuroimaging studies of reading and reading disability (developmental dyslexia).** *Mental retardation and developmental disabilities research reviews*. 2000;6(3):207-213.

[10] Norton E, Beach S, Gabrieli J. **Neurobiology of dyslexia.** *Current opinion in neurobiology*. 2015;30:73-78.

[11] Nicolson R, Fawcett A. **Developmental dyslexia, learning and the cerebellum.** *Journal of neural transmission. Supplementum.* 2005;69:19-36.

[12] Pugh K, Landi N, Preston J, et al. **The relationship between phonological and auditory processing and brain organization in beginning readers.** *Brain and language*. 2013;125(2):173-183.

[13] Ullman M, Earle F, Walenski M, Janacsek K. **The neurocognition of developmental disorders of language.** *Annual reiew of psychology*. 2020;71:389-417.

[14] Olson R. **Dyslexia: Nature and nurture.** *Dyslexia (Chichester, England)*. 2002;8(3):143-159.

[15] Scerri T, Schulte-Körne G. Genetics of developmental dyslexia. *European child & adolescent psychiatry*. 2010;19(3):179-197.

[16] Raskind W, Peter B, Richards T, Eckert M, Berninger V. **The genetics of reading disabilities: From phenotypes to candidate genes.** *Frontiers in psychology*. 2013;3:601.

[17] Becker N, Vasconcelos M, Oliveira V, et al. **Genetic and environmental risk factors for developmental dyslexia in children: Systematic review of the last decade.** *Developmental neuropsychology*. 2017;42(7-8):423-445.

[18] Mascheretti S, De Luca A, Trezzi V, et al. **Neurogenetics of developmental dyslexia: From genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms.** *Translational psychiatry*. 2017;7(1):e987.

[19] Plomin R, Kovas Y. **Generalist genes and learning disabilities.** *Psychological bulletin*. 2005;131(4):592-617.

[20] Plomin R, Kovas Y, Haworth C. **Generalist genes: Genetic links between brain, mind, and education.** *Mind, brain and education: the official journal of the International Mind, Brain, and Education Society*. 2007;1(1):11-19.

[21] Kovas Y, Plomin R. **Learning abilities and disabilities: Generalist genes, specialist environments.** *Current directions in psychological science*. 2007;16(5):284-288.

[22] Jasińska K, Molfese P, Kornilov S, et al. **The *BDNF* Val66Met polymorphism influences reading ability and patterns of neural activation in children.** *PloS one*. 2016;11(8):e0157449.

[23] Landi N, Frost S, Mencl W, et al. **The *COMT* Val/Met polymorphism is associated with reading-related skills and consistent patterns of functional neural activation.** *Developmental science*. 2013;16(1):13-23.

[24] Notaras M, Hill R, van den Buuse M. **The *BDNF* gene Val^{66}Met polymorphism as a modifier of psychiatric disorder susceptibility: Progress and controversy.** *Molecular psychiatry*. 2015;20(8):916-930.

[25] Sterner KN, Weckle A, Chugani HT, et al. **Dynamic Gene Expression in the Human Cerebral Cortex Distinguishes Children from Adults.** *PLoS ONE*. 2012;7(5): e37714.

[26] Kowiański P, Lietzau G, Czuba E, et al. ***BDNF*: A Key Factor with Multipotent Impact on Brain Signaling and Synaptic Plasticity**. *Cellular and Molecular Neurobiology*. 2018;38:579–593.

[27] Mills KL, Goddings AL, Herting MM, et al. **Structural brain development between childhood and adulthood: Convergence across four longitudinal samples.** *NeuroImage*. 2016;141:273-281.

[28] Waber DP, De Moor C, Forbes PW, et al. **The NIH MRI Study of Normal Brain Development: Performance of a Population Based Sample of Healthy Children Aged 6 to 18 Years on a Neuropsychological Battery.** *Journal of the International Neuropsychological Society*. 2007;13(5):729-246.

[29] Egan M, Kojima M, Callicott J, et al. **The *BDNF* val^{66}met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function.** *Cell*. 2003;112(2):257-269.

[30] Dincheva I, Glatt C, Lee F. **Impact of the *BDNF* Val^{66}Met polymorphism on cognition: Implications for behavioral genetics.** *The Neuroscientist: A revew journal bringing neurobiology, neurology and psychiatry*. 2012;18(5):439-451.

[31] Hariri A, Goldberg T, Mattay V, et al. **Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance.** *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2003;23(17):6690-6694.

[32] Hashimoto R, Moriguchi Y, Yamashita F, et al. **Dose-dependent effect of the Val66Met polymorphism of the brain-derived neurotrophic factor gene on memory-related hippocampal activity.** *Neuroscience research*. 2008;61(4):360-367.

[33] Simmons T, Flax J, Azaro M, et al. **Increasing genotype-phenotype model determinism: Application to bivariate reading/language traits and epistatic interactions in language-impaired families.** *Human heredity*. 2010;70(4):232-244.

[34] Bartlett C, Flax J, Logue M, et al. **A major susceptibility locus for specific language impairment is located on 13q21.** *American Journal of Human Genetics*. 2002;71(1):45-55.

[35] Snowling M. **From language to reading and dyslexia.** *Dyslexia (Chichester, England)*. 2001;7(1):36-46.

[36] Goswami U. **Why theories about developmental dyslexia require developmental designs.** *Trends in cognitive sciences*. 2003;7(12):534-540.

[37] Shaywitz B, Shaywitz S, Blachman B, et al. **Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention.** *Biological psychiatry*. 2004;55(9):926-933.

[38] Gabrieli J. **Dyslexia: A new synergy between education and cognitive neuroscience.** *Science (New York, N.Y.)*. 2009;325(5938):280-283.

[39] **Melby-Lervåg M, Lyster SAH, Hulme C. Phonological Skills and Their Role in Learning to Read: A Meta-Analytic Review. *Psychological Bulletin*. 2012;138(2):322-52.**

[40] Gottesman I, Gould T. **The endophenotype concept in psychiatry: Etymology and strategic intentions.** *The American Journal of Psychiatry*. 2003;106(4):636-645.

[41] Flint J, Munafò M. **The endophenotype concept in psychiatric genetics.** *Psychological medicine*. 2007;37(2):163-180.

[42] Braff D. **The importance of endophenotypes in schizophrenia research.** *Schizophrenia research*. 2015;163(1-3):1-8.

[43] Szatmari P, Maziade M, Zwaigenbaum L, et al. **Informative phenotypes for genetic studies of psychiatric disorders.** *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*. 2007;144B(5):581-588.

[44] Flint J, Timpson N, Munafò M. **Assessing the utility of intermediate phenotypes for genetic mapping of psychiatric disease.** *Trends in neurosciences*. 2014;37(12):733-741.

[45] Kendler K, Neale M. **Endophenotype: A conceptual analysis.** *Molecular psychiatry*. 2010;15(8):789-797.

[46] Munafò M. **Candidate gene studies in the 21st century: Meta-analysis, mediation, moderation.** *Genes, brain, and behavior*. 2006;5(Suppl 1):3-8.

[47] Wechsler D. *Wechsler abbreviated scale of intelligence, second edition (WASI-II).* San Antonio, TX: NCS Pearson; 2011.

[48] Frost S, Landi N, Mencl W, et al. **Phonological awareness predicts activation patterns for print and speech.** *Annals of dyslexia*. 2009;59(1):78-97.

[49] Preston J, Molfese P, Frost S, et al. **Print-speech convergence predicts future reading outcomes in early readers.** *Psychological science*. 2016;27(1):75-84.

[50] Malins J, Pugh K, Buis B, et al. **Individual differences in reading skill are related to trial-by-trial neural activation variability in the reading network.** *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2018;38(12):2981-29989.

[51] Perdue M, Mascheretti S, Kornilov S, et al. **Common variation within the SETBP1 gene is associated with reading-related skills and patterns of functional neural activation.** *Neuropsychologia*. 2019;130:44-51.

[52] Cox R. **AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages.** *Computers and biomedical research, a biomedical journal*. 1996;29(3):162-173.

[53] Ashburner J. **A fast diffeomorphic image registration algorithm.** *NeuroImage*. 2007;38(1):95-113.

[54] Holmes C, Hoge R, Collins L, Woods R, Toga A, Evans A. **Enhancement of MR images using registration for signal averaging.** *Journal of computer assisted tomography*. 1998;22(2):324-344.

[55] Chen G, Saad Z, Britton J, Pine D, Cox R. **Linear mixed-effects modeling approach to FMRI group analysis.** *NeuroImage*. 2013;73:176-190.

[56] Woodcock R, McGrew K, Mather N. *Woodcock-johnson III.* Itasca, IL: Riverside Publishing; 2001.

[57] Wagner R, Torgesen J, Naglieri J, Goldstein S. *Using the comprehensive test of phonological processing (CTOPP) to assess reading-related phonological processes.* Hoboken, NJ: John Wiley & Sons Inc; 2009.

[58] Stekhoven D, Bühlmann P. **MissForest--non-parametric missing value imputation for mixed-type data.** *Bioinformatics (Oxford, England)*. 2012;28(1):112-118.

[59] Eckert M, Vaden KJ, Gebregziabher M, Dyslexia Data Consortium. **Reading profiles in multi-site data with missingness.** *Frontiers in psychology*. 2018;9:644.

[60] Norman G, Streiner D. Principal components and factor analysis. In: *Biostatistics - The bare essentials. Third Edition ed.* Hamilton, Ontario: Norman, GR; Streiner, DL.; 2008:194-209.

[61] Muthén L, Muthén B. *Mplus user's guide (8th edition).* Los Angeles, CA: Muthén, LK; Muthén, BO.; 2017.

[62] Fritz M, Mackinnon D. **Required sample size to detect the mediated effect.** *Psychological science*. 2007;18(3):233-239.

[63] Tofighi D, MacKinnon D. **RMediation: An R package for mediation analysis confidence intervals.** *Behavior research methods*. 2011;43(3):692-700.

[64] Crowley S, Fan X. **Structural equation modeling: Basic concepts and applications in personality assessment research.** *Journal of personality assessment*. 1997;68(3):508-531.

[65] Saygin ZM, Osher DE, Norton ES, et al. **Connectivity precedes function in the development of the visual word form area.** *Nature Neuroscience.* 2016;19(9):1250-1255.

[66] Chyl K, Kossowski B, Dębska A, et al. **Reading acquisition in children: Developmental processes and Dyslexia-specific effects.** *Journal of the American Academy of Child and Adolescent Psychiatry*. 2019;58(10):948-960.

[67] Stoodley CJ, Stein JF. **Cerebellar Function in Developmental Dyslexia.** *Cerebellum.* 2013;12(2):267-276.

[68] Ramus F, Altarelli I, Jednoróg K, et al. **Neuroanatomy of Developmental Dyslexia: Pitfalls and Promise.** *Neuroscience and biobehavioral reviews.* 2018;84:434-452.

[69] Xia Z, Zhang L, Hoeft F, et al. **Neural Correlates of Oral Word Reading, Silent Reading Comprehension, and Cognitive Subcomponents.** *International journal of behavioral development.* 2018;42(3):342-356.

[70] Bourguignon NJ. **A Rostro-Caudal Axis for Language in the Frontal Lobe: The Role of Executive Control in Speech Production.** *Neuroscience and biobehavioral reviews.* 2014;47:431-44.

[71] Martin A, Schurz M, Kronbichler M, et al. **Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies.** *Human Brain Mapping.* 2015;36:1963–1981.

[72] Chyl K, Kossowski B, Dębska A, et al. **Prereader to Beginning Reader: Changes Induced by Reading Acquisition in Print and Speech Brain Networks.** *Journal of child psychology and psychiatry, and allied disciplines.* 2018;59(1):76-87.

[73] Shaywitz S, Shaywitz B, Fulbright R, et al. **Neural systems for compensation and persistence: Young adult outcome of childhood reading disability.** *Biological psychiatry*. 2003;54(1):25-33.

[74] Hancock R, Richlan F, Hoeft F. **Possible roles for fronto-striatal circuits in reading disorder.** *Neuroscience and biobehavioral reviews*. 2017;72:243-260.

[75] Castles A, Coltheart M. **Is there a causal link from phonological awareness to success in learning to read?**. *Cognition*. 2004;91(1):77-111.

[76] Black J, Myers C, Hoeft F. **The utility of neuroimaging studies for informing educational practice and policy in reading disorders.** *New directions for child and adolescent development*. 2015;2015(147):49-56.

[77] Pennington B. **From single to multiple deficit models of developmental disorders.** *Cognition*. 2006;101(2):385-413.

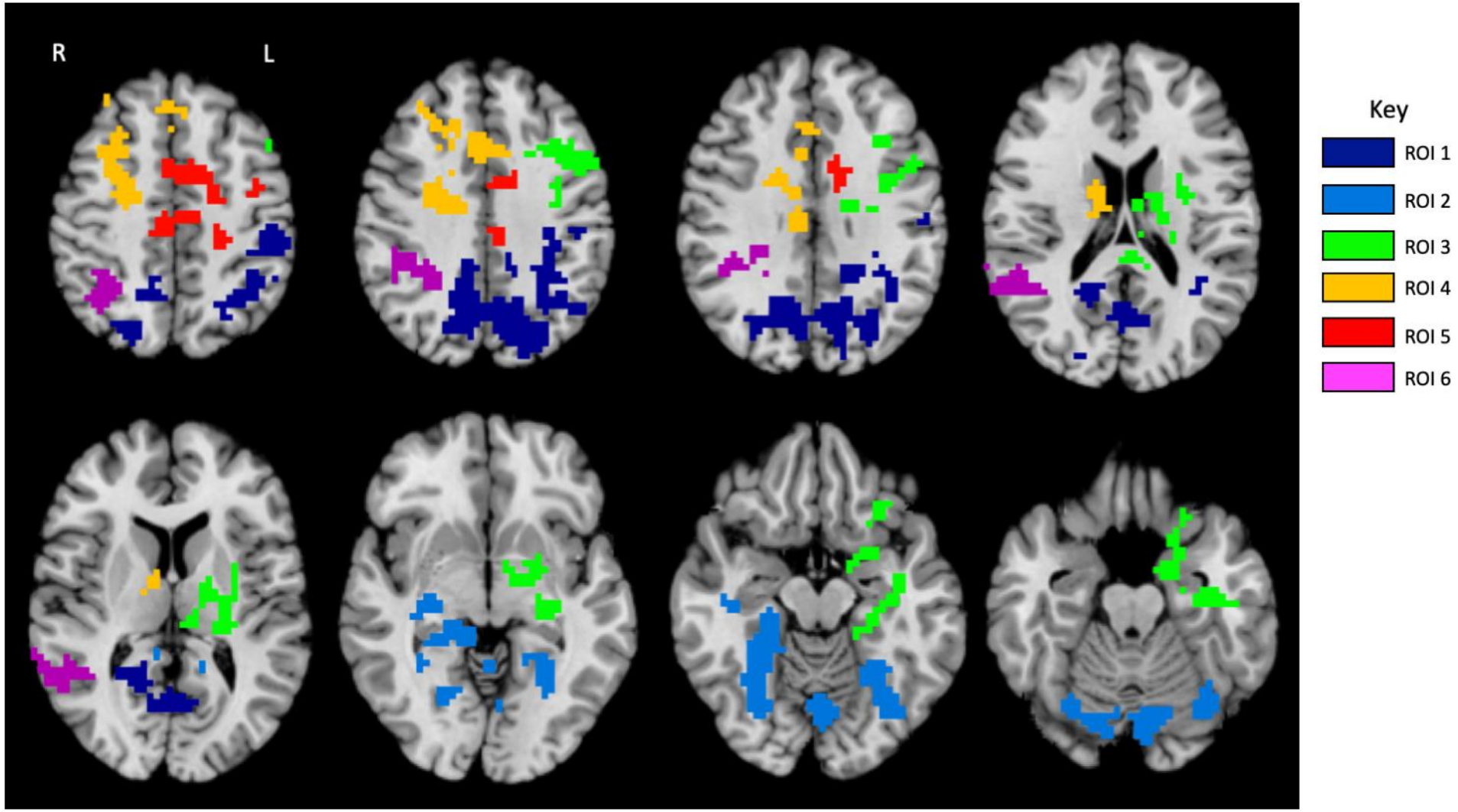
[78] McGrath L, Peterson R, Pennington B. **The multiple deficit model: Progress, problems, and prospects.** *Scientific Studies of Reading*. 2020;24(1):7-13.

[79] Preacher K, Hayes A. **Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models.** *Behavior research methods*. 2008;40(3):879-891.

[80] Muthén L, Muthén B. **How to use a monte carlo study to decide on sample size and determine power.** *Structural equation modeling: A multidisciplinary journal*. 2002;9:599-620.

Figures' titles and legends.

Figure 1. ROIs with significant main effect of genotype (Met allele carriers > Val/Val homozygotes) (Jasińska et al., 2016)



ROI 1=bilateral precuneus extending into left inferior parietal lobule, ROI 2=bilateral hippocampus/parahippocampal gyrus/fusiform gyrus/cerebellum, ROI 3=left middle frontal gyrus/inferior frontal gyrus/thalamus, ROI 4=right cingulate/middle frontal gyrus/superior frontal gyrus, ROI 5=left cingulate/medial frontal gyrus/middle frontal gyrus/precentral gyrus, ROI 6=right superior temporal gyrus/inferior parietal lobule/superior parietal lobule.

Figure 2. The serial multiple mediation model

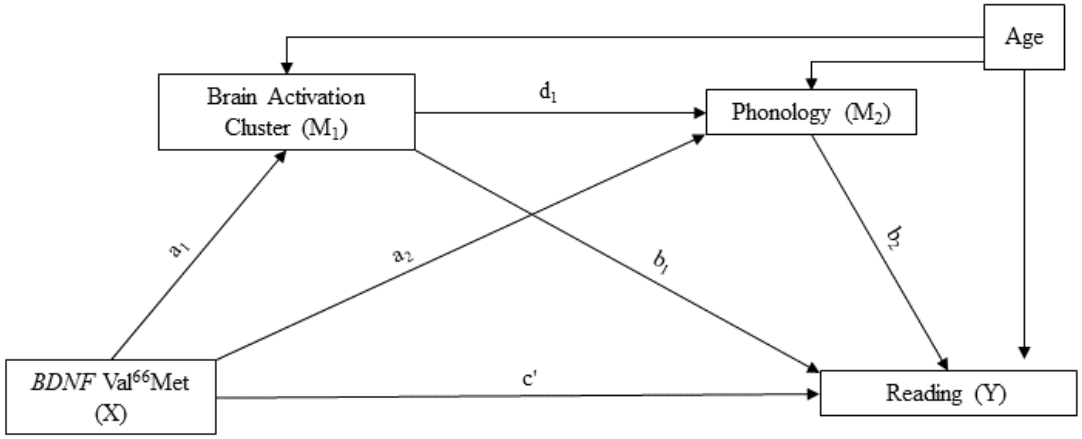


Table 1. Descriptive statistics of the language and reading skills (raw scores) in the total sample (n=94).

	Min	Max	Mean	Standard Deviation	Skewness	Kurtosis
WJ-Letter-word decoding	19.000	69.000	47.505	12.191	-0.418	-0.500
WJ-Word Attack	3.000	31.000	18.785	7.233	-0.293	-0.809
WJ-Spelling	13.000	49.000	29.827	8.919	0.291	-0.736
WJ-Passage comprehension	4.000	37.000	26.172	7.327	-0.890	0.226
WJ-Oral Comprehension	12.000	30.000	20.500	4.438	0.219	-0.747
WJ-Picture Vocabulary	16.000	34.000	25.032	3.914	0.107	0.003
CTOPP-Blending words	0.000	20.000	13.225	3.589	-0.648	0.992
CTOPP-Memory for digits	8.000	20.000	13.640	2.773	0.121	-0.491
CTOPP-Non-word repetition	3.000	16.000	9.549	2.676	0.123	0.027
CTOPP-Blending non-words	2.000	15.000	9.690	2.767	-0.401	0.783

WJ=Woodcock-Johnson Achievement Battery III (Woodcock et al., 2001); CTOPP=The Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 2009).
For both WJ and CTOPP subtests, higher scores correspond to better performance.

Table 2. Rotated Component Matrix (extraction method: principal component analysis; rotation method: promax)

	Components	
	Reading	**Phonology**
CTOPP_Blending Words	0.003	**0.809**
CTOPP_Memory for Digits	0.037	**0.582**
CTOPP_Non-Word Repetition	-0.035	**0.615**
CTOPP_Blending Non-Words	-0.038	**0.874**
WJ_Letter-Word decoding	**0.936**	0.028
WJ_Word Attack	**0.720**	0.234
WJ_Spelling	**0.865**	0.062
WJ_Passage Comprehension	**0.960**	-0.031
WJ_Oral Comprehension	**0.788**	0.002
WJ_PictureVocabulary	**0.886**	-0.161

CTOPPT=The Comprehensive Test of Phonological Processing (Wagner et al., 2009); WJ=Woodcock-Johnson Achievement Battery III (Woodcock et al., 2001).

Table 3. Descriptive statistics of the demographics, brain activation clusters and behavioral components in the total sample (n=94).

	Min	Max	Mean	Standard Deviation	Skewness	Kurtosis
Age	5.900	13.070	8.376	1.328	0.931	0.975
Full Scale IQ	76.000	153.000	113.392	15.668	0.154	-0.182
ROI 1	-0.870	2.224	0.327	0.557	0.554	1.007
ROI 2	-1.757	2.152	0.368	0.731	-0.090	0.613
ROI 3	-0.998	2.485	0.462	0.496	0.597	2.842
ROI 4	-0.968	1.509	0.272	0.418	0.525	1.516
ROI 5	-0.805	3.078	0.684	0.579	0.953	3.530
ROI 6	-0.724	2.171	0.384	0.525	0.638	0.980
Reading	-2.429	1.786	0.000	1.000	-0.338	-0.521
Phonology	-2.628	2.155	0.000	1.000	0.142	-0.329

ROI 1=left/right precuneus/inferior parietal lobule; ROI 2=left/right fusiform gyrus, hippocampus, cerebellum; ROI 3=left middle frontal gyrus, inferior frontal gyrus, thalamus; ROI 4=right middle frontal gyrus, superior frontal gyrus, cingulate; ROI 5=left cingulate, middle frontal gyrus, precentral gyrus; ROI 6=right superior temporal gyrus, inferior parietal lobule, superior parietal lobule.

Table 4. Correlation among *BDNF* Val^{66}Met, associated ROIs and our PCA identified behavioral metrics (n=94).

	ROI 1	ROI 2	ROI 3	ROI 4	ROI 5	ROI 6	Reading	Phonology
***BDNF* Val^{66}Met**	-0.384**	-0.319**	-0.416**	-0.431**	-0.315**	-0.352**	0.111	0.125
ROI 1	1	0.810**	0.805**	0.801**	0.788**	0.884**	0.288**	0.212*
ROI 2		1	0.828**	0.735**	0.730**	0.737**	0.354**	0.281**
ROI 3			1	0.830**	0.869**	0.723**	0.278*	0.257*
ROI 4				1	0.813**	0.745**	0.155	0.100
ROI 5					1	0.731**	0.277*	0.251*
ROI 6						1	0.213*	0.098
Reading							1	0.573*

** $p < 0.01$ (two-tails); * $p < 0.05$ (two-tails).

Table 5. Indirect effects of mediators and direct effects (unstandardized βs) of brain activation clusters and Phonology on Reading in the serial multiple mediation model.

Paths in Figure 2			β (95% CI)*					
			ROI 1	**ROI 2**	**ROI 3**	**ROI 4**	**ROI 5**	**ROI 6**
X on	**M_1**	a_1	*-0.897 (-1.322 / -0.500)*	*-0.752 (-1.182 / -0.320)*	*-0.923 (-1.352 / -0.521)*	*-0.736 (-1.224 / -0.314)*	*-0.988 (-1.440 / -0.560)*	*-0.822 (-1.291 / -0.390)*
	M_2	a_2	*0.477 (0.028 / 0.967)*	*0.479 (0.064 / 0.951)*	*0.558 (0.110 / 1.034)*	0.447 (0.015 / 0.892)	0.384 (-0.037 / 0.896)	0.335 (-0.088 / 0.800)
	Y	c'	0.100 (-0230 / 0.429)	0.117 (-0.204 / 0.451)	0.157 (-0.196 / 0.525)	0.057 (-0.262 / 0.394)	0.029 (-0.328 / 0.379)	0.066 (-0.264 / 0.399)
M_1 on	**M_2**	d_1	*0.281 (0.096 / 0.457)*	*0.337 (0.159 / 0.523)*	*0.360 (0.172 / 0.594)*	*0.301 (0.122 / 0.517)*	0.160 (-0.048 / 0.395)	0.133 (-0.068 / 0.330)
	Y	b_1	0.130 (-0.015 / 0.277)	*0.172 (0.020 / 0.328)*	*0.182 (0.042 / 0.347)*	0.101 (-0.024 / 0.232)	0.052 (-0.099 / 0.203)	0.105 (-0.026 / 0.242)
M_2 on	**Y**	b_2	*0.469 (0.344 / 0.589)*	*0.447 (0.312 / 0.571)*	*0.443 (0.308 / 0.570)*	*0.472 (0.340 / 0.593)*	*0.493 (0.362 / 0.617)*	*0.487 (0.362 / 0.611)*
X on Y *via* M_1		a_1b_1	-0.117 (-0.260 / 0.013)	*-0.129 (-0.289 / -0.013)*	*-0.168 (-0.337 / -0.036)*	-0.074 (-0.181 / 0.014)	-0.051 (-0.211 / 0.100)	-0.086 (-0.226 / 0.022)
X on Y *via* M_2		a_2b_2	*0.224 (0.013 / 0.478)*	*0.214 (0.027 / 0.439)*	*0.247 (0.050 / 0.476)*	*0.211 (0.007 / 0.448)*	0.189 (-0.018 / 0.457)	0.163 (-0.041 / 0.409)
X on Y *via* M_1 and M_2		$a_1d_1b_2$	*-0.118 (-0.227 / -0.034)*	*-0.113 (-0.227 / -0.035)*	*-0.147 (-0.279 / -0.057)*	*-0.104 (-0.206 / -0.028)*	-0.078 (-0.222 / 0.020)	-0.053 (-0.160 / 0.024)
X on Y *via* M1, M2		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2$	-0.011 (-0.230 / 0.238)	-0.028 (-0.257 / 0.199)	-0.068 (-0.301 / 0.160)	0.032 (-0.173 / 0.264)	0.060 (-0.180 / 0.334)	0.024 (-0.192 / 0.264)
X on Y		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2 + c'$	0.089 (-0.292 / 0.498)					

*Significant coefficients are reported in italics and underlined.

Supplementary Table 1. Age frequency in the total sample (n=94), Val/Val homozygotes (n=65) and in Met allele carriers (n=29).

	Total sample (n=94)			Val/Val homozygotes (n=65)			Met allele carriers (n=29)		
	Frequency	**Percentage**	**Cumulative percentage**	**Frequency**	**Percentage**	**Cumulative percentage**	**Frequency**	**Percentage**	**Cumulative percentage**
5-6	1	1.1	1.1	-	-	-	1	3.4	3.4
6-7	10	10.6	11.7	6	9.2	9.2	4	13.8	17.2
7-8	33	35.1	46.8	25	38.5	47.7	8	27.6	44.8
8-9	25	26.6	73.4	14	21.5	69.2	11	37.9	82.8
9-10	12	12.8	86.2	9	13.8	83.1	3	10.3	93.1
10-11	9	9.6	95.7	8	12.3	95.4	1	3.4	96.6
11-12	2	2.1	97.9	1	1.5	96.9	1	3.4	100.0
12-13	1	1.1	98.9	1	1.5	98.5	-	-	-
13-14	1	1.1	100.0	1	1.5	100.0	-	-	-

Supplementary Table 2. Indirect effects of mediators and direct effects (unstandardized βs) of brain activation clusters and Reading on Phonology in the serial multiple mediation model.

			β (95% CI)*					
			ROI 1	**ROI 2**	**ROI 3**	**ROI 4**	**ROI 5**	**ROI 6**
X on	**M_1**	a_1	*-0.897 (-1.322 / -0.500)*	*-0.752 (-1.183 / -0.320)*	*-0.923 (-1.352 / -0.521)*	*-0.735 (-1.224 / -0.313)*	*-0.988 (-1.440 / -0.560)*	*-0.822 (-1.291 / -0.390)*
	M_2	a_2	0.324 (-0.110 / 0.734)	0.332 (-0.070 / 0.743)	0.405 (-0.034 / 0.836)	0.268 (-0.132 / 0.681)	0.218 (-0.213 / 0.692)	0.229 (-0.183 / 0.664)
	Y	c'	0.269 (-0.081 / 0.625)	0.272 (-0.060 / 0.631)	0.309 (-0.040 / 0.667)	0.278 (-0.065 / 0.631)	0.239 (-0.102 / 0.607)	0.181 (-0.154 / 0.531)
M_1 on	**M_2**	d_1	*0.262 (0.096 / 0.430)*	*0.323 (0.165 / 0.500)*	*0.342 (0.181 / 0.535)*	*0.243 (0.089 / 0.415)*	0.131 (-0.076 / 0.325)	*0.170 (0.003 / 0.332)*
	Y	b_1	0.112 (-0.058 / 0.275)	0.136 (-0.036 / 0.310)	0.149 (-0.015 / 0.332)	*0.147 (0.002 / 0.309)*	0.073 (-0.073 / 0.242)	0.018 (-0.156 / 0.192)
M_2 on	**Y**	b_2	*0.643 (0.431 / 0.852)*	*0.624 (0.393 / 0.840)*	*0.617 (0.403 / 0.831)*	*0.633 (0.428 / 0.834)*	*0.666 (0.467 / 0.862)*	*0.673 (0.459 / 0.881)*
X on Y *via* M_1		a_1b_1	-0.101 (-0.273 / 0.051)	-0.102 (-0.281 / 0.023)	-0.138 (-0.338 / 0.013)	*-0.108 (-0.256 / -0.002)*	-0.072 (-0.265 / 0.069)	-0.015 (-0.179 / 0.126)
X on Y *via* M_2		a_2b_2	0.208 (-0.065 / 0.508)	0.207 (-0.043 / 0.505)	0.250 (-0.021 / 0.566)	0.170 (-0.085 / 0.463)	0.145 (-0.137 / 0.490)	0.154 (-0.120 / 0.473)
X on Y *via* M_1 and M_2		$a_1d_1b_2$	*-0.151 (-0.295 / -0.043)*	*-0.151 (-0.307 / -0.049)*	*-0.195 (-0.359 / -0.080)*	*-0.113 (-0.225 / -0.027)*	-0.086 (-0.261 / 0.047)	*-0.094 (-0.225 / -0.002)*
X on Y *via* M1, M2		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2$	-0.043 (-0.324 / 0.275)	-0.046 (-0.321 / 0.261)	-0.083 (-0.382 / 0.233)	-0.052 (-0.328 / 0.248)	-0.013 (-0.335 / 0.298)	0.045 (-0.238 / 0.358)
X on Y		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2 + c'$	0.226 (-0.165 / 0.628)					

*Significant coefficients are reported in italics and underlined.

The mediation model for each brain activation cluster (ROIs 1-6) explained 37.2%, 37.7%, 37.9%, 38.0%, 36.7% and 36.3% of the variance in the Phonology outcome, respectively. Using 2000 bootstrapping analyses and bias-corrected 95% CI, the significant indirect effects of X on Y *via* M1 and M2 (i.e. $a_1d_1b_2$) were the paths from the BDNF Val66Met polymorphism to Reading outcome via Phonology and activation in ROIs 1, 2, 3, 4 and 6. Inspection of beta scores revealed that the indirect effect along this pathway is negative. Specifically, Val/Val homozygotes have lower activation relative to Met allele carriers in brain ROIs 1, 2, 3, 4 and 6. Brain activation is positively associated with Reading, which is positively related to Phonology.

In addition, the indirect effect from the *BDNF* $Val^{66}Met$ polymorphism to Phonology outcome *via* ROI4 is significant (i.e. a_1b_1). Inspection of beta scores revealed that the indirect effect along this pathway is negative. That is, Val/Val homozygotes have lower activation relative to the Met allele carriers in brain ROI4; brain activation is positively related to Phonology.

Supplementary Table 3. Indirect effects of mediators and direct effects (unstandardized βs) of brain activation clusters and Phonology on Reading in serial multiple mediation models including only Caucasian participants (n=81).

Paths in Figure 2			β (95% CI)*					
			ROI 1	**ROI 2**	**ROI 3**	**ROI 4**	**ROI 5**	**ROI 6**
X on	**M_1**	a_1	*-0.709 (-1.302 / -0.343)*	*-0.622 (-1.119 / -0.165)*	*-0.840 (-1.344 / -0.381)*	*-0.631 (-1.198 / -0.133)*	*-0.910 (-1.430 / -0.456)*	*-0.720 (-1.240 / -0.242)*
	M_2	a_2	0.491 (-0.001 / 0.997)	*0.487 (0.016 / 0.994)*	*0.525 (0.039 / 1.025)*	0.427 (-0.035 / 0.914)	0.390 (-0.070 / 0.901)	0.374 (-0.085 / 0.870)
	Y	c'	0.131 (-0.204 /0.485)	0.126 (-0.204 /0.499)	0.172 (-0.182 /0.567)	0.069 (-0.269 /0.430)	0.092 (-0.264 /0.486)	0.086 (-0.239 /0.445)
M_1 on	**M_2**	d_1	*0.278 (0.078 / 0.466)*	*0.349 (0.164 / 0.544)*	*0.303 (0.116 / 0.537)*	*0.249 (0.078 / 0.458)*	0.132 (-0.068 / 0.338)	0.145 (-0.066 / 0.346)
	Y	b_1	0.154 (-0.003 / 0.319)	*0.180 (0.017 / 0.346)*	*0.190 (0.041 / 0.375)*	*0.103 (-0.040 / 0.242)*	0.101 (-0.065 / 0.253)	0.118 (-0.029 / 0.268)
M_2 on	**Y**	b_2	*0.417 (0.267 / 0.554)*	*0.395 (0.233 / 0.538)*	*0.402 (0.246 / 0.545)*	*0.431 (0.281 / 0.567)*	*0.444 (0.300 / 0.581)*	*0.440 (0.297 / 0.584)*
X on Y *via* M_1		a_1b_1	-0.123 (-0.285 / 0.001)	*-0.112 (-0.279 / -0.005)*	*-0.160 (-0.341 / -0.029)*	-0.065 (-0.187 / 0.020)	-0.092 (-0.273 / 0.054)	-0.085 (-0.233 / 0.015)
X on Y *via* M_2		a_2b_2	0.205 (0.000 / 0.435)	*0.192 (0.007 / 0.409)*	*0.211 (0.015 / 0.434)*	0.184 (-0.017 / 0.407)	0.173 (-0.030 / 0.429)	0.165 (-0.039 / 0.395)
X on Y *via* M_1 and M_2		$a_1d_1b_2$	*-0.092 (-0.190 / -0.021)*	*-0.086 (-0.181 / -0.019)*	*-0.102 (-0.200 / -0.031)*	*-0.068 (-0.149 / -0.007)*	-0.053 (-0.173 / 0.024)	-0.046 (-0.139 / 0.016)
X on Y *via* M1, M2		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2$	-0.010 (-0.232 / 0.229)	-0.005 (-0.233 / 0.222)	-0.051 (-0.285 / 0.180)	0.051 (-0.152 / 0.291)	0.028 (-0.199 / 0.283)	0.034 (-0.189 / 0.270)
X on Y		$a_1*b_1 + a_2*b_2 + a_1*d_1*b_2 + c'$	0.120 (-0.264 / 0.525)					

*Significant coefficients are reported in italics and underlined.