



A genome-wide association study reveals a substantial genetic basis underlying the Ebbinghaus illusion

Z a Zh ¹ · B g Che ^{2,3} · Re Na² · Wa Fa g^{2,4} · We a Zha g² · Q Zh ⁵ · Sha b Zh ⁶ · Ha Le ⁵ ·
A g H a g⁵ · T g e Che ⁵ · D g he g N⁷ · Y g G ⁷ · J a g L ⁷ · Y Ra ^{2,4} · Fa g Fa g^{2,8}

Received: 26 May 2020 / Revised: 3 August 2020 / Accepted: 4 August 2020 / Published online: 16 September 2020
© The Author(s), under exclusive licence to The Japan Society of Human Genetics 2020

Abstract

The Ebbinghaus illusion (EI) is an optical illusion of relative size perception that reflects the contextual integration ability in the visual modality. The current study investigated the genetic basis of two subtypes of EI, EI overestimation, and EI underestimation in humans, using quantitative genomic analyses. A total of 2825 Chinese adults were tested on their magnitudes of EI overestimation and underestimation using the method of adjustment, a standard psychophysical protocol. Heritability estimation based on common single nucleotide polymorphisms (SNPs) revealed a moderate heritability (34.3%) of EI overestimation but a nonsignificant heritability of EI underestimation. A meta-analysis of two phases (phase 1: $n = 1986$, phase 2: $n = 839$) of genome-wide association study (GWAS) discovered 1969 and 58 SNPs reaching genome-wide significance for EI overestimation and EI underestimation, respectively. Among these SNPs, 55 linkage-disequilibrium-independent SNPs were associated with EI overestimation in phase 1 with genome-wide significance and their associations could be confirmed in phase 2 cohort. Gene-based analyses found seven genes to be associated with EI overestimation at the genome-wide level, two from meta-analysis, and five from classical two-stage analysis. Overall, this study provided consistent evidence for a substantial genetic basis of the Ebbinghaus illusion.

These authors contributed equally: Zijian Zhu, Biqing Chen, Ren Na

S **a** **a** The online version of this article (<https://doi.org/10.1038/s10038-020-00827-4>) contains supplementary material, which is available to authorized users.

✉ Yi Rao
yrao@pku.edu.cn
✉ Fang Fang
ffang@pku.edu.cn

- ¹ School of Psychology, Shaanxi Normal University, 710062 Xi'an, China
- ² PKU-IDG/McGovern Institute for Brain Research, and Peking-Tsinghua Center for Life Sciences, Peking University, 100871 Beijing, China
- ³ Central Laboratory, Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, 210029 Nanjing, China

Introduction

The Ebbinghaus illusion (EI) is a visual phenomenon that the perceived size of a central disk differs from its physical “reality” due to a size contrast between the disk and its surrounding disks. Specifically, an identical disk is perceived larger than its physical size when surrounded by small disks (EI overestimation) and smaller when surrounded by large disks (EI underestimation) (Fig. 1a). The

- ⁴ Beijing Innovative Center for Genomics, Peking University School of Life Sciences, and National Institute of Biological Sciences, 102206 Beijing, China
- ⁵ College of Laboratory Medicine, Chongqing Medical University, 400016 Chongqing, China
- ⁶ University-Town Hospital of Chongqing Medical University, 401331 Chongqing, China
- ⁷ Division of Molecular Nephrology and Creative Training Center for Undergraduates, M.O.E. Key Laboratory of Medical Diagnostics, College of Laboratory Medicine, Chongqing Medical University, 400016 Chongqing, China
- ⁸ School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, 100871 Beijing, China

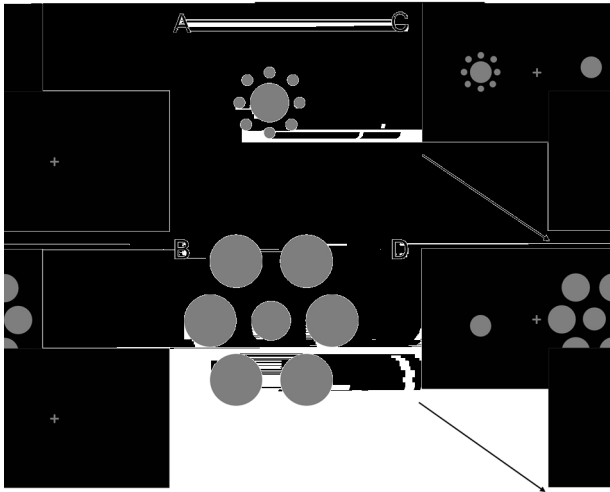


Fig. 1 Experimental stimuli and procedure for EI estimation. In the stimuli for EI overestimation (a) and EI underestimation (b), the

magnitude of EI, which is suggested to reflect the contextual integration ability in the visual modality, varies widely across individuals. Particularly, populations with mental disorders such as autism (1, 2, but also see 3) and disorganized schizophrenia/schizotypy [4] tend to show a smaller magnitude of EI, possibly due to a reduced contextual integration ability in these diseases [5]. Therefore, investigating the molecular mechanism of EI would improve our understanding of this visual phenomenon.

What are the sources of the individual differences in EI? Evidence has suggested the contribution of environmental factors to EI. For instance, children show a much lower degree of EI than adults do [6], and people within a culture show a smaller variation in EI than people across cultures [7]. On the other hand, studies have implicated a genetic origin of EI, as EI is observed in infants as young as 5–8 months [8]. To date, there is only one study that has explored the heritability of EI. Coren and Porac (1979) performed a family study, which tested the relative behavioral resemblance between family or nonfamily members. A significant familial resemblance was found, providing evidence for the influence of genetic factors to EI [9]. Indirect brain imaging evidence showed that the magnitude of EI is associated with the surface area of human primary visual cortex (V1) [10], which is explainable by additive effects of SNPs dispersed throughout the genome with a large heritable effect ($h^2 \approx 0.45$) [11]. However, to what

extent genetic factors contribute to EI and which genes underlie this phenomenon remain unknown.

The current study aimed to understand the source of individual differences and the molecular mechanism of EI by combining the genome-wide association studies (GWAS) technique with psychophysics. We first assessed the common single nucleotide polymorphism (SNP) heritability of two EI subtypes (i.e., EI overestimation and EI underestimation) using the GCTA program [12]. We then applied a meta-analysis of two phases of GWAS procedure at the levels of single nucleotide polymorphism (SNPs), gene, and pathway to identify genetic elements associated with EI overestimation and underestimation. Finally, we examined the relation between the validated SNPs and the anatomical characteristics, including the surface area, volume, and cortical thickness, of early visual cortical areas (i.e., V1, V2, and V3) using magnetic resonance imaging in 123 independent subjects. Our results provided strong support the involvement of specific genetic variants in the EI.

Methods

Participants

The GWAS cohorts consisted of college students recruited from the Chongqing Medical University in south China. Phase 1 cohort consisted of 1987 participants, with a mean age of 19 years (standard deviation (SD) = 1), 78% female, 90% Han. Phase 2 cohort consisted of 839 participants, with a mean age of 20 years (SD = 1), 88% female, 91% Han. No significant difference in ethnicity or place of origin was found between the two phases by Pearson's Chi-square tests. The MRI (magnetic resonance imaging) cohort of 123 participants (Mean = 21, SD = 2, 66% female, 96% Han) were recruited from Peking University in China. The psychiatric and medication history of each participant was screened by medical examination at the time of college entrance and was inquired by a self-reported questionnaire at the beginning of the study. None of the participants reported having neurological diseases or vision disorders. All participants had normal or corrected-to-normal visual acuity, which was measured with the Freiburg Vision Test implemented in the software program FrACT (Version 3.8.2) (see [13]).

The experimental procedures were approved by the human subject review committee in the School of Psychological and Cognitive Sciences at Peking University. All participants were naive to the experimental purpose. They gave written informed consents before the experiment. The behavioral data were collected between September in 2014 and June in 2015, along with other behavioral tests that have been published elsewhere [13, 14].

Behavioral assay

Stimuli

Two stimuli of the EI were created, one for the EI overestimation (Fig. 1a) and the other for the EI underestimation test (Fig. 1b). Each EI stimulus was composed of one central disk surrounded by six larger or eight smaller inducing disks (luminance: 77 cd/m²), and was presented on a dark background (luminance: 31 cd/m²). For the overestimation EI stimulus (Fig. 1a), the central disk was 2.64° in diameter, the eight inducing disks were 0.71° in diameter, and the center-to-center distance was 1.70°. For the underestimation EI stimulus (Fig. 1b), the central disk was 2.64° in diameter, the six inducing disks were 3.55° in diameter, and the center-to-center distance was 3.93°.

The stimuli were generated and controlled using MATLAB and PsychToolbox. The stimuli were displayed on a 21-inch LCD monitor with a refresh rate of 60 Hz. The viewing distance was kept at 40 cm using a chinrest. Luminance was measured with a MINOLTACS-100A Chroma Meter (Konica Minolta Sensing Americas, Inc, USA).

Procedure

The method of adjustment was used to measure the perceived size of the central disk (Fig. 1c, d). In each trial, the EI stimulus was presented 12.5° to the left or right of the central fixation, and a test disk was presented 12.5° to the other side of the central fixation. The central disk was kept 2.64° in diameter for all trials, while the initial size of the test disk ranged from 0.5–0.68° smaller or larger than the central disk. The spatial location of the test disk was jittered in the vertical direction, which was 0.325–3.25° lower or higher than the central fixation. Randomization of the initial size and location of the test disk was performed to avoid perceptual adaptation. Participants were instructed to adjust the size of the test disk using the Up and Down keys in the keyboard. They pressed Enter to end adjusting when they perceived the size of the test disk identical to that of the central disk in the EI stimulus. The participants were allowed to estimate the size of the disks only by viewing. There was no time limit for adjusting.

Each participant completed four blocks of 25 adjusting trials, two blocks for the overestimation and two blocks for the underestimation measurements. Therefore, 50 trials were performed for each EI subtype. The magnitude of EI was quantified as the difference between the actual diameter of the central disk and the adjusted diameter of the test disk averaged across the 50 trials (EI overestimation: *perceived size* – *physical size*; EI underestimation: *physical size* – *perceived size*).

Genotyping and quality control

DNA was extracted from peripheral blood of participants using the QuickGene whole-blood genome DNA extract system (Kurabo Industries Ltd., Japan), and was genotyped for 894,517 common SNPs using the HumanOmniZhongHua-8 Beadchip v1.2 (Illumina, Inc., San Diego, CA, USA) in phase 1 and the HumanOmniZhongHua-8 Beadchip v1.1 in phase 2. Common quality control parameters (call rate ≥ 0.95 , minor allele frequency (MAF) ≥ 0.01 , and Hardy–Weinberg equilibrium test with $p \geq 10^{-4}$) were applied to the genotype data, which retained 830,937 SNPs in both cohorts. Differences in allele frequencies between the two phases were examined using paired-sample *t*-tests, which showed no significant deviations. Potential duplicates or close relatives were examined by calculating the identity-by-state (IBS) similarity matrix implemented in PLINK [15]; none was excluded due to IBS distance ≤ 0.75 (IBS distance for any pair of individuals > 0.90). Population stratification within each cohort was examined with EIGENSTRAT [16] via the principal component (PC) analysis. Genetic outliers were detected and excluded automatically with the default mode. The maximum number of outlier removal iterations was 5. One individual outside six SDs from the population mean in terms of genetic PC was removed as an outlier. The genetic background of the ethnic groups in the remaining populations did not significantly differ from that of the Han individuals. After quality control, a total of 1986 participants were included in the final analysis of the phase 1 cohort and 839 individuals were retained in the phase 2 cohort. Array-based data has been deposited at the European Genome–phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS00001003639.

Statistical genetic analyses

Heritability estimation

Estimation of the phenotypic variance explained by all common SNPs (SNP-based heritability) was carried out by GCTA version 1.24 [12]. This method measures the variance in a trait that is due to linkage disequilibrium (LD) between genotyped SNPs and unknown causal variants. The two cohorts were combined for heritability analysis. We first computed the genetic relationship matrix on every two of the participants using all autosomal markers that were genotyped and checked for quality control as described above (altogether 830,937 SNPs). Twenty-nine individuals were excluded due to estimated genetic relatedness larger than 0.025 [17], leaving 2796 individuals for further analysis. We then estimated the proportion of the total

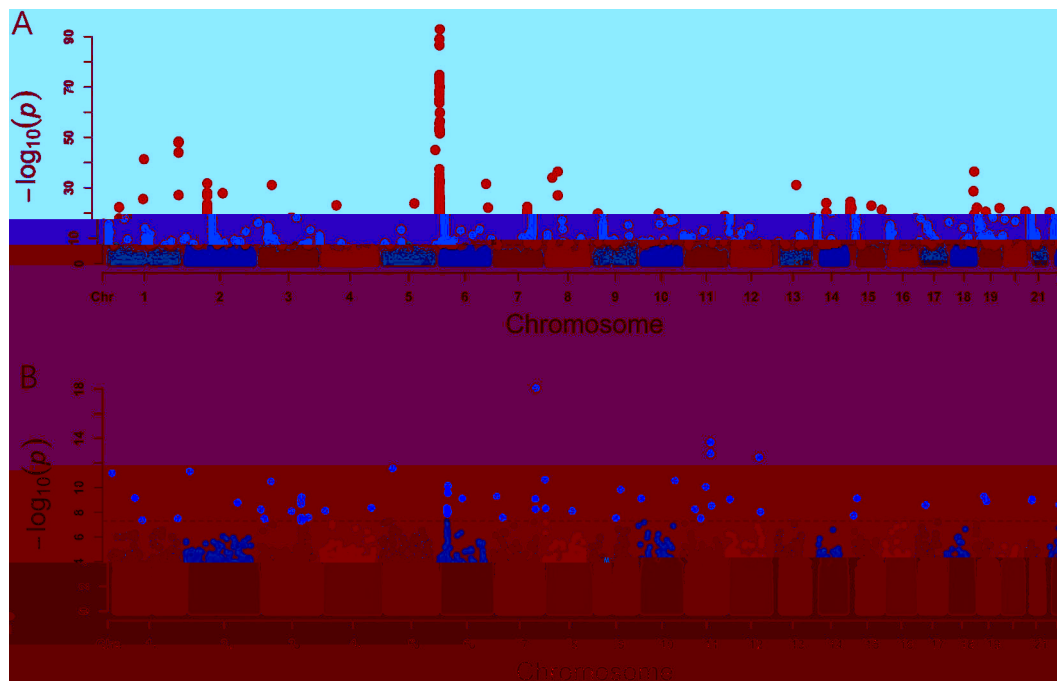


Fig. 2 Manhattan plots of the meta-analysis of genome-wide association study (GWAS) results. GWAS results [$-\log_{10}(p)$ value] are shown in a chromosomal order for individually genotyped quality-control-positive single nucleotide polymorphisms (SNPs) that were tested for linear regression with EI overestimation (a) and EI

underestimation (b). Results were based on an additive genetic model. Each dot represents one SNP. The red dotted line indicates the genome-wide significance level (5×10^{-8}). Genome-wide significant SNPs are shown as enlarged red dots. Chromosomes are shown in different colors for clarity

phenotypic variance explained by all common SNPs using the restricted maximum likelihood (REML) algorithm, with the first 20 eigenvectors from GCTA-PCA included as covariates. The online GCTA-GREML Power Calculator was used to calculate the statistical power [18].

Imputation

Genotypes were pre-phased into haplotypes with SHAPEIT [19, 20]. Imputation was then performed using IMPUTE v2.3.1 [21] based on 1000 Genomes haplotype data (Phase I integrated variant set release (SHAPEIT2) in NCBI build 37/UCSC *hg19*) with 36,820,992 SNPs, 1,384,273 short biallelic indels, and 14,017 structural variations. Quality control retained variants with missing data proportion $< 1 \times 10^{-6}$, $\text{INFO} > 0.8$, and $\text{MAF} > 0.01$ for subsequent association analysis.

GWAS

The magnitude of EI was used as the quantitative trait. Effects of demographic factors including sex and ethnicity showed no association with the phenotypes as estimated by linear regression. Because all participants were college freshmen and sophomores within a very narrow age range, age was not considered as a covariate in the analysis. Thus, no covariate was included in subsequent association tests.

Association tests of imputed SNPs were performed with SNPTEST v2.5 [22], based on the Score Method in the Frequentist Test framework. An additive model was assumed. Basic statistical analyses were performed, the genomic inflation factor λ was calculated, quantile–quantile (Q–Q) plots for observed versus expected p values were created (Fig. 2), and Manhattan plots (Fig. 3) illustrating all the association results were generated, all by R v3.2.1 (<https://www.R-project.org>). The genome-wide significance threshold was set at 5×10^{-8} . Regional association plots were drawn for two mega-base around top candidate SNPs by the web-based LocusZoom (<http://locuszoom.sph.umich.edu/locuszoom>), with variants that passed quality control. Quanto v1.2 was used for power calculation [23]. All the genomic positions reported in the main text were based on the *hg19* unless otherwise stated.

Meta-analysis

To summarize results for the two phases, we carried out a sample size weighted meta-analysis of all the QC passed imputed genotypes using METAL [24].

Gene-based and pathway-based analyses

Gene- and pathway-based association analyses were carried out by MAGMA [25]. The gene-based statistics were

derived using the GWAS results of each phenotype. The SNP to gene mapping was based on the NCBI37 (hg19) built with no additional boundary placed around the genes, resulting in 17,959 genes being analyzed. The South Chinese panel of the 1000 Genomes data (phase 1, release 3) was used as a reference panel to account for LD. Pathway databases included gene ontology (GO), KEGG, REACTOME, BIOCARTA, and PANTHER (9740 pathways in total). The genome-wide significance threshold was set as 0.05 (Bonferroni corrected), which was $0.05/17,803 = 2.8 \times 10^{-6}$ for the gene-based analysis and $0.05/9,740 = 5.13 \times 10^{-6}$ for the pathway-based analysis.

Candidate selection

Linkage disequilibrium (LD) independent SNPs, those with genome-wide significance and low LD ($r^2 < 0.1$) to a more significantly associated SNP within a 500 kb window, were selected from GWAS analysis of phase 1, and then associations were tested in the phase 2 cohort by SNPTEST v2.5 under a general linear regression model. LD scores for each SNP were calculated using individuals of Chinese (Han

was uhcount7-2.4867999carnelmaW-93600006Tf0fypinaslysis.

rank correlations between the two testing blocks for each EI stimulus. We observed high correlations between each pair of blocks (Spearman's rank correlation coefficient phase 1: $r_{1,2 \text{ overestimate}} = 0.707, p < 0.001$; $r_{1,2 \text{ underestimate}} = 0.781, p < 0.001$; phase 2: $r_{1,2 \text{ overestimate}} = 0.724, p < 0.001$; $r_{1,2 \text{ underestimate}} = 0.792, p < 0.001$), confirming the high test-retest reliability of the measurements. However, the Spearman's rank correlation between the overestimation and underestimation stimuli was relatively low ($r_{\text{phase1}} = 0.221, p < 0.001$; $r_{\text{phase2}} = 0.254, p < 0.001$), indicating dissociable processes underlying the two illusion effects. Genetic analysis was therefore performed separately for the two EI effects.

Heritability estimation

The GREML method was used to estimate the SNP heritability of the magnitude of EI [12]. This model estimated the common-SNP-based heritability through quantifying the proportion of the phenotypic variance explained by the 830,937 genotyped autosomal SNPs. We found a significant heritability of EI overestimation ($p = 0.006$). Specifically, 34.3% (SEM = 14.1%) of the phenotypic variance in EI overestimation could be explained by the common SNPs. In contrast, the heritability of EI underestimation was not significant, with only 3.2% (SEM = 14.3%, $p = 0.412$) of the phenotypic variance being explained by the common SNPs. This SNP-based analysis revealed a moderate heritability for the overestimation subtype of the EI.

Genome-wide study of single-marker association

A total of 1986 and 839 participants were retained in phase 1 and phase 2 after quality control. A quantitative GWAS was performed on the imputed markers after quality control and meta-analysis of the two phases was performed. The genomic inflation factor λ was 1.007 for EI overestimation and 0.991 for EI underestimation, confirming little influence of population stratification. 1969 and 58 SNPs reached genome-wide significance ($p < 5 \times 10^{-8}$) for EI overestimation and underestimation (Fig. 3a, b and Table S1). Among them, 70 and two LD independent SNPs were identified to be associated with EI overestimation and underestimation, respectively, at a genome-wide significance in phase 1. They were aggregated together in terms of locations on the chromosome (e.g., chr 6,7,16,19,21,22). Fifty-five of them showed significance ($p < 6.9 \times 10^{-4} = 0.05/70$) in phase 2 (Table 1).

Gene-level and pathway-level association analyses

Gene-based GWA analysis tests the joint association of SNP markers belonging to the same gene with the

phenotype. It reduces the number of tests and makes it possible to detect effects consisting of multiple weaker associations. Five genes reached the genome-wide significance ($ps < 2.57 \times 10^{-6}$) in phase 1 and showed Bonferroni-corrected significance ($p < 0.01 = 0.05/5$) in phase 2 for EI overestimation (Table 2). We then performed a meta-analysis on the two cohorts and identified another two genes (*IGL* and *LOC100652901*) that showed genome-wide association with EI overestimation (Table 2). Notably, among the seven genome-wide significant genes, three (*PSG1*, *PSG6*, and *PSG7*) belong to the *PSG* family, hinting a potential link between this gene family with the visual illusion.

We further carried out a pathway analysis to explore potential associations between biological pathways and the two EI subtypes. None of the pathways reached significance after Bonferroni correction (threshold = 5.13×10^{-6}). Table S2 listed the pathways that showed nominal significance in both phases ($ps < 0.05$).

Relationships between the top associated SNPs and the anatomical features of early visual cortical areas

We performed a functional validation based on the brain anatomical information of early visual cortical areas (i.e., V1, V2, and V3). Structural MRI data were collected from an independent sample of 123 participants. The volume, surface area, and cortical thickness were calculated for each hemisphere of each area. The 55 SNPs significantly associated with EI overestimation in both phases in the SNP-based GWAS were selected as candidate SNPs. Because rs76473868 has no polymorphism among the 123 participants, 54 SNPs were included in the final analysis. The corrected significance threshold was thus $p = 3.43 \times 10^{-5}$ ($0.05/(54 \times 9 \times 3)$) (Table S3).

Discussion

This study is the first systematic whole-genome analysis of the EI in humans. A substantial genetic basis was found for EI at multiple levels. First, the GREML method revealed a moderate heritability (34.3%) of EI overestimation based on common SNPs. Second, SNP-based GWAS meta-analysis identified 1969 SNPs associated with EI overestimation at the genome-wide level, among which 55 independent SNPs were discovered and replicated in a two-stage analysis; GWAS meta-analysis on EI underestimation identified 58 genome-wide significant SNPs. Third, gene-based GWAS identified seven genes associated with EI overestimation. These results provided consistent evidence for the genetic basis of the overestimation subtype of the EI.

Table 1 Association results for top genome wide significant independent SNPs

SNP	Position	MAF	$\beta_{\text{phase 1}}$	$SE_{\text{phase 1}}$	$P_{\text{phase 1}}$	$\beta_{\text{phase 2}}$	$SE_{\text{phase 2}}$	$P_{\text{phase 2}}$	Alleles	META
<i>Underestimation</i>										
rs150946534	chr4:166472009	0.02	-0.70	0.13	4.71E-08	-0.16	0.18	3.56E-01	A/G	3.73E-07
rs9386728	chr6:97127920	0.08	-0.32	0.06	2.73E-08	0.06	0.09	5.21E-01	T/C	1.63E-05
<i>Overestimation</i>										
rs11264578	chr1:156837416	0.21	0.33	0.04	6.73E-19	0.33	0.06	6.33E-09	T/C	2.68E-26
rs1934073	chr1:159936733	0.17	0.44	0.04	8.51E-29	0.49	0.06	5.48E-15	T/G	4.28E-42
rs2275558	chr1:164529120	0.30	0.20	0.03	1.95E-09	0.19	0.05	1.93E-04	G/A	1.62E-12
rs60732244	chr2:7960904	0.10	-0.67	0.05	2.38E-34	-0.68	0.08	5.05E-16	G/T	1.05E-48
rs2139376	chr2:112143413	0.17	-0.37	0.04	4.76E-20	-0.30	0.06	5.14E-06	T/C	2.63E-24
rs79895937	chr2:127100834	0.10	0.28	0.05	4.39E-08	0.29	0.08	1.73E-04	T/C	3.21E-11
rs61095853	chr2:168510878	0.13	-0.44	0.05	3.82E-22	-0.36	0.07	4.42E-08	C/A	1.32E-28
rs6807284	chr3:6225539	0.29	0.25	0.03	2.49E-13	0.20	0.05	7.44E-05	C/T	1.08E-16
rs2276800	chr3:8675539	0.16	-0.32	0.04	6.71E-14	-0.26	0.07	1.42E-04	G/A	6.48E-17
rs6796799	chr3:75377277	0.19	-0.38	0.04	4.11E-23	-0.39	0.06	2.69E-10	T/A	7.62E-32
rs28716006	chr4:3935145	0.11	-0.27	0.05	4.64E-08	-0.30	0.07	4.00E-05	G/A	9.08E-12
rs11732191	chr4:84005377	0.26	-0.30	0.04	9.31E-17	-0.31	0.05	1.43E-08	A/G	8.28E-24
rs11957895	chr5:148216692	0.16	0.37	0.04	9.38E-19	0.33	0.06	2.61E-07	G/A	1.61E-24
rs2472788	chr6:14735129	0.18	-0.43	0.04	5.25E-29	-0.54	0.06	8.50E-19	A/T	9.58E-46
rs2523808	chr6:29849657	0.11	-0.41	0.04	5.32E-21	-0.45	0.07	5.54E-10	T/C	1.97E-29
rs76473868	chr6:29870251	0.40	0.36	0.03	1.24E-25	0.33	0.05	5.57E-10	A/G	5.39E-34
rs9260679	chr6:29926377	0.20	0.57	0.03	7.3E-60	0.63	0.05	6.24E-32	A/G	8.76E-90
rs2571377	chr6:29938571	0.43	0.44	0.03	3.39E-50	0.47	0.05	4.72E-25	G/T	2.00E-73
rs12661394	chr6:29954115	0.20	0.39	0.04	1.42E-27	0.41	0.06	3.59E-12	C/G	3.85E-38
rs140514288	chr6:29968734	0.16	-0.43	0.04	4.93E-24	-0.41	0.07	7.86E-09	A/AAAGG	3.15E-31
rs10947055	chr6:30093364	0.12	0.31	0.05	3.82E-12	0.27	0.07	1.47E-04	T/C	3.01E-15
rs2516677	chr6:30424623	0.32	-0.24	0.03	1.69E-13	-0.26	0.05	2.23E-07	A/C	2.20E-19
rs3130431	chr6:31219342	0.39	-0.17	0.03	4.94E-09	-0.07	0.05	1.08E-01	A/G	7.48E-09
rs2524070	chr6:31244520	0.17	-0.23	0.04	5.48E-09	-0.17	0.06	8.74E-03	G/A	2.64E-10
rs5025314	chr6:31343710	0.36	-0.22	0.03	4.28E-13	-0.17	0.05	1.58E-04	T/C	4.13E-16
rs4959070	chr6:31345178	0.12	0.32	0.04	2.13E-13	0.23	0.06	2.05E-04	G/C	2.89E-16
rs77523206	chr6:31436465	0.18	0.65	0.04	1.06E-70	0.60	0.06	5.81E-25	A/T	1.20E-93
rs116770642	chr6:31439498	0.16	-0.63	0.04	1.06E-46	-0.56	0.07	1.25E-15	G/A	2.20E-60
rs2905741	chr6:31452292	0.17	0.47	0.04	5.37E-38	0.47	0.06	5.69E-17	G/A	2.73E-53
rs3093994	chr6:31490372	0.17	-0.28	0.04	4.57E-13	-0.26	0.06	2.66E-06	G/A	6.29E-18
rs1055385	chr6:31501548	0.21	0.32	0.04	2.48E-19	0.27	0.05	6.37E-07	G/A	1.17E-24
rs34636308	chr6:31777687	0.12	0.32	0.04	2.03E-13	0.15	0.06	1.21E-02	C/T	5.19E-14
rs4714901	chr6:46039788	0.14	0.26	0.05	3.10E-08	0.10	0.07	1.31E-01	A/G	4.65E-08
rs3933566	chr6:158034980	0.21	0.23	0.04	5.95E-10	0.26	0.06	1.29E-05	C/A	3.77E-14
rs719423	chr7:7355115	0.23	0.30	0.04	1.83E-15	0.33	0.06	5.72E-09	G/C	7.42E-23
rs10225026	chr7:149820951	0.23	0.24	0.04	3.8E-10	0.23	0.06	4.87E-05	T/C	8.40E-14
rs6464052	chr7:149866114	0.17	0.38	0.04	4.97E-17	0.37	0.07	1.19E-07	T/G	3.47E-23
rs10093037	chr8:16988955	0.10	0.34	0.05	2.15E-11	0.21	0.08	9.47E-03	C/A	2.10E-12
rs1010754	chr8:70867648	0.25	-0.25	0.04	1.32E-12	-0.28	0.06	1.00E-06	T/C	7.16E-18
rs1889039	chr9:139525652	0.14	0.29	0.05	1.81E-09	0.23	0.07	1.68E-03	C/T	1.43E-11
rs2913115	chr10:80028339	0.16	0.28	0.04	2.75E-11	0.30	0.06	1.34E-06	C/T	2.07E-16
rs744680	chr10:131741695	0.12	0.36	0.05	2.00E-14	0.28	0.07	6.17E-05	G/A	8.12E-18
rs11032612	chr11:34338180	0.12	-0.30	0.05	2.42E-10	-0.17	0.07	1.63E-02	A/C	3.64E-11

In our SNP-based GWAS analysis, as many as 1969 SNPs and 58 SNPs were discovered to be associated with EI overestimation and underestimation at the genome-wide level, respectively. The large number of the genome-wide significant SNPs suggests a strong power of the present test. One possible reason is the high sensitivity and stability of the testing protocol. First, the method of adjustment we employed is a classical psychophysical measurement that is particularly appropriate for perceptual illusion estimation; as reflected by the high test-retest reliability, it provided a precise measure of the magnitude of EI. Second, each subtype was tested with 50 independent trials, which ensured a stable measurement of the perceptual threshold. Third, interference factors such as perceptual adaptation were excluded by randomizing the initial size and location of the test disk. The current results recommend future GWAS studies on cognitive abilities to increase the sensitivity and reliability of the behavioral test.

Although the function of the discovered SNPs remains elusive, relationships between some of these SNPs and related cognitive processes are implicated by the current study. For instance, 587 SNPs that were associated with EI overestimation span 1928 kilobase (kb) on chromosome 6 (29,849,657–31,777,687). This locus is close to the locus on chromosome 6 (28,303,247–28,712,247) that has been found to be associated with schizophrenia in populations with European ancestry [26], which is consistent with findings that people suffering from schizophrenia tend to show lower magnitude of EI [4, 27, 28]. However, consistent common variant genetic architecture of schizophrenia across EAS and EUR samples is found only outside of the Major histocompatibility complex (MHC) region [29]. Notably, the association between the locus on chromosome 6 and schizophrenia was mainly found in populations with European ancestry [29].

Table 2 Genes with genome-wide significance by MAGMA

Gene	Genomic position	nSNPs_phase 1	nSNPs_phase 2	P_phase 1	P_phase 2	P_meta-analysis
<i>HNRNPCL1</i>	chr1:12907261-12908578	1	1	9.44E-09	4.83E-06	5.20E-13
<i>PDE4DIP</i>	chr1:144851424-145076186	2	2	1.13E-10	2.34E-03	3.47E-12
<i>PSG1</i>	chr19:43370613-43383921	36	37	5.99E-07	2.06E-05	1.43E-10
<i>PSG7</i>	chr19:43428284-43441330	19	19	7.94E-07	5.45E-05	4.30E-10
<i>PSG6</i>	chr19:43406234-43422076	24	27	1.16E-06	1.07E-04	1.14E-09
<i>IGL</i>	chr22:22380474-23265085	992	996	4.85E-05	8.93E-05	5.46E-08
<i>LOC100652901</i>	chr22:25739549-25741017	5	5	3.91E-05	3.60E-03	8.93E-07

were revealed by relatively small samples [30, 31]. The MHC region is also indicated in the association studies of Autistic Spectrum Disorder (ASD) [32], echoing the findings that ASD patients tend to show a smaller magnitude of EI [1, 2]. This raises a hypothesis that EI might serve as a common cognitive basis for schizophrenia and autism, which can be explored by future research in larger Asian populations. The MHC gene family encodes cell surface molecules for immunological recognition. In recent decades, accumulating data has suggested expression of MHC-I in an isoform- and region-specific manner in the brain, such as lateral geniculate nucleus, cortex, and the hippocampus [33]. It is mainly involved in the development and maintenance of neuronal circuitry [34]. The MHC pathway has been suggested to be necessary in the synaptic plasticity in the visual processing areas of the brain [35]. Although its function in the high-level visual cortex and cognition has not been explicated, findings of the current study imply a potential role of the MHC pathway in human visual cognition. Finally, the significant SNP rs11264578 on chromosome 1 is located in the intron of gene *NTRK1*. Gene *NTRK1* has been identified to encode a member of the neurotrophic tyrosine kinase receptor family, which is involved in specifying sensory neuron subtypes and is related to the development of white matter microstructure [36]. Mutations in *NTRK1* have been associated with self-mutilating behavior [37] and cognitive disabilities [38]. The current finding further implies its possible function in EI or EI related cognitive functions.

The two-stage GWAS revealed five genes associated with EI overestimation at the genome-wide level. Among the five genes, gene *HNRNPCL1*, heterogeneous nuclear ribonucleoprotein C-like 1, has been predicted to play a role in nucleosome assembly and is a candidate gene in Caucasian Alzheimer's disease and mild cognitive impairment [39]. Gene *PDE4DIP*, phosphodiesterase 4D interacting protein, is a protein involved in cardiac muscle contraction and has been suggested to play a role in cardiomyopathy [40]; it is among the genes that were affected by 1q21.1 microdeletion and 1q21.1 microduplication syndromes [41]. The remaining genes, pregnancy-specific glycoproteins

(PSGs), are a complex consisting of carbohydrate and protein and are present in the mammalian body specifically during pregnancy [42]. Their relationships with EI related processes are unknown.

We performed functional validation by examining the relationship between the significant SNPs and the surface area, the volume, and the cortical thickness of early visual cortical areas (i.e., V1, V2, and V3). This is based on previous findings that the degree of EI can be predicted by the surface area of V1 in the left hemisphere [10]. We measured and compared these anatomical attributes of V1 among people with different genotypes of the 55 validated SNPs. However, we did not find difference in volume, surface area, or cortical thickness for V1-V3 in either or both hemispheres. This might be due to the effect size being not enough. It is also possible that, while these SNPs are associated with EI performance, they did not influence EI through affecting the anatomical features of the early visual cortical areas. A larger sample size is needed to further test the associations between these SNPs and the cortical anatomical attributes.

To our knowledge, there has been only one study exploring the heritability of EI by measuring behavioral resemblance between family members [9]. For EI underestimation, they found significant correlations in mother-father-offspring triads but nonsignificant correlations in sibling pairs. For EI overestimation, no correlation was found between family members. This result suggested the heritability of EI underestimation as well as functional dissociation between the two EI subtypes. The reason why different heritability results were observed in the current study might be attributed to the method differences. The heritability estimation in the current study was based only on common SNPs, while the family-study technique tested the contribution of heritable factors overall. It is possible that while common SNPs contributed to EI overestimation, rare SNPs contributed to EI underestimation. This possibility could be examined in future studies using other heritability estimation techniques.

So far, a variety of visual functions have been found to be influenced by genetic factors. A moderate to high degree

of heritability has been found for binocular rivalry [13, 43], biological motion [44], face recognition [45, 46], visual contour integration [14], visual contrast sensitivity [47], and so on. Some visual functions show genetic associations with certain mental disorders. For instance, shared genetic factors have been found to account for 75% of the covariance between the local biological motion processing ability and autistic traits, and 83% of the covariance between the local biological motion processing ability and communication scores [44]. Despite the strong heritability evidence, only a few studies have further explored the genetic variants underlying visual functions. To the best of our knowledge, GWAS studies have been conducted only on facial expression recognition [48, 49], bistable perception [13, 50], and visual contour integration [14]. Possibly due to the relatively strict requirement of the testing environment, the sample sizes have been relatively small. Given the importance of visual functions to daily life, their associations with many neural disorders, and the consistent findings of genetic contributions, future studies should further explore specific genetic variants underlying visual functions using the GWAS approach.

To sum up, a substantial genetic basis underlies the EI. The quantitative GWAS analysis revealed a significant heritability of EI overestimation and suggested some critical SNPs and genes associated with this EI subtype. These suggested SNPs and genes may help discover biomarkers and therapeutic targets for psychological disorders such as autism and schizophrenia. Our study provides not only evidence for the genetic basis of the EI but also valuable suggestions to improve the efficiency of GWAS studies on human cognition.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Projects 31930053, 31421003 and 31671168). We are grateful to Zhangyan Guan and Huizhen Yang for help with DNA preparation.

Compliance with ethical standards

25. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *Plos Comput Biol*. 2015;11:e1004219.
26. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421.
27. King DJ, Hodgekins J, Chouinard PA, Chouinard V-A, Sperandio I. A review of abnormalities in the perception of visual illusions in schizophrenia. *Psychonomic Bull Rev*. 2017;24:734–51.
28. Uhlhaas PJ, Silverstein SM, Phillips WA, Lovell PG. Evidence for impaired visual context processing in schizotypy with thought disorder. *Schizophrenia Res*. 2004;68:249–60.
29. Lam M, Chen C-Y, Li Z, Martin AR, Bryois J, Ma X, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet*. 2019;51:1670.
30. Li Z, Chen J, Yu H, He L, Xu Y, Zhang D, et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat Genet*. 2017;49:1576.
31. Yu HG, Yan HK, Li JM, Li ZP, Zhang XD, Ma YC, et al. Common variants on 2p16.1, 6p22.1 and 10q24.32 are associated with schizophrenia in Han Chinese population. *Mol Psychiatr*. 2017;22:954–60.
32. Needleman LA, McAllister AK. The major histocompatibility complex and autism spectrum disorder. *Dev Neurobiol*. 2012;72:1288–301.
33. Huh GS, Boulanger LM, Du H, Riquelme PA, Brotz TM, Shatz CJ. Functional requirement for class I MHC in CNS development and plasticity. *Sci (N. Y., NY)*. 2000;290:2155–9.
34. Higenell V, Ruthazer ES. Layers upon Layers: MHC Class I acts in the retina to influence thalamic segregation. *Neuron*. 2010;65:439–41.
35. Lee H, Brott BK, Kirkby LA, Adelson JD, Cheng S, Feller MB, et al. Synapse elimination and learning rules co-regulated by MHC class I H2-D-b. *Nature*. 2014;509:195.
36. Braskie MN, Jahanshad N, Stein JL, Barysheva M, Johnson K, McMahon KL, et al. Relationship of a variant in the NTRK1 gene to white matter microstructure in young adults. *J Neurosci*. 2012;32:5964–72.
37. Sawal HA, Ullah MI, Ahmad A, Nasir A, Amar A, Khan EA, et al. Homozygous mutations in NTRK1 gene underlie congenital insensitivity to pain with anhidrosis in Pakistani families. *Neurology Asia*. 2016;21:129–36.
38. Liu Z, Liu J, Liu G, Cao W, Liu S, Chen Y, et al. Phenotypic heterogeneity of intellectual disability in patients with congenital insensitivity to pain with anhidrosis: a case report and literature review. *J Int Med Res*. 2018;46:2445–57.
39. Swaminathan S, Kim S, Shen L, Risacher SL, Foroud T, Pankratz N, et al. Genomic copy number analysis in alzheimer's disease and mild cognitive impairment: an ADNI study. *Int J Alzheimer's Dis*. 2011;2011:729478.
40. Uys GM, Ramburan A, Loos B, Kinnear CJ, Korkie LJ, Mouton J, et al. Myomegalin is a novel A-kinase anchoring protein involved in the phosphorylation of cardiac myosin binding protein C. *Bmc Cell Biol*. 2011;12:18.
41. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or609998on.BJTJ0-1.17159998TD[de]iation with autistic traits. *P998 Natl Acad Sci USA*. 2018;115:1937–42.
45. Shakeshaft NG, Plomin R. Genetic specificity of face recognition. *P998 Natl Acad Sci USA*. 2015;112:12887–92.
46. Wilmer JB, Germine L, Chabris CF, Chatterjee G, Williams M, Loken E, et al. Human face recognition ability is specifically and highly heritable. *P998 Natl Acad Sci USA*. 2010;107:5238–41.
47. Haak KV. Genetic influence on contrast sensitivity in young adults. *Acta Ophthalmologica* 2019;97:E663–4.
48. Coleman JRI, Lester KJ, Keers R, Munafo MR, Breen G, Eley TC. Genome-wide association study of facial emotion recognition in children and association with polygenic risk for mental health disorders. *Am J Med Genet Part B-Neuropsychiatr Genet*. 2017;174:701–11.
49. Warrier V, Grasby KL, Uzefovsky F, Toro R, Smith P, Chakrabarti B, et al. Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Mol Psychiatr*. 2018;23:1402–9.
50. Bosten JM, Goodbourn PT, Lawrance-Owen AJ, Bargary G, Hogg RE, Mollon JD. A population study of binocular function. *Vis Res*. 2015;110:34–50.