Published in final edited form as:

Brain Cogn. 2008 December ; 68(3): 462–475. doi:10.1016/j.bandc.2008.08.027.

Does performance on the standard antisaccade task meet the co-familiarity criterion for an endophenotype?

Deborah L. Levy^{a,*}, Elizabeth A. Bowman^b, Larry Abel^b, Olga Krastoshevsky^a, Verena Krause^a, and Nancy R. Mendell^c

^a *Psychology Research Laboratory, McLean Hospital, 115 Mill Street, Belmont, MA 02478–9106, USA*

^b *Department of Optometry and Vision Sciences, University of Melbourne, Carlton, Vic., Australia*

^c *Department of Applied Math and Statistics, State University of New York at Stony Brook, Stony Brook, NY, USA*

Abstract

The “co-familiarity” criterion for an endophenotype has two requirements: (1) clinically unaffected relatives as a group should show both a shift in mean performance and an increase in variance compared with controls; (2) performance scores should be heritable. Performance on the antisaccade task is one of several candidate endophenotypes for schizophrenia. In this paper we examine whether the various measures of performance on the standard version of the antisaccade task meet the co-familiarity criterion for an endophenotype. The three measures of performance—reflexive saccade errors, latency of correct antisaccades, and gain—show a wide range of effect sizes and variance ratios as well as evidence of significant or near significant heterogeneity. The estimated mean effect sizes [Cohen's *d*: error rate: 0.34 (*SD*: 0.29); latency: 0.33 (*SD*: 0.30); gain: 0.54 (*SD*: 0.38)] are significantly greater than 0, but the magnitude of the departures from 0 is relatively small, corresponding to modest effect sizes. The width of the 95% confidence intervals for the estimated effect sizes (error rate: 0.2–0.49; latency: 0.17–0.50; gain: 0.23–0.85) and the coefficients of variation in effect sizes (error rate: 85.3%; latency: 90.9%; gain: 68.4%) reflect heterogeneity in effect sizes. The effect sizes for error rate showed statistically significant heterogeneity and those for latency ($P = .07$) and gain ($P = .09$) showed a trend toward heterogeneity. These results indicate that the effect sizes are not consistent with a single mean and that the average effect size may be a biased estimate of the magnitude of differences in performance between relatives of schizophrenics and controls. Relatives of schizophrenics show a small but significant increase in variance in error rate, but the confidence interval is broad, perhaps reflecting the heterogeneity in effect size. The variance ratios for latency and gain did not differ in relatives of schizophrenics and controls. Performance, as measured by error rate, is moderately heritable. The data do not provide compelling support for a consistent shift in mean or variance in relatives of schizophrenia patients compared with nonpsychiatric controls, both of which are required for a major gene involved in co-familial transmission. This set of findings suggests that although intra-familial resemblance in antisaccade performance is due in part to genetic factors, it may not be related to a schizophrenia genotype. Based on the current literature, it would be premature to conclude that any of the measures of antisaccade performance unambiguously meets the co-familiarity criterion for an endophenotype.

Keywords

Antisaccades; Schizophrenia; Endophenotype; Genetics; Relatives of schizophrenics; Meta-analysis; Effect size

* Corresponding author. Fax: +1 617 855 2778. E-mail address: dlevy@mclean.harvard.edu (D.L. Levy)..

1. Introduction

Gottesman and Gould (2003) define an endophenotype (see glossary) as a heritable trait that co-segregates with a disease, is state-independent, and has a significantly higher recurrence risk in unaffected relatives of individuals with the disease than in the general population. A primary rationale for incorporating endophenotypes into linkage studies (see glossary) is to improve power to detect common susceptibility alleles (Holzman & Matthyse, 1990; Matthyse & Parnas, 1992). This improvement in power would occur because more accurate identification of non-penetrant gene carriers through a segregating endophenotype lowers the risk of false negatives (Botstein & Risch, 2003; Matthyse & Parnas, 1992). Since endophenotypes also have a much higher recurrence than the disease in families with an affected individual (Gottesman & Gould, 2003), a pleiotropic (see glossary) locus that is one of several for a complex disease would have a higher probability of being detected through its effect on a more penetrant endophenotype than through its effect on a less penetrant disease phenotype (Freedman, Adler, & Leonard, 1999; Holzman, 1994; Holzman & Matthyse, 1990; Lander, 1988; Matthyse & Parnas, 1992). Endophenotypes are thought: (1) to be potentially closer to the biological actions of a single gene; (2) to have a simpler genetic architecture than complex diseases (i.e., are more likely to be monogenic); (3) to be useful for clarifying the neurobiological effects of genes and the pathophysiology of the disease; and (4) to provide a method for subtyping potentially more homogeneous forms of a clinical disorder (Braff, Freedman, Schork, & Gottesman, 2007; Carlson, Eberle, Kruglyak, & Nickerson, 2004; Freedman et al., 1999; Gottesman & Gould, 2003; Holzman & Matthyse, 1990; Jablensky, 2006; Matthyse & Parnas, 1992; Szatmari et al., 2007). For more extensive discussions of endophenotypes, their value in genetic studies of schizophrenia, and their various definitions, (see also Adler, Freedman, Ross, Olincy, & Waldo, 1999; Almasy & Blangero, 2001; Bearden & Freimer, 2006; Calkins et al., this issue; Adler, Freedman, Ross, Olincy, & Waldo, 1999; Almasy & Blangero, 2001; Bearden & Freimer, 2006; Flint & Munafo, 2006; Glahn et al., 2007; Greenberg, 1992; Iacono, 1998; Keating et al., 1991; Keri & Janka, 2004; Leboyer et al., 1988; Meyer-Lindenberg & Weinberger, 2006; Thaker, 2008; Turetsky et al., 2007; Walters & Owen, 2007).

The “co-familiality” criterion refers to whether a potential endophenotype is over-represented in clinically unaffected relatives of probands compared with nonpsychiatric controls. This over-representation can be reflected in a significant shift in the mean value for a quantitative trait or in a significant difference in the frequency of a qualitative trait compared with controls. Since first-degree relatives as a group are a mixture of gene carriers and non-gene carriers, the endophenotype would also be expected to have a larger variance in clinically unaffected relatives than in controls. The reason is that the subgroup of relatives with deviant performance, the putative gene carriers, would increase the variance for the group of relatives as a whole.

Genes may or may not underlie the familial aggregation of a trait. As Seymour Kety was fond of noting, many things run in families—the silver, politics, religion—that do not have a genetic basis. For this reason it is also important to demonstrate that the trait thought to be an endophenotype shows familial aggregation because it is heritable. In addition, a trait may aggregate in families and be heritable, but not be related to risk for a disease genotype. In that case, the trait would show a significant within-family correlation but relatives of affected individuals would not show a shift in mean or an increase in variance.

A number of candidate endophenotypes for schizophrenia have been provisionally identified. These include eye tracking dysfunction (Levy, Holzman, Matthyse, & Mendell, 1993), decreased P50 inhibition (Freedman et al., 1999), reduced P300 amplitude (Hall et al., 2007),

reflexive saccade errors on the antisaccade (AS) task (Levy, Mendell, & Holzman, 2004a), and a variety of neurocognitive measures (Hallmayer et al., 2003).

Poor performance on the antisaccade (AS) task is so strongly associated with schizophrenia that there has never been a failure to replicate (see section by Gooding et al., this issue). The strength of this association in a wide range of schizophrenia populations (acutely ill, remitted, first-episode, chronic) made it logical to assess whether poor performance on the AS task is also over-represented in clinically unaffected relatives of schizophrenics (RelSZ). Some studies (see below) have reported that unaffected RelSZ perform the AS task significantly worse than controls, a difference that has been interpreted as support for considering this trait to be a schizophrenia endophenotype. Other studies do not support the familial aggregation of poor performance in RelSZ or the usefulness of performance on this task as an endophenotype (see below).

In this paper we examine the evidence as to whether the various measures of AS performance meet the co-familiality criterion for an endophenotype in biological RelSZ.

If AS performance taps processes related to genetic vulnerability for schizophrenia, it must fulfill at least two criteria. First, genetic effects should contribute to variability in performance (i.e., it should be heritable). Second, clinically unaffected RelSZ as a group should show both a shift in mean performance and an increase in variance compared with controls that is independent of poor performance in RelSZ with diagnoses of schizophrenia, psychosis or schizophrenia spectrum personality disorders. The magnitude of this shift in mean and variance bears on how potentially useful an endophenotype is for identifying non-penetrant gene carriers. Evaluating the broader question of the endophenotype status of AS performance would be premature until the co-familiality criterion has been satisfactorily established.

2. The antisaccade task

The AS task requires a subject to inhibit a visually guided saccade to a novel peripheral target, and instead, to generate a voluntary eye movement to the mirror image location in the contralateral visual field (Hallett, 1978). This voluntary eye movement is the antisaccade. Performance is typically assessed using one or more of several dependent measures. The most widely used measure of performance is the proportion of pro-saccades *to* the novel peripheral target instead of *away* from it, the error rate. This score is thought to reflect a failure to inhibit prepotent responses to the target (see also (Massen, 2004) for a different formulation) and is commonly interpreted to reflect an impairment in prefrontal cortical functioning (for a more extensive discussion of this interpretation, see (Levy et al., 2004b)). Other measures of performance include latency (i.e., oculomotor reaction time) of correct and incorrect antisaccades and gain. Gain, sometimes also called accuracy, compares the amplitude of a correct antisaccade to the amplitude of the distance between central fixation and the mirror image location.

In the most commonly used paradigm in studies of schizophrenia patients and their relatives, the standard AS task, the offset of the central fixation point occurs simultaneously with the onset of the peripheral target. Other variations include (1) the overlap paradigm (see glossary), in which the central fixation point remains illuminated for a short period of time after the peripheral target appears, and (2) the gap paradigm (see glossary), in which the central fixation point is extinguished shortly before the appearance of the novel peripheral target. Timing parameters and the number of eccentricities (e.g., one at $\pm 10^\circ$ versus several at $\pm 5^\circ$, 10° , 15°) vary. In addition, sometimes the correct location for the antisaccade is illuminated at the end of each trial and sometimes it is not.

3. Family studies of antisaccade performance

Based on a PubMed search of the literature, we identified 21 family studies of AS performance to date. Table 1 lists, in chronological order, the studies, their sample sizes, inclusion/exclusion criteria, type of AS paradigm used, and dependent measures. Primary findings with respect to RelSZ-nonpsychiatric control (NC) or sibling of schizophrenic (SibSZ)-NC comparisons are also briefly summarized in the table. Effect sizes, using Cohen's d (Cohen, 1977, 1988), r (Rosenthal & Rosnow, 1984), and Glass' delta (Glass, 1976), are included. Cohen's d assumes equal variances [$d = (\text{mean}_{\text{RelSZ}} - \text{mean}_{\text{NC}}) / \text{pooled standard deviation}$] and Glass' delta [$(\text{mean}_{\text{RelSZ}} - \text{mean}_{\text{NC}}) / \text{standard deviation}_{\text{NC}}$] does not. When groups differ in variance, which would be expected if RelSZ are a mixture of gene carriers and non-gene carriers, Cohen's d will underestimate effect size relative to Glass' delta. The effect size, r , is the correlation between group membership and performance.

Seventeen of these 21 studies used a standard AS paradigm (Boudet et al., 2005; Brownstein et al., 2003; Clementz, McDowell, & Zisook, 1994; Crawford, Sharma, Puri, Murray, & Lewis, 1998; Curtis, Calkins, Grove, Feil, & Iacono, 2001a; Curtis, Calkins, & Iacono, 2001b; de Wilde, Bour, Dingemans, Boeree, & Linszen, 2008; Ettinger et al., 2004, 2006; Karoumi et al., 2001; Katsanis, Kortenkamp, Iacono, & Grove, 1997; Louchart-de la Chapelle et al., 2005; MacCabe et al., 2005; McDowell & Clementz, 1997; Price et al., 2005; Thaker, Cassady, Adami, Moran, & Ross, 1996; Thaker et al., 2000), sixteen of which involved independent samples (subjects included in the standard paradigm results reported in the Curtis et al. (2001b) study were a subset of those from the Curtis et al. (2001a) study data from the former study are not included in the discussion that follows). Two other studies used a gap paradigm only (Raemaekers, Ramsey, Vink, van den Heuvel, & Kahn, 2006; Ross et al., 1998). Two studies used an overlap paradigm only (McDowell, Myles-Worsley, Coon, Byerley, & Clementz, 1999; Petrovsky et al., 2008). Two of the studies that used a standard paradigm also used an overlap paradigm (Curtis et al., 2001b; McDowell & Clementz, 1997). Since the most data are available from the sixteen independent studies that used a standard paradigm, the discussion here will emphasize results from those studies.

Below we address the following issues: (1) Is performance on the standard AS paradigm heritable, and on what measures? (2) Do RelSZ show a shift in mean and an increase in variance on the various measures of AS performance compared with nonpsychiatric controls?

3.1. Is performance heritable?

Two studies have addressed the heritability of performance on the AS task. Both studies support the conclusion that genetic effects contribute to variability in error rate. No studies have examined the heritability of latency and gain, which have been used to measure performance less often than error rate. Malone and Iacono (2002) carried out a large twin study on performance on the standard AS task in a community-based epidemiological sample. They assessed AS errors in 207 11-year old and 203 17-year-old female monozygotic (MZ) twin pairs and 126 11-year old and 102 17-year-old dizygotic (DZ) twin pairs. The proportions of variance in error rate attributable to genetic and shared and specific environmental effects were estimated by fitting biometric models to the observed data. Variance in error rate attributable to additive genes was 57% in both age groups combined, despite a significant effect of age on error rate (effect size for age effect: 0.81). In both age cohorts of twins, the intra-class correlations (ICC) were statistically significant (11-year old MZ: ICC = 0.54; 11-year old DZ: ICC = 0.29; 17-year old MZ: ICC = 0.62; 17-year old DZ: ICC = 0.31), reflecting substantial phenotypic similarity within both types of twin pairs. Degree of resemblance was significantly higher in MZ twins than in DZ twins, consistent with genetic involvement in performance. The proportion of the variance accounted for by non-shared environment was 43% for both age groups combined. Shared environment did not contribute to phenotypic similarity. Of note,

Malone and Iacono commented on the magnitude of individual differences in performance: "... standard deviations in the two [age] cohorts were at least as great as the magnitude of the difference in mean error rate between the groups." (p. 670).

The results recently reported by Greenwood et al. (2007) for the overlap AS paradigm are largely consistent with those of Malone and Iacono (2002). They obtained an estimated heritability of 0.42 (*SE*: 0.08; 95% confidence interval: 0.27–0.57) for error rate in a sample of 525 members of schizophrenia pedigrees. The authors noted that shared environment could have impacted the heritability estimates, but the Malone and Iacono results suggest that such an effect would not be substantial. The results of these two studies provide consistent evidence that a sizeable proportion of the variance in error rate on the standard and overlap AS paradigms is heritable.¹

These data supporting moderate levels of heritability are consistent with other findings showing significant but modest within-family correlations in performance (Brownstein et al., 2003; Crawford et al., 1998; Curtis et al., 2001a; McDowell & Clementz, 1997), although not all studies have reported significant intra-familial association (Ettinger et al., 2004; MacCabe et al., 2005). The key question is what heritable familial resemblance has to do with a schizophrenia genotype (see (Brownstein et al., 2003) for a more extensive discussion of this distinction). The heritability estimates for AS error rate described above, like heritability estimates for schizophrenia, do not necessarily imply that the mode of inheritance involves a major gene. A polygenic trait, for example, can have high heritability.

3.2. Do relatives of schizophrenia patients show the shift in mean performance and increase in variance compared with nonpsychiatric controls that would be expected from a genetic model involving co-familial transmission?

3.2.1. Error rate scores—The error rate measure of AS performance was used in every study of RelSZ. Table 2 presents means, standard deviations and effect sizes (ES) for the sixteen independent studies that used the standard paradigm. Five studies reported that RelSZ made statistically significantly more errors on the AS task than NC did (Clementz et al., 1994; Curtis et al., 2001a; Karoumi et al., 2001; Katsanis et al., 1997; McDowell & Clementz, 1997). Effect sizes in these studies ranged from 0.50–0.78 for Cohen's *d* and from 0.635–1.03 for Glass' delta. Eleven studies reported that RelSZ did not make significantly more errors than NC did (Boudet et al., 2005; Brownstein et al., 2003; Crawford et al., 1998; de Wilde et al., 2008; Ettinger et al., 2004, 2006; Louchart-de la Chapelle et al., 2005; MacCabe et al., 2005; Price et al., 2005; Thaker et al., 1996, 2000). Effect sizes in these eleven studies ranged from –0.13 to 0.60 for Cohen's *d* and 0.10 to 0.80 for Glass' delta. In three of these studies the difference between groups was at the trend level for statistical significance: $P < .08$, ES: 0.60/0.80 (Ettinger et al., 2004); $P < .074$, ES: 0.28/0.30 (Price et al., 2005); $P < .08$, ES: 0.49/0.55 (de Wilde et al., 2008). In three other studies the difference between RelSZ and NC was in the wrong direction, resulting in negative effect sizes.

The design of some family studies has been predicated on the rationale that multiplex families, which are thought to have an increased genetic loading for schizophrenia, would be expected to show a higher rate of deviance in performance on a putative endophenotype than simplex families. Comparisons of simplex and multiplex families have been inconsistent. Two studies

¹Ettinger et al. (2006) also examined performance on an AS task in eight schizophrenic probands and nine of their clinically discordant MZ co-twins. The comparison group consisted of 10 pairs of MZ twins with no Axis I disorder and no family history of psychosis in first- or second-degree relatives. No intra-class correlations for the discordant schizophrenia pairs or the control twins were presented, conceivably because the small sample sizes (which are not unusual in scarce populations such as twins discordant for schizophrenia) would have resulted in broad confidence intervals. Thus, it is more appropriate to treat this study as a family study than as a twin study *per se*.

of multiplex families that used a standard AS paradigm reported either small or negative effect sizes (Crawford et al., 1998; MacCabe et al., 2005). One of those studies included both simplex and multiplex families (MacCabe et al., 2005). Error rate effect sizes were *larger* in simplex families than in multiplex families. The effect sizes for multiplex families were the most strongly negative ($-.046/-0.41$) to date and the effect sizes for simplex families were weakly positive ($0.15/0.14$). Another study of multiplex families showed a small effect size for the error rate score (Crawford et al., 1998). In contrast, a study that used the overlap paradigm found that two samples of multiplex families had substantially *larger* (2.5–3 times larger) effect sizes than was found in a sample of simplex families. Far-overlap targets maximized separation of the groups (McDowell et al., 1999). Another way to examine the effect of genetic loading is to compare the performance of most-likely gene carrier (MLC) parents and least-likely gene carrier (LLC) parents. Using the gap AS paradigm, Ross et al. (1998) compared small samples of most-likely gene carrier (MLC) parents of schizophrenics with least-likely gene carrier (LLC) parents. MLC parents, but not LLC parents, made significantly more errors than controls. Using an overlap paradigm, Petrovsky et al. (2008) reported that both the small group of MLC parents and the LLC parents had significantly higher error rates than controls. The data are thus mixed as to whether clinically unaffected members of multiplex families or individuals at greater genetic risk by virtue of a family history of psychosis have an increased probability of making more errors on the AS task. It is useful to point out that the discussion above has focused on the probability of deviance on the AS task given genetic loading, not on the probability of schizophrenia-related illness given AS deviance.

It is possible that particular features of the AS paradigm are critical for eliciting group differences. For example, the far-overlap condition or the gap condition may reveal an effect of genetic loading that is not apparent with the standard paradigm. It is important to note, however, that the expectation of higher recurrence of poor AS performance in multiplex families would hold only if the same gene that increases liability for the endophenotype also increases risk for schizophrenia in those families. If the same gene that contributes to AS performance does not contribute to susceptibility for schizophrenia, for example, there would be no reason to expect multiplex families to differ from simplex families. Relevant in this regard, suggestive evidence of linkage to a marker on 22q11–12 has been reported for a composite phenotype consisting of performance on the far-overlap task and P50 sensory gating in multiplex families (Myles-Worsley et al., 1999). COMT is located in this region. Variation in the valine¹⁵⁸ methionine genotype was not significantly associated with AS error rate in one population sample (Stefanis et al., 2004) but was related to performance in another sample of patients and controls (Haraldsson et al., 2008) and to BOLD response (Ettinger et al., 2008).

As Table 2 shows, effect sizes for a difference between RelSZ and NC are quite variable and cover a very broad range. The mean for Cohen's d is 0.34 (SD : 0.29) (95% confidence interval: 0.2–0.49) and the median is 0.36. The mean for Glass' delta is 0.43 (SD : 0.39) (95% confidence interval: 0.24–0.63) and the median is 0.39. The mean for r is 0.16 (SD : 0.14) (95% confidence interval: 0.10–0.23) and the median is 0.17. One sample t -tests on the effect size values were all statistically significant (Cohen's d : $t_{(15)} = 4.7$, $P = .0003$; r : $t_{(15)} = 4.8$, $P = .0003$; Glass' delta: $t_{(15)} = 4.4$, $P = .0005$), indicating that the mean values for each measure of effect size are significantly greater than 0, although on average, the magnitude of the departure of the estimated effect size from 0 is relatively small and corresponds to a modest effect size using Cohen's criteria (Cohen, 1988).

We calculated whether the effect sizes were consistent with values drawn from a single distribution using heterogeneity analysis (Rosenthal & Rubin, 1982). The results showed statistically significant heterogeneity ($\chi^2_{15} = 27.99$, $P = .022$), indicating that the effect sizes are not consistent with a single mean. The average effect size may thus not be a meaningful indicator of the magnitude of the difference in mean error rate between RelSZ and NC. The

width of the confidence intervals, which is inversely related to the number of studies, also reflects uncertainty about the estimated mean effect size. Consistent with this interpretation, the coefficient of variation for the error rate effect sizes is 85.3%, indicating that the standard deviation of the effect sizes is almost as large as the mean effect size. The actual mean effect size may thus be smaller or larger than the average of these effect sizes. The average of these effect sizes may also be an overestimate because of publication bias favoring positive findings. Factors that may contribute to this heterogeneity are discussed below.

Table 3 presents the variance ratios for the fifteen standard paradigm studies that provided means and standard deviations. Under the null hypothesis of equal variance, the variance ratio is distributed as F and the mean value of $F = 1.0$. The variances of the error scores of RelSZ were significantly larger than those of controls ($t_{(14)} = 1.87, P = .04$ one sided; Wilcoxon Signed Rank Test, $S = 33, P = .03$, one-sided). On average, RelSZ had a variance that was 32% larger than that of NC. The 90% confidence interval ranges from 0.016 to 0.544, indicating that the mean variance ratio could be as low as 1.02 (indicating a 2% increase in variance) or as high as 1.72 (indicating a 72% increase in variance) in RelSZ relative to controls. A 32% mean increase in variance corresponds to a 15% mean increase in standard deviation.

The reason(s) for this variability in effect size and variance ratio are uncertain, but are essential for clarifying the value of the AS task in genetic studies. We did not try to identify those factors that contribute to the heterogeneity of effect sizes. In a previous meta-analysis on the nine published family studies at that time (Levy et al., 2004a), the one factor that was consistent with the variability in effect size was the symmetry/asymmetry of the inclusion/exclusion criteria applied to RelSZ and NC. Studies that obtained large effect sizes used asymmetric exclusion criteria in that more restrictive inclusion/exclusion criteria were applied to controls than to RelSZ. Effect sizes were significantly smaller in studies that applied comparable inclusion/exclusion criteria to both groups. In contrast, neither differences in number of trials nor number of target eccentricities contributed significantly to variability in effect size.

Although the results of that meta-analysis indicated that symmetrical/asymmetrical inclusion/exclusion criteria may be *sufficient* to explain variability in effect size when studies are considered as a group, data from individual studies suggest that asymmetrical inclusion/exclusion criteria may not be *necessary* for obtaining large effect sizes. In two studies whose inclusion/exclusion criteria were asymmetrical, RelSZ performed significantly worse than NC independent of co-morbid psychopathology even when more equivalent inclusion/exclusion criteria were subsequently applied to the two groups (Calkins, Curtis, Iacono, & Grove, 2004; Curtis et al., 2001a; Katsanis et al., 1997).

Taken together, the results for error rate support three general conclusions regarding the standard AS paradigm. First, as discussed in the previous section, heritability contributes to variability in performance. Second, the estimated mean effect size for a difference between RelSZ and NC is modest and the coefficient of variation is large. In addition, there is significant heterogeneity in effect size. Third, on average, RelSZ show a small increase in variance, but the confidence interval is very broad and may reflect the heterogeneity in effect sizes. These results do not provide strong support for the inference that error rate on the standard AS paradigm would be a strong indicator of gene carrier status in RelSZ or would increase the power of linkage analysis of a multidimensional phenotype.

3.2.2. Latency—Latency of correct AS is the next most widely used measure of performance. Table 4 presents means, standard deviations and effect sizes for the fourteen independent studies that used the standard paradigm. Three studies reported significantly increased AS latency in all or a subset of RelSZ (Ettinger et al., 2006; Thaker et al., 1996, 2000). In one report RelSZ *without* schizophrenia spectrum personality symptoms (SSPS) accounted for the

difference in latency between RelSZ and NC (Thaker et al., 1996); in another report from the same group, only SSPS RelSZ had significantly increased latency (Thaker et al., 2000) (see also section below on effects of schizotypal symptoms). Effect sizes in these three studies ranged from 0 to 0.67 (Thaker et al., 1996), 0.2 to 1.6 (Thaker et al., 2000), and 0.95 to 1.03 (Ettinger et al., 2006), respectively (see Table 1). Eleven studies did not find significant differences in latency between RelSZ and NC (Boudet et al., 2005; Brownstein et al., 2003; Crawford et al., 1998; Curtis et al., 2001a; de Wilde et al., 2008; Ettinger et al., 2004; Karoumi et al., 2001; Katsanis et al., 1997; Lou-chart-de la Chapelle et al., 2005; MacCabe et al., 2005; McDowell & Clementz, 1997). Effect sizes ranged from -0.08 to 0.81 for Cohen's d and from -0.10 to 0.98 for Glass' delta.

As Table 4 shows, effect sizes for a difference in latency between RelSZ and NC are quite variable and cover a very broad range, as was the case for the error rate score. The mean for Cohen's d is 0.33 ($SD: 0.30$) (95% confidence interval: $0.17-0.50$) and the median is 0.36 . The mean for Glass' delta is 0.37 ($SD: 0.35$) (95% confidence interval: $0.18-0.55$) and the median is 0.36 . The mean for r is $.16$ ($SD: 0.14$) (95% confidence interval: $0.08-0.24$) and the median is 0.17 . One sample t -tests on the effect size values were all statistically significant (Cohen's d : $t_{(13)} = 4.1$, $P = .0012$; r : $t_{(13)} = 4.2$, $P = .001$; Glass' delta: $t_{(13)} = 4.0$, $P = .002$), indicating that the mean values for each measure of effect size are significantly greater than 0, although as with error rate, on average, the magnitude of the departure of the estimated mean effect size from 0 is relatively small and corresponds to a small effect size using Cohen's criteria (Cohen, 1988).

The results of the heterogeneity analysis show a trend toward significant heterogeneity ($\chi^2_{13} = 20.99$, $P = .07$), indicating that the effect sizes may not be consistent with a single mean. The mean effect size thus may not be interpretable as a reflection of the magnitude of the difference in average latency between RelSZ and NC. The coefficient of variation for the latency effect sizes is 90.9%, which is consistent with evidence in favor of heterogeneity. As with the error rate score, the actual mean effect size for latency may be smaller or larger than the average of these effect sizes. It is not possible to rule out the possibility that the mean of these effect sizes may be an overestimate because of publication bias favoring positive findings. The heterogeneity analysis did not reach statistical significance, but the result does not provide strong evidence of homogeneity of effect sizes for latency.

Table 5 presents the variance ratios for the thirteen standard AS paradigm studies that provided means and standard deviations. The variances of the latency scores of RelSZ were not significantly larger than those of controls (Signed Rank $S = 10.5$, $P = .25$ one-sided).

These data do not support a consistently moderate-large difference in AS latency or an increase in variance between RelSZ and nonpsychiatric controls, although there could be a small difference in mean latency.

3.2.3. Gain—Fewer studies have examined AS gain, or the accuracy of AS to the mirror image location, than the other two measures of performance. Table 6 presents means, standard deviations and effect sizes for the six independent studies that used the standard paradigm. Three studies reported significantly reduced gain in RelSZ or SibSZ compared with NC (Ettinger et al., 2004, 2006; Karoumi et al., 2001), with effect sizes ranging from 0.68 to 1.0 . One study reported significantly increased gain in RelSZ (de Wilde et al., 2008) with an effect size of $0.58/0.73$. Two studies reported no significant group difference (Crawford et al., 1998; McDowell & Clementz, 1997) and effect sizes of -0.06 and $0.26/0.28$, respectively.

As Table 6 shows, effect sizes for a difference in gain between RelSZ and NC are quite variable and cover a very broad range. The mean for Cohen's d is 0.54 ($SD: 0.38$) (95% confidence

interval: 0.23–0.85) and the median is 0.63. The mean for Glass' delta is 0.57 (SD : 0.39) (95% confidence interval: 0.25–0.89) and the median is 0.71. The mean for r is .25 (SD : 0.17) (95% confidence interval: 0.11–0.39) and the median is 0.30. One sample t -tests on the effect size values were all statistically significant (Cohen's d : $t_{(5)} = 3.47$, $P = .002$; r : $t_{(5)} = 3.6$, $P = .02$; Glass' delta: $t_{(5)} = 3.57$, $P = .02$), indicating that the mean values for each measure of effect size are significantly greater than 0, although as with error rate and latency, on average the magnitude of the departure of the estimated mean effect size from 0 for gain is relatively small and corresponds to a medium effect size using Cohen's criteria (Cohen, 1988).

The results of the heterogeneity analysis show a trend toward significant heterogeneity ($\chi^2_5 = 9.66$, $P = .09$), indicating that the effect sizes may not be consistent with a single mean. The mean effect size may not be interpretable as a reflection of the magnitude of the difference in average gain between RelSZ and NC. Another reflection of the heterogeneity in effect size for gain is the large coefficient of variation, 68.4%, a result that does not provide strong evidence of homogeneity of effect sizes for gain. Indeed, the confidence intervals are broadest for this variable, reflecting uncertainty about the mean effect size for these six studies.

Table 7 presents the variance ratios for the six standard paradigm studies that provided means and standard deviations. The variances of the gain scores of RelSZ were not significantly different from those of controls (Signed Rank $S = 9.5$, $P = .34$, one-sided).

These data may support a moderate difference in mean AS gain between RelSZ and nonpsychiatric controls but do not support an increase in variance in RelSZ. In light of the small number of studies that used gain as a dependent measure, the results for this variable, in particular, should be interpreted cautiously.

Three aspects of the results on gain are worth noting. First, the effect sizes for reduced gain (hypometria, see glossary) were larger for SibSZ and for discordant MZ co-twins compared with NC than for SZ patients compared with NC in both Ettinger et al. studies (Ettinger et al., 2004, 2006) (2004—SibSZ vs NC: 0.68/0.69; SZ vs NC: 0.44/0.58; 2006—co-twins vs NC: 1.0/1.01; SZ twins vs NC: 0.62/0.72). Similarly, de Wilde et al. (2008) reported a larger effect size for increased gain (hypermetria, see glossary) in the comparison of SibSZ with NC, 0.58/0.73, than in the comparison of SZ and NC, 0.24/0.29. Moreover, the effect size for hypermetria in SibSZ compared with SZ, 0.71/0.70, was larger than the effect size for the comparison of patients or siblings with NC. Larger effects in siblings than in patients would not be expected for a co-familial trait that would serve as a useful endophenotype. Also casting doubt on the usefulness of the gain measure as a measure of co-familiality, the intra-class correlation for familial resemblance was 0.08 (de Wilde et al., 2008). Second, some studies report near perfect AS accuracy in controls. Controls in the Ettinger et al. sibling study (Ettinger et al., 2004) had a mean AS gain of 0.996 to two target locations ($\pm 15^\circ$). The Ettinger et al. twin study (Ettinger et al., 2006) reported a mean gain in controls of 1.0 to four target locations ($\pm 6^\circ$, 12°). Karoumi et al. also reported near perfect AS accuracy in controls (0.98) for six target locations ($\pm 5^\circ$, 15° , 25°) (Karoumi et al., 2001). These are remarkable levels of precision to non-visible targets. Third, several studies reported that controls showed *better* accuracy in making antisaccades—in the absence of a visual target—than for visually guided saccades. The mean gain for pro-saccades in controls was 0.94 in the Ettinger et al. sibling study (Ettinger et al., 2004) and 0.96 in the Ettinger et al. twin study (Ettinger et al., 2006), both of which were *lower* than the mean gains for antisaccades in the same groups (see above). Mean accuracy of pro-saccades and AS in controls were essentially equivalent in the Karoumi et al. study (Karoumi et al., 2001).

The estimated effect sizes for pairs of measures of performance are graphed in Figs. 1–3. None of the correlations was statistically significant, although the maximum number of studies

considered here was 14. The modest positive estimated correlation between latency and gain (Fig. 3a; $r = .39$, $P = .44$) is driven by one outlier value; the estimated correlation obtained between effect sizes for these two variables is modestly negative when the outlier value is removed (Fig. 3b; $r = .43$, $P = .47$).

4. Effects of schizotypal symptoms

If schizotypal symptoms in RelSZ increase the probability of being a carrier for schizophrenia genes and poor performance on the AS task is an alternate expression of one or more of these genes, the probability of poor AS performance could be increased in RelSZ with schizotypal symptoms compared to those without schizotypal symptoms. Only two studies have dichotomized subjects in terms of the presence or absence of schizotypal symptoms (SSPS), although a number of samples in other studies included a few schizotypal RelSZ. Thaker et al. (1996) did not find a statistically significant difference in error rate between RelSZ with or without SSPS compared with community controls (with or without SSPS), but the effect size for the comparison of RelSZ with SSPS vs NC with SSPS (0.41/0.37) was substantially larger than the effect size for the comparison of RelSZ without SSPS vs NC without SSPS (0.02/0.02). A similar pattern was found for the error rate effect sizes in a larger subsequent study (Thaker et al., 2000). RelSZ with SSPS had a larger effect size (0.68/0.75) than RelSZ without SSPS (0.15/0.17). Thus, both studies showed moderately large effect sizes for increased error rate in RelSZ with SSPS and very small effect sizes in RelSZ without SSPS.

The results for latency show an inconsistent pattern across the two studies. In the first Thaker et al. (Thaker et al., 1996) study, latency was significantly increased in all RelSZ compared with NC, although the ES were modest (0.38/0.39). However, the overall group difference was driven by the *non-SSPS* RelSZ group; the effect size for a difference between SSPS RelSZ and SSPS NC was 0, whereas the contrast between *non-SSPS* RelSZ and *non-SSPS* NC yielded an effect size of 0.66/0.67. In the later study (Thaker et al., 2000), the pattern for latency paralleled the pattern observed for error rate in the same study, but the effect size was substantially larger. SSPS RelSZ were slower to initiate correct AS than SSPS NC (0.90/1.61), whereas the contrast between *non-SSPS* RelSZ and *non-SSPS* NC produced a small effect (0.21/0.19).

Taken together, there seems to be some association between SSPS in RelSZ and poor performance on the AS task that is more strongly seen in slowed latency than in increased errors, but the results are not consistent. Although three of the discordant MZ co-twins in the Ettinger et al. (2006) sample and three RelSZ in the Crawford et al (Crawford et al., 1998) sample were diagnosed with schizotypal personality disorder, the relation between performance and this diagnosis was not addressed in either study. In the Clementz et al. (1994) study, three RelSZ had schizotypal personality disorder and one had a diagnosis of schizophrenia. The scatter plot in Fig. 3a and b of that paper makes clear that a schizophrenia spectrum disorder is not a necessary condition for poor performance even though one of these four RelSZ had an error score almost three times the standard deviation of the RelSZ group.

5. Conclusions

For the standard AS task, the estimated effect size for each of the three measures of performance was significantly greater than 0. For each variable, however, the magnitude of the departure of the estimated effect size from 0 is relatively small. Each dependent measure shows a wide range of effect sizes and variance ratios as well as evidence of significant or near significant heterogeneity. The coefficients of variation for the three measures of performance were large (all >65%), reflecting unclarity about the magnitude of the actual effect sizes and (given the probability of publication bias favoring positive effect sizes) the representativeness of the

average effect sizes as estimates of mean effect sizes. These data do not provide compelling support for a consistent shift in mean or variance in RelSZ relative to nonpsychiatric controls. Error rate on both the standard and far-overlap AS task shows moderate heritability. This combination of findings suggests that intra-familial resemblance on the AS task may well be influenced by genetic factors, but may not be related to schizophrenia susceptibility genes. Based on the current literature, it would be premature to conclude that any of the measures of performance on the standard AS task unambiguously or strongly meets the co-familiality criterion for an endophenotype. Thus, it is unlikely that performance on this task would increase power to detect linkage by improving detection of non-penetrant gene carriers.

The principal support for meeting the co-familiality criterion must be derived independent of schizophrenia-related psychopathology in RelSZ. In order to demonstrate convincingly that RelSZ show both a shift in mean and an increase in variance relative to controls, symmetrically ascertained samples are essential. This is the only way to establish that increased errors identify that subset of non-penetrant gene carriers who would be classified as unaffected based on a clinical phenotype alone. It is also the only way to estimate the magnitude of the improvement in gene carrier detection associated with an elevated error rate. Since psychotic relatives are likely to perform poorly (see (Curtis et al., 2001a; MacCabe et al., 2005)), their inclusion can be expected to increase the mean and variance of the RelSZ group. There is no reason not to test RelSZ who are psychotic or who have schizophrenia-related personality disorders on the AS task, but it is the performance of RelSZ who would be considered clinically unaffected that is most germane to fulfilling the co-familiality criterion.

Fewer studies used the overlap and gap versions of the AS task than used the standard paradigm. The far-overlap AS task has been shown to yield the largest effects sizes in distinguishing RelSZ from NC. Whether this version of the AS task fulfills the co-familiality criterion for an endophenotype remains unclear. Data on the far overlap paradigm are available from only two studies of three samples, all from a single laboratory. A previous meta-analysis found significant heterogeneity for both the near-overlap and far-overlap paradigms (Levy et al., 2004a). We purposefully did not combine studies that used different AS paradigms for two reasons. First, heterogeneity of effect size is present within paradigm (for the standard task and for the overlap task; for the gap task, there are too few studies for a heterogeneity analysis to be meaningful) (Levy et al., 2004a). Second, observed effect sizes are correlated in most studies that used the overlap task due to non-independence of the control group.

The reasons for the variability in effect size and variance ratio across studies are not clear. As mentioned above, the symmetry/asymmetry of inclusion criteria may be sufficient but not necessary to account for large effect sizes. A number of other aspects of experimental design may impact the magnitude of group differences. For example, the absence of mention of a blind in the diagnostic assessments, during data acquisition, and during data processing is conspicuous in most of this literature. In addition, a standard pro-saccade task is routinely administered before the AS task, which is equivalent to priming subjects to do exactly the opposite of what is required for correct performance on the AS task. Another unexplored potential source of variability is whether or not subjects received feedback about their performance by having a target appear at the correct response location at the end of each trial. Routine inclusion of scatter plots or histograms of data points would be one way to facilitate thorough and constructive evaluation of this literature.

Acknowledgments

This study was supported in part by NIMH Grants R01 MH071523 and MH31340, the Sidney R. Baer Jr. Foundation, the Essel Foundation, and the National Association for Research on Schizophrenia and Depression. The authors thank Dr. Mark Lenzenweger for critical comments and helpful discussion.

References

- Adler LE, Freedman R, Ross RG, Olincy A, Waldo MC. Elementary phenotypes in the neurobiological and genetic study of schizophrenia. *Biological Psychiatry* 1999;46:8–18. [PubMed: 10394470]
- Almasy L, Blangero J. Endophenotypes as quantitative risk factors for psychiatric disease: Rationale and study design. *American Journal of Medical Genetics* 2001;105:42–44. [PubMed: 11424994]
- Bearden CE, Freimer NB. Endophenotypes for psychiatric disorders: Ready for primetime? *Trends in Genetics* 2006;22:306–313. [PubMed: 16697071]
- Botstein D, Risch N. Discovering genotypes underlying human phenotypes: Past successes for Mendelian disease, future approaches for complex disease. *Nature Genetics* 2003;33(Suppl):228–237. [PubMed: 12610532]
- Boudet C, Bocca ML, Chabot B, Delamillieure P, Brazo P, Denise P, et al. Are eye-movement abnormalities indicators of genetic vulnerability to schizophrenia? *European Journal of Psychiatry* 2005;20:339–345.
- Braff DL, Freedman R, Schork NJ, Gottesman II. Deconstructing schizophrenia: An overview of the use of endophenotypes in order to understand a complex disorder. *Schizophrenia Bulletin* 2007;33:21–32. [PubMed: 17088422]
- Brownstein J, Krastoshevsky O, McCollum C, Kundamal S, Matthyse S, Holzman PS, et al. Antisaccade performance is abnormal in schizophrenia patients but not in their biological relatives. *Schizophrenia Research* 2003;63:13–25. [PubMed: 12892854]
- Calkins ME, Curtis CE, Iacono WG, Grove WM. Antisaccade performance is impaired in medically and psychiatrically healthy biological relatives of schizophrenia patients. *Schizophrenia Research* 2004;71:167–178. [PubMed: 15374584]
- Carlson CS, Eberle M, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. *Nature* 2004;429:446–452. [PubMed: 15164069]
- Clementz BA, McDowell JE, Zisook S. Saccadic system functioning among schizophrenia patients and their first-degree biological relatives. *Journal of Abnormal Psychology* 1994;103:277–287. [PubMed: 8040497]
- Cohen, J. Statistical power analysis for the behavioral sciences. Academic Press; New York: 1977.
- Cohen, J. Statistical power analysis for the behavioral sciences. Lawrence Earlbaum Associates; Hillsdale, NJ: 1988.
- Crawford TJ, Sharma T, Puri BK, Murray RM, Lewis SW. Saccadic eye movements in families multiply affected with schizophrenia: The Maudsley family study. *The American Journal of Psychiatry* 1998;155:1703–1710. [PubMed: 9842779]
- Curtis CE, Calkins ME, Grove WM, Feil KJ, Iacono WG. Saccadic disinhibition in patients with acute and remitted schizophrenia and their first-degree biological relatives. *The American Journal of Psychiatry* 2001a;158:100–106. [PubMed: 11136640]
- Curtis CE, Calkins ME, Iacono WG. Saccadic disinhibition in patients and their first-degree biological relatives. *Experimental Brain Research* 2001b;137:228–236.
- de Wilde OM, Bour L, Dingemans P, Boeree T, Linszen D. Antisaccade deficit is present in young first-episode patients with schizophrenia but not in healthy young siblings. *Psychological Medicine* 2008;38:871–875. [PubMed: 17949519]
- Ettinger U, Kumari V, Collier DA, Powell J, Luzi S, Michel TM, et al. Catechol-*O*-methyltransferase (COMT) val(158)met genotype is associated with BOLD response as a function of task characteristic. *Neuropsychopharmacol.* 2008e-pub
- Ettinger U, Kumari V, Crawford TJ, Corr PJ, Das M, Zachariah E, et al. Smooth pursuit and antisaccade eye movements in siblings discordant for schizophrenia. *Journal of Psychiatric Research* 2004;38:177–184. [PubMed: 14757332]
- Ettinger U, Picchioni M, Hall M-H, Schulze K, Touloupoulou T, Landau S, et al. Antisaccade performance in monozygotic twins discordant for schizophrenia: The Maudsley twin study. *American Journal of Psychiatry* 2006;163:543–545. [PubMed: 16513882]
- Flint J, Munafo MR. The endophenotype concept in psychiatric genetics. *Psychological Medicine* 2006;37:163–180. [PubMed: 16978446]

- Freedman R, Adler LE, Leonard S. Alternative phenotypes for the complex genetics of schizophrenia. *Biological Psychiatry* 1999;45:551–558. [PubMed: 10088045]
- Glahn D, Almasy L, Blangero J, Burk GM, Estrada J, Peralta JM, et al. Adjudicating neurocognitive endophenotypes for schizophrenia. *American Journal of Medical Genetics* 2007;144B:242–249. [PubMed: 17034022]
- Glass GV. Primary, secondary, and meta-analysis of research. *Educational Researcher* 1976;5:3–8.
- Gottesman II, Gould TD. The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry* 2003;130:636–645. [PubMed: 12668349]
- Greenberg DA. There is more than one way to collect data for linkage analysis: What a study of epilepsy can tell us about linkage strategy for psychiatric disease. *Archives of General Psychiatry* 1992;49:745–750. [PubMed: 1514880]
- Greenwood TA, Braff DL, Light GA, Cadenhead KS, Calkins ME, Dobie DJ, et al. Initial heritability analyses of endophenotypic measures for schizophrenia. *Archives of General Psychiatry* 2007;64:1242–1250. [PubMed: 17984393]
- Hall M-H, Rijdsdijk F, Picchioni M, Schulze K, Ettinger U, Touloupoulou T, et al. Substantial shared genetic influences on schizophrenia and event-related potentials. *American Journal of Psychiatry* 2007;164:804–812. [PubMed: 17475740]
- Hallett PE. Primary and secondary saccades to goals defined by instructions. *Vision Research* 1978;18:1279–1296. [PubMed: 726270]
- Hallmayer JF, Jablensky A, Michie P, Woodbury M, Salmon B, Combrinck J, et al. Linkage analysis of candidate regions using a composite neurocognitive phenotype correlated with schizophrenia. *Molecular Psychiatry* 2003;8:511–523. [PubMed: 12808431]
- Haraldsson HM, Ettinger U, Magnusdottir BB, Sigmundsson T, Sigurdsson E, Ingason A, et al. Catechol-*o*-methyltransferase val158met polymorphism and antisaccade eye movements in schizophrenia. *Schizophrenia Bulletin*. 2008e-pub
- Holzman PS. The role of psychological probes in genetic studies of schizophrenia. *Schizophrenia Research* 1994;13:1–9. [PubMed: 7947411]
- Holzman PS, Matthysse S. The genetics of schizophrenia: A review. *Psychological Science* 1990;1:279–286.
- Iacono WG. Identifying psychophysiological risk for psychopathology: Examples from substance abuse and schizophrenia research. *Psychophysiol* 1998;35:621–637.
- Jablensky A. Subtyping schizophrenia: Implications for genetic research. *Molecular Psychiatry* 2006;11:815–836. [PubMed: 16801952]
- Karoumi B, Saoud M, d'Amato T, Rosenfeld F, Denise P, Gutknecht C, et al. Poor performance in smooth pursuit and antisaccadic eye-movement tasks in healthy siblings. *Psychiatry Research* 2001;101:209–219. [PubMed: 11311924]
- Katsanis J, Kortenkamp S, Iacono WG, Grove WM. Antisaccade performance in patients with schizophrenia and affective disorder. *Journal of Abnormal Psychology* 1997;106:468–472. [PubMed: 9241948]
- Keating M, Atkinson D, Dunn C, Timothy K, Vincent GM, Leppert M. Linkage of a cardiac arrhythmia the long QT syndrome and the Harvey ras-1 gene. *Science* 1991;252:704–706. [PubMed: 1673802]
- Keri S, Janka Z. Critical evaluation of cognitive dysfunctions as endophenotypes of schizophrenia. *Acta Psychiatrica Scandinavica* 2004;110:83–91. [PubMed: 15233708]
- Lander E. Splitting schizophrenia. *Nature* 1988;336:105–106. [PubMed: 2903447]
- Leboyer M, Bellivier F, Nosten-Bertrand M, Jouvent R, Pauls DL, Mallet J. Psychiatric genetics: Search for phenotypes. *Trends in Neuroscience* 1988;21:102–105.
- Levy DL, Holzman PS, Matthysse S, Mendell NR. Eye tracking dysfunction and schizophrenia: A critical perspective. *Schizophrenia Bulletin* 1993;19:461–536. [PubMed: 8235455]
- Levy DL, Mendell NR, Holzman PS. The antisaccade task and neuropsychological tests of prefrontal cortical integrity in schizophrenia: Empirical findings and interpretative considerations. *World Psychiatry* 2004a;2:131–139.

- Levy DL, O'Driscoll G, Cook SR, Matthyse S, Holzman PS, Mendell NR. Antisaccade performance in biological relatives of schizophrenia patients: A meta-analysis. *Schizophrenia Research* 2004b; 71:113–125. [PubMed: 15374579]
- Louchart-de la Chapelle S, Nkam I, Houy E, Belmont A, Menard J-F, Roussignol A-C, et al. A concordance study of three electrophysiological measures in schizophrenia. *American Journal of Psychiatry* 2005;162:466–474. [PubMed: 15741463]
- MacCabe JH, Simon H, Zanelli JW, Walwyn R, McDonald CD, Murray RM. Saccadic distractibility is elevated in schizophrenia patients, but not in their unaffected relatives. *Psychological Medicine* 2005;35:1727–1736. [PubMed: 16300688]
- Malone SM, Iacono WG. Error rate on the antisaccade task: Heritability and developmental change in performance among preadolescent and late-adolescent female twin youth. *Psychophysiol* 2002;39:664–673.
- Massen C. Parallel programming of exogenous and endogenous components in the antisaccade task. *The Quarterly Journal of Experimental Psychology A* 2004;57:475–498.
- Matthyse S, Parnas J. Extending the phenotype of schizophrenia: Implications for linkage analysis. *Journal of Psychiatric Research* 1992;26:329–344. [PubMed: 1491357]
- McDowell JE, Clementz BA. The effect of fixation condition manipulations on antisaccade performance in schizophrenia: Studies of diagnostic specificity. *Experimental Brain Research* 1997;115:333–344.
- McDowell JE, Myles-Worsley M, Coon H, Byerley W, Clementz BA. Measuring liability for schizophrenia using optimized antisaccade stimulus parameters. *Psychophysiol* 1999;36:138–141.
- Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews. Neuroscience* 2006;7:818–827.
- Myles-Worsley M, Coon H, McDowell JE, Brenner C, Hoff M, Lind B, et al. Linkage of a composite inhibitory phenotype to a chromosome 22q locus in eight Utah families. *American Journal of Medical Genetics* 1999;88:544–550. [PubMed: 10490714]
- Petrovsky N, Weiss-Motz F, Schulze-Rauschenbach S, Lemke M, Hornung P, Ruhrmann S, et al. Antisaccade performance is related to genetic loading for schizophrenia. *Journal of Psychiatric Research*. 2008e-pub
- Price GW, Michie PT, Johnston J, Innes-Brown H, Kent A, Clissa P, et al. A multivariate electrophysiological endophenotype, from a unitary cohort, shows greater research utility than any single feature in the Western Australian Family Study of Schizophrenia. *Biological Psychiatry* 2005;60:1–10. [PubMed: 16368076]
- Raemaekers M, Ramsey NF, Vink M, van den Heuvel MP, Kahn RS. Brain activation during antisaccades in unaffected relatives of schizophrenic patients. *Biological Psychiatry* 2006;59:530–535. [PubMed: 16165103]
- Rosenthal, R.; Rosnow, RL. *Essentials of behavioral research: Methods and data analysis*. McGraw-Hill; New York: 1984.
- Rosenthal R, Rubin DB. Comparing effect sizes of independent studies. *Psychological Bulletin* 1982;92:500–504.
- Ross RG, Harris JG, Olincy A, Radant A, Adler LE, Freedman R. Familial transmission of two independent saccadic abnormalities in schizophrenia. *Schizophrenia Research* 1998;30:59–70. [PubMed: 9542789]
- Stefanis NC, van Os J, Avramopoulos D, Smyrnis N, Evdokimidis I, Hantoumi I, et al. Variation in catechol-*o*-methyltransferase val158 met genotype associated with schizotypy but not cognition: A population study in 543 young men. *Biological Psychiatry* 2004;56:510–515. [PubMed: 15450787]
- Szatmari P, Maziade M, Zwaigenbaum L, Merette C, Roy M-A, Joobar R, et al. Informative phenotypes for genetic studies of psychiatric disorders. *American Journal of Medical Genetics* 2007;144B:581–588. [PubMed: 17219386]
- Thaker G. Neurophysiological endophenotypes across bipolar and schizophrenia psychosis. *Schizophrenia Bulletin* 2008;34:760–773. [PubMed: 18502737]
- Thaker G, Cassady S, Adami H, Moran M, Ross D. Eye movements in spectrum personality disorders: Comparison of community subjects and relatives of schizophrenic patients. *American Journal of Psychiatry* 1996;153:362–368. [PubMed: 8610823]

- Thaker GK, Ross DE, Cassady SL, Adami HM, Medoff DR, Sherr J. Saccadic eye movement abnormalities in relatives of patients with schizophrenia. *Schizophrenia Research* 2000;45:235–244. [PubMed: 11042441]
- Turetsky BI, Calkins ME, Light GA, Olincy A, Radant AD, Swerdlow NR. Neurophysiological endophenotypes of schizophrenia: The viability of selected candidate measures. *Schizophrenia Bulletin* 2007;33:69–94. [PubMed: 17135482]
- Walters JTR, Owen MJ. Endophenotypes in psychiatric genetics. *Molecular Psychiatry* 2007;12:886–890. [PubMed: 17895920]

Fig. 1.
Correlation between estimated effect sizes (Cohen's d) for error rate and latency.

Fig. 2.
Correlation between estimated effect sizes (Cohen's d) for error rate and gain.

Fig. 3.

(a) Correlation between estimated effect sizes (Cohen's d) for Latency and Gain. (b) Correlation between estimated effect sizes (Cohen's d) for latency and gain.

Table 1Family studies of antisaccade performance^a

Study	Subjects	Methods	Measures	Results
Clementz et al. (1994)	32 RelSZ ^A	Standard paradigm	AS error rate	↑Errors in RelSZ - ES: 0.51/0.63
Thaker et al. (1996)	33 NC ^B	Standard paradigm	AS error rate	No group difference in AS errors (all RelSZ vs all NC); ES: 0.14/0.13
	12 RelSZ w/SSPS ^{C,D}		AS error rate	SSPS RelSZ vs SSPS NC ES: 0.41/0.37
	14 RelSZ w/o SSPS ^{C,D}		AS latency	Non-SSPS RelSZ vs non-SSPS NC ES: 0.02/0.02
	23 NC w/SSPS ^{C,D,E}		AS latency	↑AS latency (all RelSZ vs all NC); ES: 0.38/0.39
	45 NC w/o SSPS ^{C,D,E}		AS latency	SSPS RelSZ vs SSPS NC ES: 0.0/0.0 Non-SSPS RelSZ vs Non-SSPS NC ES: 0.66/0.67
Katsanis et al. (1997)	55 RelSZ ^F	Standard paradigm	AS error rate	↑Errors in RelSZ; ES: 0.78/1.03
McDowell and Clementz (1997)	38 NC ^{G,H}	Standard/overlap paradigm	AS latency	Latency ES: -0.08/-0.10
	60 RelSZ ^F		AS error rate	↑Errors in 1° RelSZ
	11 2° RelSZ ^F		AS latency	ES (standard): 0.50/0.63
	32 NC ^I		AS gain	ES (overlap): 0.39/0.73 Error rate: within-family correlation Latency (standard) ES: 0.42/0.40 Latency (overlap) ES: 0.60/0.63 Gain ES: -0.06/-0.06
Crawford et al. (1998)	50 RelSZ ^J (multiplex)	Standard paradigm	AS error rate	% Error and latency: no difference between RelSZ and NC
	38 NC ^E		AS latency	% Error: within-family correlation
			Primary saccade gain Final eye position	ES - errors: 0.23/0.26 ES - latency: 0.26/0.23 ES - gain (final): 0.26/0.28
McDowell et al. (1999)	I. 60 RelSZ ^F	Overlap paradigm (near: ±8°/far: ±16°)	AS error rate	↑AS errors in RelSZ
	II. 29 RelSZ ^K (multiplex)		AS latency	Effect sizes greater for far than near eccentricities (I) 0.43/0.70 (near/far)/0.50/1.375
	III. 41 RelSZ ^K (multiplex)		AS gain	(II) 1.23/1.97 (near/far)/1.44/3.75 (III) 1.30/1.90 (near/far)/1.39/3.5
	23 2° RelSZ ^K			Error rates higher in multiplex than simplex families
	94 NC ^L			↑Latency in RelSZ; ES: 1.16/0.8 Gain ES: 0.05/0.07
Ross et al. (1998)	16 Parents of SZ (8 pairs) - 1 most-likely gene carrier (MLC), one least-likely gene carrier (LLC)	Gap paradigm	AS error rate	↑Errors in MLC parents - ES: 0.77/0.75 but not in LLC parents - ES: -0.49/-0.54
	22 NC ^M		AS latency	No latency difference in MLC parents - ES: -0.16/-0.16
			AS accuracy	↓Latency in LLC parents - ES: -1.13/-1.27 No accuracy difference in MLC parents - ES: -0.11/-0.11 ↓Accuracy in LLC parents - ES: 0.95/1.0
Thaker et al. (2000)	21 RelSZ w/ SSPS ^{D,N,O}	Standard paradigm	AS error rate	Error rate ES for all RelSZ vs all NC: 0.19/0.21
	34 RelSZ w/o SSPS ^{D,N,O}		AS latency	↑Errors in RelSZ w/SSPS vs NC w/SSPS (ES: 0.68/0.75) but not vs all NC (ES: 0.19/0.21)
	25 NC w/ SSPS ^{D,N,O,P}			Latency ES for all RelSZ vs all NC: 0.53/0.65
	37 NC w/o SSPS ^{D,N,O,P}			↑Latency in RelSZ w/SSPS vs all other groups

Study	Subjects	Methods	Measures	Results
				SSPS RelSZ vs SSPS NC: 0.90/1.61 Non-SSPS RelSZ vs non-SSPS NC: 0.21/0.19 ↑Errors in RelSZ; ES:0.65/0.80
Curtis et al. (2001a)	116 RelSZ ^Q 109 NC ^{G,R}	Standard paradigm	AS error rate AS latency	% Error within-family correlation Latency ES: 0.15/0.25 ↑Errors in RelSZ only for overlap condition; ES: 0.23/0.29
Curtis et al. (2001b) ^T	42 RelSZ ^S 38 NC ^{G,R}	Standard/overlap paradigm (near: ±10°)	AS error rate	
Karoumi et al. (2001)	21 SibLSZ ^{N,U} 21 NC ^{N,U,V}	Standard paradigm	AS error rate AS latency AS gain AS error rate	Errors in SibSZ; ES: 0.77/1.19 ↓Gain in SibSZ; ES: 0.78/0.78 Latency in SibSZ; ES: 0.14/0.13 No group differences
Brownstein et al. (2003)	98 RelSZ ^W 24 NC ^X	Standard paradigm	AS latency	Error rate ES: -0.05/-0.05 Modest but significant within- family correlation in AS errors, accounting for 7% of variance Latency ES: -0.03/-0.03 ↑Error rate but ns (0.08); ES: 0.60/0.80 ↓Gain; ES: 0.68/0.69 Gain ES for sibs > for patients (0.44/0.58) as in 2006 twin study Latency ES: 0.33/0.35 Within-family correlation not significant No group differences in error rate or latency
Ettinger et al. (2004)	24 SibSZ (1/family) ^{J,U,Y} 24 NC ^{J,U,X,Y}	Standard paradigm	AS error rate AS latency AS gain	
Louchart-de la Chapelle et al. (2005)	31 Parents of SZ ^{Z,b} 73 NC ^{b,c}	Standard paradigm	AS error rate AS latency	Error rate ES: 0.15/0.16 Latency ES: 0.81/0.98
MacCabe et al. (2005)	31 RelSZ (multiplex) ^d 49 RelSZ (simplex) ^d 41 NC ^{d,P}	Standard paradigm	AS error rate AS latency	No difference between simplex or multiplex RelSZ and NC No difference between simplex and multiplex RelSZ Error rate ES - simplex: 0.15/0.14; multiplex: -0.46/-0.41 Latency ES - simplex: 0.53/0.53; multiplex: 0.09/0.08 No effect of genetic loading No within-family association— relatives of good performers had slightly worse scores than relatives of bad performers No group differences
Boudet et al. (2005)	31 RelSZ ^e 31 NC ^f	Standard paradigm	AS error rate AS latency	Error ES: -0.13/-0.10 Latency ES: -0.04/-0.04
Raemaekers et al. (2006)	16 SibSZ ^g 16 NC ^{g,h}	Gap paradigm	AS error rate AS latency	No difference in AS errors or latency Error rate ES: 0.50/0.68 Latency ES: 0.20/0.22
Ettinger et al. (2006)	8 MZ SZ probands 9 MZ Non-SZ co-twins ^{i,j} 10 NC twins (N treated as 20 in analyses) ^{N,k}	Standard paradigm	AS error rate AS latency AS gain	fMRI - less activation in the caudate of SibSZ during AS Error rate: ns difference between non-SZ co-twins and NC twins (ES: 0.43/0.48) ^Y Non-SZ co-twins were significantly more hypometria (ES: 1.0/1.01) and had longer latencies (ES: 0.95/1.03) than NC twins Gain ES > in co-twins than in SZ probands (0.62/0.72) as in 2004 sib study No data on ICC in MZ twin pairs Comparisons were between proband twins or co-twins of probands and both members of NC twin pairs

Study	Subjects	Methods	Measures	Results
Price et al. (2006)	53 RelSZ ^l	Standard paradigm	AS error rate	↑Error rate but ns (0.074); ES: 0.28/0.30
de Wilde et al. (2007)	44 NC ^m	Standard paradigm	AS error rate	↑Error rate but ns (0.08); ES: 0.49/0.55
	25 SibSZ ⁿ		AS latency	↑Gain (hypermetria) in SibSZ; ES: 0.58/0.73
	36 NC ^{n,o}		AS gain	Gain of sibs also significantly higher than gain of SZ (ES: 0.71/0.70). ES for gain in SZ vs NC: 0.24/0.29 Latency ES: 0.47/0.55 Similarity between patients and their siblings: error ICC: 0.33; latency ICC: 0.32; gain ICC: 0.08 All parents vs NC: ↑error rate; ES: 0.66/1.38 ↑Latency but ns (0.06); ES: 0.50/0.58 PFH parents vs NC: ↑error rate - ES: 0.93/2.05; ↑latency - ES: 0.86/1.08 NFH parents vs NC: ↑error rate - ES: 0.70/1.19; latency ns - ES: 0.42/0.45
Petrovsky et al. (2008)	41 Parents of SZ	Overlap paradigm	AS error rate	
	(9 PFH; 32 NFH) ^p		AS latency	
	22 NC ^q			

Small differences in effect sizes reported here and in Levy et al., 2004a are due to rounding.

Only group differences that are statistically significant or show a trend toward significance are included in the results column. Effect sizes included for statistically significant and non-significant comparisons as Cohen's *d*/Glass' delta.

^A Allows psychotic and nonpsychotic Axis I disorders and schizophrenia spectrum personality disorders.

^B No psychosis, major affective disorder, substance abuse/dependence or schizophrenia spectrum personality disorders.

^C No personal Axis I disorder other than one episode of past major depressive disorder (none within the preceding 2 years), for which no hospitalization, tricyclic antidepressant or electroshock treatment was received; no substance abuse within 2 years.

^D SSPS: schizophrenia spectrum personality symptoms: one trait less than needed to meet criteria for schizotypal, schizoid or paranoid personality disorder.

^E No personal or family history of "psychotic illness"/"mental illness".

^F Allows psychotic and nonpsychotic Axis I disorders (Axis II not mentioned). Same sample in McDowell and Clementz (1997) and McDowell et al. (1999).

^G No major affective disorder, psychosis, substance abuse/dependence; Axis II not mentioned.

^H No first-degree relatives treated for psychosis, major affective disorder, substance abuse/dependence; Axis II not mentioned.

^I No Axis I, no family history of psychosis, psychiatric hospitalization or suicide, no *T*-score >70 on MMPI-2 scales L, F, 2, 6, 7, 8.

^J Allows nonpsychotic-related Axis I disorders and schizophrenia spectrum personality disorders, but not psychosis.

^K None psychotic.

^L One parent SZ in remission, 1 parent unspecified functional psychosis in remission; psychiatric status of other parents not specified.

^M No personal or family history of psychotic disorders.

^N No Axis I disorder.

^O No non-schizophrenia-related personality disorders.

^P No family history of psychosis.

^Q Allows past but not current psychotic disorder, nonpsychotic Axis disorders; Axis II not mentioned.

- R* No personal or first-degree family history of treatment for a psychiatric disorder.
- S* Allows past and current psychotic disorder, nonpsychotic Axis disorders; Axis II not mentioned.
- T* Groups are subsets of Curtis et al. (2001a).
- U* No schizotypal personality disorder.
- V* No family history of Axis I disorder.
- W* Allows nonpsychotic Axis I disorders but not schizophrenia spectrum personality disorders or psychosis.
- X* No family history of psychosis.
- Y* No substance dependence in past year or lifetime history >5 years.
- Z* No schizophrenia; other Axis I disorders (psychotic and nonpsychotic) not mentioned; schizotypal personality disorder allowed ($n = 8$).
- a* Results presented for first-degree RelSZ or SibSZ and NC only unless otherwise specified; SZ vs NC comparisons not routinely included in table.
- b* No current alcohol or substance abuse.
- c* No personal or family history of “psychiatric disease” and no psychotropic drug treatment.
- d* Allows nonpsychotic Axis I disorders but not schizophrenia, schizo-affective disorder or schizophreniform disorder; Axis II not mentioned.
- e* No schizophrenia, schizo-affective disorder or paranoid, schizoid, schizotypal personality disorder; unclear if other personality disorders or other Axis I disorders (psychotic and nonpsychotic) allowed.
- f* No personal or family history of “psychiatric illness” or “psychiatric disease”.
- g* No present or past major psychiatric illness.
- h* Screened for psychiatric illnesses in first- and second-degree relatives (presumably applies to NC only).
- i* No substance abuse/dependence in previous year.
- j* No current psychotic or Axis I disorder in co-twins; 5 had diagnoses of past major depressive disorder and/or obsessive-compulsive disorder or schizotypal personality disorder. Three had diagnoses of schizotypal personality disorder.
- k* No family history of psychosis in first- or second-degree relatives.
- l* Excluded if had a diagnosis of schizophrenia.
- m* Excluded if psychotic or had a first-degree relative with a diagnosis of schizophrenia, schizo-affective disorder, schizotypal or paranoid personality disorder, non-affective psychosis (other than persistent delusional disorder) or bipolar affective disorder.
- n* No DSM-IV mood disorder, no psychotic symptoms, no substance abuse diagnosis, no use of cannabis >10 times lifetime.
- o* No first-degree relative who had been treated for a psychiatric disorder.
- p* PFH: positive family history - first or second-degree relative other than offspring who meets criteria for schizophrenia; 7 had first-degree relative; 2 had second-degree relative. NFH: negative family history - no personal history of psychosis or alcohol/substance substance (lifetime); no psychiatric or CNS-acting medications.
- q* No first-, second- or third-degree relative with history of psychosis; no personal history of psychosis or alcohol/substance substance (lifetime); no psychiatric or CNS-acting medications.

Table 2

Means, standard deviations and effect sizes of error rate score in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients ordered by effect size (Cohen's *d*)

Study	First-degree relatives or siblings of schizophrenia patients		Nonpsychiatric controls		Effect size	
	N	Error rate score mean (SD)	N	Error Rate Score Mean (SD)	Cohen's <i>d</i>	Glass' delta
MacCabe et al. (2005) (multiplex families) ^a	31	18.6(12.9)	41	26.6(19.34)	-0.46	-0.41
Boudet et al. (2005)	31	13.0(9.98)	31	14.27(12.96)	-0.13	-0.10
MacCabe et al. (2005)	80	25.22(17.2)	41	26.6(19.4)	-0.08	-0.07
Brownstein et al. (2003)	98	23.2(18.5)	24	24.2(19.1)	-0.05	-0.05
Thaker et al. (1996)	26	26.7 (10.8)	68	24.3(19.2)	0.14	0.13
MacCabe et al., 2005 (simplex families) ^a	49	29.4(19.4)	41	26.6(19.4)	0.15	0.14
Louchart-de la Chapelle et al. (2005)	31	10.29(8.71) ^b	73	9.07(7.5) ^b	0.15	0.16
Thaker et al. (2000)	55	28.96(23.2)	62	24.98(18.7)	0.19	0.21
Crawford et al. (1998)	50	33.0(29.0)	38	27.0(23.0)	0.23	0.26
Price et al. (2005)	53	36.0(22.0)	44	30.0(20.0)	0.28	0.30
Eitinger et al. (2006)	9	30.68(25.29)	20	21.54(19.18)	0.43	0.48
de Wilde et al. (2007)	28	34.7(20.53)	36	25.41(17.03)	0.49	0.55
McDowell and Clementz (1997)	60	17.0(21.0)	32	7.0(16.0)	0.50	0.635
Clementz et al. (1994)	32	32.1(23.2)	33	22.0(15.9)	0.51	0.63
Eitinger et al. (2004)	24	24.27(17.42)	24	15.04(11.53)	0.60	0.80
Curtis et al. (2001a)	116	38.2(22.3)	109	24.6(17.0)	0.65	0.80
Karoumi et al. (2001)	21	36.6(24.1)	21	20.4(13.6)	0.77	1.19
Katsanis et al. (1997)	55	45.0(26.2) ^c	38	25.0(19.5) ^c	0.78	1.03

^a Separate values (based on age and gender-unadjusted scores because SDs of adjusted scores not provided) were calculated separately for the simplex and multiplex families in the MacCabe et al. (2005) study, but are not included in the summary statistics or *t*-test results presented in the text.

Table 3

Variance ratios^a for error rate score in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients (ordered from smallest to largest variance ratio)^b

Study	Variance ratio ^e	P-value	In (variance ratio) ^f
Thaker et al. (1996)	$F = 0.32$ $df = 25, 67$	.99	-1.14
MacCabe et al. (2005) (multiplex families) ^c	$F = 0.44$ $df = 30, 40$	.99	-0.82
Boudet et al. (2005)	$F = 0.59$ $df = 30, 30$	.92	-0.52
MacCabe et al. (2005) ^c	$F = 0.79$ $df = 79, 40$	.81	-0.24
Brownstein et al. (2003)	$F = 0.90$ $df = 97, 23$	.65	-0.06
MacCabe et al. (2005) (simplex families) ^c	$F = 1.0$ $df = 48, 40$	.50	0
Price et al. (2005)	$F = 1.21$ $df = 52, 43$	.26	0.19
Louchart-de la Chapelle et al. (2005) ^d	$F = 1.35$ $df = 30, 72$	.15	0.30
de Wilde et al. (2007)	$F = 1.45$ $df = 27, 35$	.15	0.37
Thaker et al. (2000)	$F = 1.54$ $df = 54, 61$	.051	0.43
Crawford et al. (1998)	$F = 1.59$ $df = 49, 37$	.072	0.46
McDowell and Clementz (1997)	$F = 1.72$ $df = 59, 31$	.051	0.54
Curtis et al. (2001a)	$F = 1.72$ $df = 115, 108$	.002	0.54
Ettinger et al. (2006)	$F = 1.74$ $df = 8, 19$	.15	0.55
Clementz et al. (1994)	$F = 2.13$ $df = 31, 32$	.018	0.76
Ettinger et al. (2004)	$F = 2.28$ $df = 23, 23$	.027	0.82
Karoumi et al. (2001)	$F = 3.14$ $df = 20, 20$	.007	1.14

^a $S_{\text{RelSZ}}^2/S_{\text{NC}}^2$ (one-sided to the right).

^b Katsanis et al. (1997) not included because means and standard deviations not provided.

^c Based on age and gender-unadjusted scores (*SDs* of adjusted scores not provided).

^d *SD* of error score, not of % error score.

^e Separate values for the simplex and multiplex families in the MacCabe et al. (2005) study are presented for completeness but are not included in the summary statistics or in computing the Signed Rank Test (see text). Mean: 1.50, *SD*: 0.71; median 1.54.

^f Mean: 0.28; *SD*: 0.57; median: 0.43; *SE*: 0.15.

Table 4

Means, standard deviations and effect sizes of latency score of correct antisaccades in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients ordered by effect size (Cohen's *d*)

Study	First-degree relatives or siblings of schizophrenia patients		Nonpsychiatric controls		Effect size		
	N	Latency mean (SD)	N	Latency mean (SD)	Cohen's d	r	Glass' delta
Katsanis et al. (1997) ^a	55	290(58.5)	38	294(42)	-0.08	-.04	-0.10
Boudet et al. (2005)	31	344.4(57.46)	31	347.16(67.09)	-0.04	-.02	-0.04
Brownstein et al. (2003)	98	326.5(54.6)	24	328.2(50.1)	-0.03	-.02	-0.03
MacCabe et al. (2005) (multiplex families) ^b	31	406.7(55.7)	41	400(84.6)	0.09	.05	0.08
Karoumi et al. (2001)	21	362(61)	21	353(67)	0.14	.07	0.13
Curtis et al. (2001a)	116	319.8(99.4)	109	308.1(47.7)	0.15	.07	0.25
Crawford et al., 1998	50	342(65)	38	323(84)	0.26	.13	0.23
Eitinger et al. (2004)	24	299.35(54.99)	24	282.22(48.31)	0.33	.16	0.35
Thaker et al. (1996)	26	351(67)	68	326 (64)	0.38	.19	0.39
MacCabe et al., 2005 ^b	80	430.3(71.4)	41	400(84.6)	0.40	.19	0.36
McDowell and Clementz (1997)	60	324.5(43)	32	305.2(48.7)	0.42	.21	0.40
de Wilde et al. (2007)	28	375.88(74.58)	36	345.33(55.19)	0.47	.23	0.55
Thaker et al. (2000)	55	314.62(60.2)	62	285.8(44.2)	0.53	.26	0.65
MacCabe et al., 2005 (simplex families) ^b	49	445.2(79.6)	41	400(84.6)	0.53	.26	0.53
Louchart-de la Chapelle et al. (2005) ^c	31	347.9(66.54)	73	300.25(48.68)	0.81	.38	0.98
Eitinger et al. (2006)	9	302.07(29.19)	20	270.23(30.88)	0.95	.43	1.03

Separate values for the simplex and multiplex families in the MacCabe et al. (2005) study are presented for completeness but are not included in the summary statistics or *t*-test results presented in the text.

^aEstimated based on interquartile range as follows: (3/4) (*Q*₃ - *Q*₁), where *Q*₃ and *Q*₁ were the upper and lower levels of the interquartile range, respectively.

^bbased on age and gender-unadjusted scores, because *SD*s of adjusted scores not provided.

^cMean/*SD* of error score, not of % error score.

Table 5

Means, standard deviations and variance ratios^a of correct AS latency in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients (ordered from smallest to largest) variance ratio

Study	Variance ratio ^b	P-value	ln (variance ratio) ^c
MacCabe et al. (2005) (multiplex) ^b	$F = 0.43$ $df = 30, 40$	.99	-0.84
Crawford et al. (1998)	$F = 0.60$ $df = 49, 37$	.95	-0.51
Boudet et al. (2005)	$F = 0.73$ $df = 30, 30$	.80	-0.31
MacCabe et al. (2005) ^b	$F = 0.71$ $df = 79, 40$	.90	-0.34
McDowell and Clementz (1997)	$F = 0.78$ $df = 59, 31$	.80	-0.25
Karoumi et al. (2001)	$F = 0.83$ $df = 20, 20$	.66	-0.19
MacCabe et al. (2005) (simplex) ^b	$F = 0.89$ $df = 48, 40$	.65	-0.12
Ettinger et al. (2006)	$F = 0.89$ $df = 8, 19$	.54	-0.11
Thaker et al. (1996)	$F = 1.1$ $df = 25, 67$	.37	0.09
Brownstein et al. (2003)	$F = 1.19$ $df = 97, 23$	.33	0.17
Ettinger et al. (2004)	$F = 1.30$ $df = 23, 23$	.27	0.26
de Wilde et al. (2007)	$F = 1.83$ $df = 27, 35$	.05	0.60
Thaker et al. (2000)	$F = 1.86$ $df = 54, 61$	.01	0.62
Louchart-de la Chapelle et al. (2005)	$F = 1.87$ $df = 30, 72$	.02	0.63
Curtis et al. (2001a)	$F = 4.34$ $df = 115, 108$	.00000	1.47

^a S_{RelSZ}^2/S_{NC}^2 (one-sided to the right).

^b Separate values for the simplex and multiplex families in the MacCabe et al. (2005) study are presented for completeness but are not included in these summary statistics (mean: 1.39; SD: 1.0; median: 1.1) or in computing the Signed Rank Test (see text).

^c Mean: 0.16; standard deviation: 0.55; median: 0.09.

Table 6
Means, standard deviations and effect sizes of gain score in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients ordered by effect size (Cohen's *d*)

Study	First-degree relatives or siblings of schizophrenia patients		Nonpsychiatric controls		Effect size		<i>r</i>	Glass' delta
	<i>N</i>	Gain mean (<i>SD</i>)	<i>N</i>	Gain mean (<i>SD</i>)	Cohen's <i>d</i>			
McDowell and Clementz (1997)	60	0.92(0.17)	32	0.91(0.16)	-0.06		-.03	-0.06
Crawford et al. (1998)	50	1.16(0.31)	38	1.08(0.29)	0.26		.13	0.28
de Wilde et al. (2007)	28	1.38(0.60)	36	1.08(0.41)	0.58		.28	0.73
Erttinger et al. (2004)	24	0.86(0.19)	24	1.00(0.203)	0.68		.32	0.69
Karoumi et al. (2001)	21	0.80(0.20)	21	0.98(0.23)	0.78		.36	0.78
Erttinger et al. (2006)	9	0.83(0.115)	20	1.00(0.173)	1.0		.45	1.01

Table 7

Means, standard deviations and variance ratios^a of gain in antisaccade studies (standard paradigm) of first-degree relatives of schizophrenia patients (ordered from smallest to largest) variance ratio

Study	Variance ratio ^b	P-value	ln (variance ratio) ^c
Ettinger et al. (2006)	$F = 0.44$ $df = 8, 19$	.88	-0.82
Karoumi et al. (2001)	$F = 0.76$ $df = 20, 20$	.73	-0.28
Ettinger et al. (2004)	$F = 0.88$ $df = 23, 23$	.62	-0.19
McDowell and Clementz (1997)	$F = 1.13$ $df = 59, 31$	.36	0.12
Crawford et al. (1998)	$F = 1.14$ $df = 49, 37$	.34	0.13
de Wilde et al. (2007)	$F = 2.14$ $df = 24, 34$	.034	0.76

^a $S_{\text{RelSZ}}^2 / S_{\text{NC}}^2$ (one-sided to the right).

^b Mean: 1.08; standard deviation: 0.58; median: 1.00.

^c Mean: -0.045; standard deviation: 0.52; median: -0.035.