

© 2012. This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

**TITLE: Executive functioning in schizophrenia spectrum disorder patients and their unaffected siblings: a ten-year follow-up study.**

**AUTHORS: Ana M. Sánchez-Torres, Virginia Basterra, Lucía Moreno-Izco, Araceli Rosa, Lourdes Fañanás, Amalia Zarzuela, Víctor Peralta, Manuel J. Cuesta\***

**Psychiatric Unit B  
Complejo Hospitalario de Navarra  
c/ Irunlarrea, 4  
31008 Pamplona (Spain)**

**\* Correspondence:**

**Dr. Manuel J. Cuesta  
Psychiatric Unit B, Complejo Hospitalario de Navarra  
c/ Irunlarrea 4,  
E-31008 Pamplona, Spain  
Tel/Fax: +34 848 422488  
Email: [mcuestaz@cfnavarra.es](mailto:mcuestaz@cfnavarra.es)**

**KEY WORDS: Psychosis, schizophrenia spectrum disorders, executive function, siblings, longitudinal study**

## ABSTRACT

Executive dysfunction represents a core deficit that is associated with schizophrenia spectrum disorders (SSDs). However, the longitudinal course of executive deficits in SSDs is still controversial.

The aim of this study was to examine the executive performance of 34 SSD patients in relation to 34 of their unaffected siblings over a period of 10 years. Both groups completed psychopathological and executive assessments. Thirteen healthy controls were assessed using the same instruments.

At baseline, the SSD patients differed significantly from siblings and controls in their performance on Trail Making Test-B (TMT-B) and the number of categories in which they succeeded in the Wisconsin Card Sorting Test (WCST). They also differed significantly from the controls in the total number of errors in the WCST. The siblings did not differ in executive functioning from the controls over the follow-up. Longitudinally, the patients demonstrated significant improvement only for the TMT-B.. However, only 14.71% of the patients showed reliable and clinically significant improvements for the TMT-B, and 8.82% made more errors on the WCST at the follow-up evaluation. Less than 3% of the patients showed either improved or worse results on the remaining measures of the WCST. A stabilisation pattern for the WCST was observed in the three groups

The patients performed worse than their siblings and controls on both executive tests. Some patients exhibited significant improvements in the TMT-B over time, but this improvement was reliable and clinically significant for less than 15% of the sample. Thus, we conclude that the patients exhibited stable impairments over time in the executive functions assessed.

## 1. INTRODUCTION

Patients with schizophrenia spectrum disorders (SSDs) show marked cognitive impairments in most clinical neuropsychological tests and very few preserved domains of cognitive performance (Gold et al., 2009). Executive function, attention and memory are the most severely affected cognitive functions in SSD patients (Goldberg, 2003). Between 27% and 46% of schizophrenia patients exhibit patterns of selective executive dysfunction and between 54% and 90% have at least one executive deficit (Chan et al., 2006a; Johnson-Selfridge and Zalewski, 2001). Executive deficits have a substantial impact on functional outcomes (Altshuler et al., 2007; Bowie et al., 2008; Semkovska et al., 2004) and are closely related to chronicity (Greenwood et al., 2008). In recent years, social cognition has drawn researchers' interest due to its relationship with functional outcomes (Fett et al., 2011)

Similar to other cognitive impairments, executive dysfunction in SSDs represents a core deficit that is present long before the onset of symptoms (Cannon et al., 2006; Wozniak et al., 2008). Moreover, it has been shown that first-degree unaffected relatives show executive impairment profiles that are, although attenuated, similar to those of their affected relatives (Cannon et al., 2000; Harvey et al., 2010; Schulze et al., 2011).

Although there is agreement regarding the prevalence and functional relevance of executive impairment, controversial findings have been reported concerning the course of these deficits in SSD patients. Although significant age-related declines in abstraction functions have been reported among schizophrenia patients (Fucetola et al., 2000), most studies suggest that these initial impairments are stable (Heaton et al., 2001; Mur et al., 2008; Rund, 1998; Townsend and Norman, 2004; Wozniak et al., 2008) or might even improve over time (Gold et al., 1999; Hoff et al., 2005). Longitudinal studies have reported improvement (Gold et al., 1999; Sweeney et al., 1991; Townsend et al., 2002) or no significant change over time (Censits et al., 1997; Hoff et al., 1992) on the Wisconsin Card Sorting Test (WCST). Similar results were found in other patient populations with psychosis, including euthymic bipolar patients, for whom executive functioning and processing speed were reported as the only two cognitive domains that were impaired over a period of two years (Mur et al., 2008).

The aim of the current study was to examine the executive function performance of a group of SSD patients relative to their unaffected siblings and a group of healthy controls over a period of 10 years. Our main goals were to examine the change over time in the executive performance of patients compared to their unaffected siblings and to determine the magnitude of the change and its clinical significance relative to a healthy control group.

We hypothesised that there would be a discrepancy between patients and their unaffected relatives regarding executive impairment, whereby patients would show an initial impairment that would remain stable over time and their unaffected siblings would show a similar trajectory to that of the healthy controls.

## **2. METHODS**

### **2.1. Sample**

The initial sample population included 89 nuclear families. Between 1999 and 2001, 89 patients who were affected by DSM-IV SSDs (APA, 1994) were recruited from consecutive admissions to the Psychiatric Unit of Virgen del Camino Hospital in Pamplona, Spain; the admissions were due to psychotic exacerbations. We subjected the patients' parents and one healthy sibling (sample described elsewhere, (Rosa et al., 2004)) to a comprehensive evaluation that included psychopathological, motor, and neuropsychological assessments.

For the present study, only the patients and their siblings were invited to participate in the second evaluation in 2009. The mean time between the two evaluations was 9 years, 6 months (range: 7-11 years). At the follow-up evaluations, the sample population consisted of 34 sibling pairs (38% of the initial sample). The reasons for participant discontinuation were as follows: the death of one sibling (8 pairs; 7 patients and 1 sibling); traumatic brain injury to the patient (1 pair); the patients moved or were unreachable (11 pairs); and one of the siblings declined to participate (35 pairs).

We also included a healthy control group comprising 26 volunteers (16 men and 10 women). Inclusion criteria for the control subjects were as follows: the absence of major psychiatric disorders, neurological illness or brain injury; the absence of drug or alcohol abuse disorders; the absence of first-degree relatives with major psychiatric illness; and the absence of any drug treatment. At the follow-up evaluation, the control group consisted of 13 subjects (9 men and 4 women). The reasons for participant discontinuation were as follows: 3 subjects declined to participate and 10 had moved or were unreachable. The mean time between the two evaluations was 13 years, 3 months (range: 13-15 years)

All subjects provided written informed consent for participation in the study, and the study was approved by the local ethics committee.

### **2.2. Procedures**

#### **2.2.1. Clinical assessments**

The demographic and clinical variables were assessed according to the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992) at baseline and at the follow-up evaluations. The patients and their siblings underwent psychopathological assessments at both examinations to determine the presence of current or lifetime psychopathological symptoms.

Drug abuse by patients and their siblings was assessed using a structured interview that was based on an adaptation of the Composite International Diagnostic Interview (CIDI) (Organization, 1993). Consumption habits at baseline and at follow-up were determined using the available sources of information (participants, family, and charts). The ratings were classified from 0 (no consumption) to 5 (dependence).

The clinical assessments were carried out by an experienced psychiatrist who was blinded to the participants' cognitive status (VB). At baseline, the assessments were made during an index admission once the patients reached clinical stability. The siblings were also evaluated during the patient's admission. At follow-up, the patients were clinically stabilised for at least 6 months prior to the assessment, and both patients and siblings were evaluated during the same week.

### **2.2.2. Neuropsychological assessments**

The participants were administered the following two executive tests: form B of the Trail Making Test (TMT-B) (Reitan and Wolfson, 1993), which assesses cognitive flexibility, alternating and divided attention, and the computerised version of the Wisconsin Card Sorting Test (WCST-64) (Heaton, 1993), which assesses the cognitive flexibility, the ability to form abstract concepts, to shift and maintain set, and to use feedback (Miyake et al., 2000). The TMT-B is scored according to the time required to draw lines connecting characters that are sequentially alternating between numbers (1-13) and letters (A-L) (e.g., 1-A-2-B). In the WCST, participants are instructed to sort a set of cards and match them with one of four cards presented on the computer on the basis of a target criterion (colour, shape, or number), but they are not told the criterion; they are only told whether their choice is wrong or right. After ten successful trials, the sorting criterion changes without notice, and the participants are required to adjust to the new sorting criterion and continue sorting the cards. The indices of test performance included the total number of categories, total number of errors, number of perseverative errors, conceptual level responses, and failure to maintain set.

An experienced neuropsychologist (AMS) who was blinded to clinical status assessed each participant (patients, siblings and controls).

### 2.3. Data analysis

To compare the demographic characteristics between groups, we applied the t-test and chi-squared test. A one-way ANOVA was performed to compare executive performance between patients, siblings, and controls at baseline and follow-up.

Repeated measures of the multivariate analysis of variance (MANOVA) were performed for each neuropsychological variable to assess time and group effects regarding executive performance. The testing occasion was the within-subjects factor (baseline and follow-up), and the group was the between-subjects factor (patients, siblings, and controls). A multivariate analysis of covariance (MANCOVA) was performed for each variable according to age, gender, drug abuse, and diagnosis to account for potential differences between groups.

To ensure that the study subjects in the later evaluations were representative of the initial cohort, we computed independent t-tests comparing the patients (n=34), siblings (n=34) and controls (n=13) of the current study to those in the original sample who did not participate in the current study (55 patients and siblings and 26 controls) in terms of demographic and clinical variables.

The Reliable Change Index (RCI) was estimated to characterise changes in the executive tests. The RCI is a statistical tool that is used in many areas of medicine to help determine whether an individual's performance on a neuropsychological test has changed from a previous assessment of the same test. It evaluates whether a patient change is unlikely to be caused by a single measure of unreliability, such as measurement error or practice effect (Jacobson and Truax, 1991). It requires a matched, healthy control group to be reassessed using the same instruments within the same timeframe as the patient group.

In addition, we determined the Clinically Significant Change (CSC), which represents the extent to which change over time is clinically meaningful. We calculated the cut-off point according to the C criterion of Evans et al. (Evans et al., 1998). The control group allowed us to test whether the patients moved from a clinical distribution to a normative distribution over time. The cut-off point represents the probability that placement in either distribution is equal and was calculated using the following equation:

$$CSC = [(mean_{patients} \times SD_{controls}) + (mean_{controls} \times SD_{patients})] / (SD_{controls} + SD_{patients})$$

All analyses were conducted using SPSS version 17.0 (SPSS Inc., 2008).

### 3. RESULTS

There were no significant differences between patients who agreed and patients who declined to participate in terms of age ( $27.35\pm 5.61$  vs.  $26.38\pm 6.02$ ;  $t=-0.77$ ,  $p=0.45$ ), years of education ( $11.68\pm 3.12$  vs.  $12.04\pm 3.81$ ;  $t=0.46$ ;  $p=0.64$ ), years since onset of illness ( $6.29\pm 5.61$  vs.  $4.07\pm 5.12$ ;  $t=-1.92$ ,  $p=0.06$ ), number of episodes ( $4.26\pm 5.96$  vs.  $2.81\pm 2.59$ ;  $t=-1.57$ ,  $p=0.12$ ), and level of functioning as measured by the Global Assessment of Functioning scale (GAF; APA, 1994) ( $86.12\pm 10.05$  vs.  $83.83\pm 12.31$ ;  $t=-0.91$ ,  $p=0.37$ ). Likewise, the current and original sibling samples did not differ significantly in terms of age ( $28.29\pm 6.08$  vs.  $27.09\pm 7.18$ ;  $t=-0.82$ ,  $p=0.42$ ), years of education ( $12.63\pm 3.61$  vs.  $12.62\pm 4.39$ ;  $t=-0.01$ ,  $p=0.99$ ) or level of functioning ( $95.52\pm 7.45$  vs.  $92.35\pm 8.06$ ;  $t=-1.46$ ,  $p=0.15$ ). Neither the current and original healthy control samples differ significantly in terms of age ( $30\pm 5.75$  vs.  $32.31\pm 9.69$ ;  $t=0.74$ ,  $p=0.47$ ) nor years of education ( $11.92\pm 2.5$  vs.  $12.54\pm 1.71$ ;  $t=0.73$ ,  $p=0.47$ ). The subsequent results refer to those subjects who completed both assessments (34 sibling pairs and 13 healthy controls).

The patients and siblings were not significantly different in terms of age or years of education, although they differed in terms of gender distribution. The patients and siblings were younger than the controls at follow-up, but the three groups did not differ in years of education. Neither the patients nor the siblings differed from the controls in terms of gender (Table 1).

The siblings and controls did not differ significantly in any of the executive measures at baseline. The performance of the patients at baseline was significantly lower than that of their siblings in the TMT-B and the total number of categories in the WCST. Compared to the controls, the performance of the patients was significantly lower in the TMT-B, the total number of categories and the total number of errors in the WCST. At follow-up, the performance of the patients was significantly lower than that of their siblings for all measures except failure to maintain set. Compared to controls, the performance of the patients was significantly lower in the TMT-B and the total number of categories in the WCST. The siblings and controls differed significantly only in total number of errors in the WCST (Supplementary table 1).

Nine patients (26.5%) and 13 siblings (38.2%) reported cannabis use at baseline, but only 4 (11.8%) and 2 (5.9%), respectively, demonstrated a pattern of dependence. At the follow-up evaluations, 12 patients (35.3%) and 10 siblings (29.4%) reported cannabis use, but only 2 patients (5.9%) and 2 siblings (5.9%) demonstrated a dependence pattern. The pattern of cocaine and amphetamine use was occasional. None of the control subjects had a lifetime drug abuse or dependence diagnosis according to DSM-IV criteria (APA, 1994).

Comparisons between the patients, siblings and controls showed a significant effect based on group for the TMT-B and each of the WCST measures, with the exception of the failure to maintain set. There was no significant effect of time. The interaction between time and group was significant only for the TMT-B, although there was a trend towards significance for total number of errors, perseverative errors and conceptual level responses on the WCST. Only the sibling group improved in terms of performance on the WCST, but this improvement was not significant. We introduced age as a covariate because there were found significant differences between groups at follow-up (Table 2).

Further analyses were performed to account for the effects of additional variables, including gender, drug use, and diagnosis, on the course of executive performance. However, no variable conveyed a significant effect.

To estimate the RCI, we calculated the Pearson correlations for the executive variables in the control group. These data were used as reliability indexes of the measures. The Pearson correlations for the executive measures were as follows: 0.73 for the TMT-B and -0.25 (total number of categories), -0.20 (total number of errors), -0.17 (number of perseverative errors), -0.04 (conceptual level responses), and -0.18 (failure to maintain set) for the WCST. Only the TMT-B showed a significant correlation of the controls' performance at baseline and follow-up. In the WCST performance, controls showed low stability over time.

The results of applying the RCI and CSC to patient and sibling performance to the executive variables over time and in various combinations are shown in Table 3. Twenty-six (76.47%) and 7 (20.59%) patients achieved reliably stable and improving patterns, respectively, over time based on their TMT-B scores. Only 1 patient (2.94%) required more time to complete the task at follow-up, which indicated a reliable, worsening pattern. Only 5 patients (14.71%) showed reliable and clinically significant improvements of their TMT-B scores over time, and no patient showed clinically significant worsening of his or her TMT-B score over time. Improvement patterns were limited in siblings at the individual level because they displayed normal performance at baseline. Only 2 (5.88%) siblings showed reliable improvement patterns regarding controls, but this improvement was not clinically significant. Regarding the WCST, predominantly stable patterns were observed for each of the measures. Reliable worsening patterns were found in 3 patients (8.82%) for the total number of errors and 1 patient (2.94%) for conceptual level responses. Reliable improvement patterns were only observed in 1 patient (2.94%) for perseverative errors.

The clinical significance for the comparison to the control group was achieved in 3 patients (8.82%) for the total number of categories, 11 patients (32.35%) for total



errors, 5 patients (14.71%) for perseverative errors, 1 patient (2.94%) for conceptual level responses and 6 patients (17.65%) for the failure to maintain set. In addition, each of these patients demonstrated a worsening pattern of performance. Clinically significant improvement patterns were observed in 3 patients (8.82%) for the total number of categories, 2 patients (5.88%) for perseverative errors and 13 patients (38.24%) for the failure to maintain set. However, the percentage of patients with reliable and clinically significant worsening patterns was approximately 9% for the total number of errors. Reliable and clinically significant improvement patterns were observed only in 1 (2.94%) patient for the perseverative errors and conceptual level responses.

Reliable and clinically significant improvement patterns were found in 2 (5.88%) siblings for perseverative errors, and 1 (2.94%) sibling for conceptual level responses and failure to maintain set.

#### **4. DISCUSSION**

Our results support three main conclusions. First, significantly lower performance on the TMT-B endured over time for SSD patients regarding their unaffected siblings and the healthy control group, even though the patients showed a significant improvement in this test in the follow-up. In addition, there was a trend towards significance over time in time per group interaction in total number of errors, perseverative errors and conceptual level responses of the WCST. Between group differences were due to significant differences between patients and their unaffected siblings. Second, most of the cross-sectional differences in the executive measures between the three groups found at baseline were evidenced at follow-up. Third, the patients' individual trajectories suggested a stable impairment in executive functioning because only a minority of them showed reliable and clinically significant improvement or worsening patterns on both executive tests.

Our results imply that SSD patients might improve in executive functioning over time, due to the fact that they have room for improvement, because of their notably low performance at baseline. However, despite their improvement they were unable to reach the level of performance of their siblings or healthy subjects. Moreover, siblings and controls remained stable since they only could improve their performance by practice and this putative effect lasting 10 years seems to be unlikely.

It is widely assumed that executive impairment is a core feature of psychosis and it could be a useful target for treatment (Reeder et al., 2004; Wykes et al., 2007). However, there are not many studies addressing the long-term outcome of patients in

their executive functioning and comparing their outcome to their unaffected siblings or to a healthy control group for more than 10 years as it was carried out in this study. Whether executive dysfunction should be accepted as specific for psychosis should rely in demonstrating that this deficit persists over time and that its trajectory differs from control subjects, as it was concluded from our results.

Furthermore, the long-term follow up of executive functioning can account for the examination of unanswered questions, such as the mixed state-trait properties of these impairments in psychosis. In this regard, the extent to which executive deficits are trait or state markers of psychosis is still in debate. Longitudinal data support the idea that executive impairment shares both state and trait properties in psychosis patients because these deficits seem to improve during the first two years and this improvement tends to be sustained (Rund, 1998; Szoke et al., 2008). The initial evidence for executive impairment as a trait marker was drawn from studies of first-episode, medication-naïve patients (Chan et al., 2006b) and from studies that controlled for the use of conventional antipsychotic drugs (Liu et al., 2011) to demonstrate the independence of medication effects. Other authors, including Hughes et al. (Hughes et al., 2005), have suggested that impairments in the WCST represent a disease-related state marker because performance on this test improves with atypical medication (Cuesta et al., 2009; Cuesta et al., 2001; Meltzer and McGurk, 1999) and cognitive remediation (Wykes et al., 1999).

Our results support the conclusion that executive deficits are a trait marker for psychotic disorders because we found a stable impairment over time compared to the siblings and healthy controls. The patients' performance can be interpreted as an arrest of executive function or as a stabilisation of impairments, but it must be noted that some patients had significantly improved performance over time, which also suggests that these deficits owe state properties.

Another finding was the improvement on the TMT-B observed in the siblings' group compared to the controls. Although the differences did not reach significance, the siblings showed worse performance than the controls at baseline. Therefore, the controls had less room for improvement over time, showing a ceiling effect at baseline.

In general, our patients showed a slight improvement in the executive functioning, such as it was found in different short-term studies, but the executive impairment persist as a core feature of psychosis at long-term. However, the extent to which the improvements described in longitudinal studies are, at least in the short term, attributable to a practice effect must be determined (Goldberg et al., 2010; Szoke et al., 2008). Previous studies accounted for a pronounced practice effect on measures of executive function due to the inherent characteristics of the majority of these tests

(Bartels et al., 2010; Basso et al., 1999; Beglinger et al., 2005), which imply the existence of novel problem-solving contexts and serve to assess the capacity to generate different solutions and recognise abstract concepts. Repeated testing generates situations in which novelty is no longer present. The question of whether it is possible to determine a practice effect after a period of 10 years remains unanswered. Regarding the main features of the TMT, it seems unlikely that the improvement shown by both groups is due to a practice effect. In contrast, the WCST presents a novel task, and individuals might internalise the rules. Rather than a practice effect, any effect reflects the retention of the test strategy, which may be recalled when the individuals are exposed to the test a second time. However, the current results demonstrate a stabilisation on the WCST performance in the three groups. Thus, these results suggest that retention of the test strategy seems unlikely at long term.

### **Strengths and Limitations**

The main strength of this study was the comprehensive follow-up period. There have been few longitudinal studies with such an extended period between evaluations (Hoff et al., 2005; Stirling et al., 2003), and the extended follow-up period may shed light on the actual course of executive functioning in these patients.

The patient population sample was an important aspect of this study because it comprised both patients and their unaffected siblings, as well as controls. To the best of our knowledge, few longitudinal studies have assessed cognitive performance in both patients and siblings.

The clinical and cognitive heterogeneity of SSD patients makes further analysis at the individual level useful, and this type of analysis was conducted in the current study using the RCI and CSC. These useful clinical tools may help clinicians to identify patterns of improvement and decline because they allow a method of analysis that extends beyond the group level. However, one limitation of this approach is related to those patients who performed well at baseline and maintained their performance over time. In these cases, there is a ceiling effect that can obscure global results. Thus, these methods can supplement but not replace clinical judgement.

This study has three main limitations. First, the sample size was limited due to the high proportion of discontinuation. This limitation represents the drawback to the lengthy follow-up period.

The second limitation refers to the limited scope of the executive evaluation. Executive function is a complex term that encompasses a set of cognitive processes, including planning, working memory, set shifting, flexibility, response generation and abstract thinking (Strauss et al., 2006). We addressed only a portion of executive

functioning by applying two of the most widely used executive tests to research and clinical practice. Although a more thorough assessment would be desirable, we obtained longitudinal data for the WCST and the TMT-B.

Finally, caution is warranted in interpreting the WCST results over time because of the high variability shown by the control group.

## REFERENCES

- Altshuler, L., Tekell, J., Biswas, K., Kilbourne, A.M., Evans, D., Tang, D., Bauer, M.S., 2007. Executive function and employment status among veterans with bipolar disorder. *Psychiatr Serv* 58(11), 1441-1447.
- Andreasen, N.C., Flaum, M., Arndt, S., 1992. The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Arch Gen Psychiatry* 49(8), 615-623.
- APA, 1994. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. American Psychiatric Association, Washington.
- Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., Ehrenreich, H., 2010. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neurosci* 11, 118.
- Basso, M.R., Bornstein, R.A., Lang, J.M., 1999. Practice effects on commonly used measures of executive function across twelve months. *Clin Neuropsychol* 13(3), 283-292.
- Beglinger, L.J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D.A., Crawford, J., Fastenau, P.S., Siemers, E.R., 2005. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol* 20(4), 517-529.
- Bowie, C.R., Leung, W.W., Reichenberg, A., McClure, M.M., Patterson, T.L., Heaton, R.K., Harvey, P.D., 2008. Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biol Psychiatry* 63(5), 505-511.
- Cannon, M., Moffitt, T.E., Caspi, A., Murray, R.M., Harrington, H., Poulton, R., 2006. Neuropsychological performance at the age of 13 years and adult schizophreniform disorder: prospective birth cohort study. *Br J Psychiatry* 189, 463-464.
- Cannon, T.D., Bearden, C.E., Hollister, J.M., Rosso, I.M., Sanchez, L.E., Hadley, T., 2000. Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophr Bull* 26(2), 379-393.
- Censits, D.M., Ragland, J.D., Gur, R.C., Gur, R.E., 1997. Neuropsychological evidence supporting a neurodevelopmental model of schizophrenia: a longitudinal study. *Schizophr Res* 24(3), 289-298.
- Cuesta, M.J., García de Jalon, E., Campos, M.S., Peralta, V., 2009. Cognitive effectiveness of olanzapine and risperidone in first-episode psychosis. *Br J Psychiatry* 194(5), 439-445.
- Cuesta, M.J., Peralta, V., Zarzuela, A., 2001. Effects of olanzapine and other antipsychotics on cognitive function in chronic schizophrenia: a longitudinal study. *Schizophr Res* 48(1), 17-28.

Chan, R.C., Chen, E.Y., Cheung, E.F., Chen, R.Y., Cheung, H.K., 2006a. The components of executive functioning in a cohort of patients with chronic schizophrenia: a multiple single-case study design. *Schizophr Res* 81(2-3), 173-189.

Chan, R.C., Chen, E.Y., Law, C.W., 2006b. Specific executive dysfunction in patients with first-episode medication-naive schizophrenia. *Schizophr Res* 82(1), 51-64.

Evans, C., Margison, F., Barkham, M., 1998. The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evid Based Mental Health* 1, 70-72.

Fett, A.K., Viechtbauer, W., Dominguez, M.D., Penn, D.L., van Os, J., Krabbendam, L., 2011. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neurosci Biobehav Rev* 35(3), 573-588.

Fucetola, R., Seidman, L.J., Kremen, W.S., Faraone, S.V., Goldstein, J.M., Tsuang, M.T., 2000. Age and neuropsychologic function in schizophrenia: a decline in executive abilities beyond that observed in healthy volunteers. *Biol Psychiatry* 48(2), 137-146.

Gold, J.M., Hahn, B., Strauss, G.P., Waltz, J.A., 2009. Turning it upside down: areas of preserved cognitive function in schizophrenia. *Neuropsychol Rev* 19(3), 294-311.

Gold, S., Arndt, S., Nopoulos, P., O'Leary, D.S., Andreasen, N.C., 1999. Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. *Am J Psychiatry* 156(9), 1342-1348.

Goldberg, T.E., David, A., Gold, J.M., 2003. Neurocognitive deficits in schizophrenia, in: Hirsch, S.R., Weinberger, D. (Ed.), *Schizophrenia*, 2nd ed. Blackwell Science, Oxford, pp. 168-184.

Goldberg, T.E., Keefe, R.S., Goldman, R.S., Robinson, D.G., Harvey, P.D., 2010. Circumstances under which practice does not make perfect: a review of the practice effect literature in schizophrenia and its relevance to clinical treatment studies. *Neuropsychopharmacol* 35(5), 1053-1062.

Greenwood, K.E., Morris, R., Sigmundsson, T., Landau, S., Wykes, T., 2008. Executive functioning in schizophrenia and the relationship with symptom profile and chronicity. *J Int Neuropsychol Soc* 14(5), 782-792.

Harvey, P.D., Wingo, A.P., Burdick, K.E., Baldessarini, R.J., 2010. Cognition and disability in bipolar disorder: lessons from schizophrenia research. *Bipolar Disord* 12(4), 364-375.

Heaton, R.K., Chelune, G.J., Talley, J.L., Talley, J.L., Kay, G.G., Curtiss, G., 1993. *Wisconsin Card Sorting Test (WCST)-CV-64 Psychological Assessment Resources*, Odessa, FL.

- Heaton, R.K., Gladsjo, J.A., Palmer, B.W., Kuck, J., Marcotte, T.D., Jeste, D.V., 2001. Stability and course of neuropsychological deficits in schizophrenia. *Arch Gen Psychiatry* 58(1), 24-32.
- Hoff, A.L., Riordan, H., O'Donnell, D.W., Morris, L., DeLisi, L.E., 1992. Neuropsychological functioning of first-episode schizophreniform patients. *Am J Psychiatry* 149(7), 898-903.
- Hoff, A.L., Svetina, C., Shields, G., Stewart, J., DeLisi, L.E., 2005. Ten year longitudinal study of neuropsychological functioning subsequent to a first episode of schizophrenia. *Schizophr Res* 78(1), 27-34.
- Hughes, C., Kumari, V., Das, M., Zachariah, E., Ettinger, U., Sumich, A., Sharma, T., 2005. Cognitive functioning in siblings discordant for schizophrenia. *Acta Psychiatr Scand* 111(3), 185-192.
- Jacobson, N.S., Truax, P., 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 59(1), 12-19.
- Johnson-Selfridge, M., Zalewski, C., 2001. Moderator variables of executive functioning in schizophrenia: meta-analytic findings. *Schizophr Bull* 27(2), 305-316.
- Liu, K.C.M., Chan, R.C.K., Chan, K.K.S., Tang, J.Y.M., Chiu, C.P.Y., Lam, M.M.L., Chan, S.K.W., Wong, G.H.Y., Hui, C.L.M., Chen, E.Y.H., 2011. Executive function in first-episode schizophrenia: A three-year longitudinal study of an ecologically valid test. *Schizophr Res* 126(1-3), 87-92.
- Meltzer, H.Y., McGurk, S.R., 1999. The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 25(2), 233-255.
- Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn Psychol* 41(1), 49-100.
- Mur, M., Portella, M.J., Martinez-Aran, A., Pifarre, J., Vieta, E., 2008. Long-term stability of cognitive impairment in bipolar disorder: a 2-year follow-up study of lithium-treated euthymic bipolar patients. *J Clin Psychiatry* 69(5), 712-719.
- Reeder, C., Newton, E., Frangou, S., Wykes, T., 2004. Which executive skills should we target to affect social functioning and symptom change? A study of a cognitive remediation therapy program. *Schizophr Bull* 30(1), 87-100.
- Reitan, R., Wolfson, D., 1993. *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Neuropsychology Press, Tucson, AZ.
- Rosa, A., Peralta, V., Papiol, S., Cuesta, M.J., Serrano, F., Martinez-Larrea, A., Fananas, L., 2004. Interleukin-1beta (IL-1beta) gene and increased risk for the depressive symptom-dimension in schizophrenia spectrum disorders. *Am J Med Genet B Neuropsychiatr Genet* 124B(1), 10-14.

- Rund, B.R., 1998. A review of longitudinal studies of cognitive functions in schizophrenia patients. *Schizophr Bull* 24(3), 425-435.
- Schulze, K.K., Walshe, M., Stahl, D., Hall, M.H., Kravariti, E., Morris, R., Marshall, N., McDonald, C., Murray, R.M., Bramon, E., 2011. Executive functioning in familial bipolar I disorder patients and their unaffected relatives. *Bipolar Disord* 13(2), 208-216.
- Semkovska, M., Bedard, M.A., Godbout, L., Limoge, F., Stip, E., 2004. Assessment of executive dysfunction during activities of daily living in schizophrenia. *Schizophr Res* 69(2-3), 289-300.
- Stirling, J., White, C., Lewis, S., Hopkins, R., Tantam, D., Huddy, A., Montague, L., 2003. Neurocognitive function and outcome in first-episode schizophrenia: a 10-year follow-up of an epidemiological cohort. *Schizophr Res* 65(2-3), 75-86.
- Strauss, E., Sherman, E.M.S., Spreen, O., 2006. A compendium of neuropsychological tests, 3rd ed. Oxford University Press, New York.
- Sweeney, J.A., Haas, G.L., Keilp, J.G., Long, M., 1991. Evaluation of the stability of neuropsychological functioning after acute episodes of schizophrenia: one-year followup study. *Psychiatry Res* 38(1), 63-76.
- Szoke, A., Trandafir, A., Dupont, M.E., Meary, A., Schurhoff, F., Leboyer, M., 2008. Longitudinal studies of cognition in schizophrenia: meta-analysis. *Br J Psychiatry* 192(4), 248-257.
- Townsend, L.A., Norman, R.M., 2004. Course of cognitive functioning in first episode schizophrenia spectrum disorders. *Expert Rev Neurother* 4(1), 61-68.
- Townsend, L.A., Norman, R.M., Malla, A.K., Rychlo, A.D., Ahmed, R.R., 2002. Changes in cognitive functioning following comprehensive treatment for first episode patients with schizophrenia spectrum disorders. *Psychiatry Res* 113(1-2), 69-81.
- World Health Organisation (WHO), 1993. Composite International Diagnostic Interview, version 1.1. WHO, Geneva.
- Wozniak, J.R., Block, E.E., White, T., Jensen, J.B., Schulz, S.C., 2008. Clinical and neurocognitive course in early-onset psychosis: a longitudinal study of adolescents with schizophrenia-spectrum disorders. *Early Interv Psychiatry* 2(3), 169-177.
- Wykes, T., Reeder, C., Corner, J., Williams, C., Everitt, B., 1999. The effects of neurocognitive remediation on executive processing in patients with schizophrenia. *Schizophr Bull* 25(2), 291-307.
- Wykes, T., Reeder, C., Landau, S., Everitt, B., Knapp, M., Patel, A., Romeo, R., 2007. Cognitive remediation therapy in schizophrenia: randomised controlled trial. *Br J Psychiatry* 190, 421-427.



### **Acknowledgements**

All of the participants and support staff made this study possible.

### **Conflict of interest**

All authors report no conflict of interest.

## Contributors

Manuel J. Cuesta, Victor Peralta and Lourdes Fañanas designed the study, analysed the data and supervised the draft completion. Virginia Basterra, Araceli Rosa Amalia Zarzuela, Lucía Moreno-Izco and Ana M. Sánchez-Torres managed the literature searches, collected the clinical and cognitive data and contributed to data analysis. Ana Sánchez-Torres wrote the first draft of the manuscript. All authors contributed to and approved the final draft of the paper.

### Role of the Funding Source

This study was partly funded by the Plan Nacional sobre Drogas (grant 2008/I/030, the Department of Health of the Government of Navarra (grant 55/2007) and the Ministerio de Ciencia e Innovación (grant SAF2008-05674-C03-02).

Table 1. Socio-demographic, clinical and diagnostic variables of patients, unaffected siblings and control group.

		Patients (n=34)	Healthy siblings (n=34)	Control group (n=13)	P. vs. S.	S. vs. C.	P. vs. C.
Age	Baseline	27.35(5.61)	28.09(6.28)	30(5.75)	p=0.510	p=0.387	p=0.158
	Follow-up	36.94(5.55)	37.74(6.29)	44.15(5.8)	p=0.583	<b>p=0.003*</b>	<b>p&lt;0.001*</b>
Gender (% men/women)	Baseline and follow-up	70.6/29.4	41.2/58.8	69.2/30.8	<b>p=0.015*</b>	p=0.089	p=0.928
Years of education	Baseline	11.68(3.12)	12.79(3.56)	11.92(2.5)	p=0.173	p=0.424	p=0.800
	Follow-up	11.76(3.19)	12.76(3.62)	11.92(2.5)	p=0.231	p=0.446	p=0.873
Age at onset		21.53(5.65)					
Years since illness onset		15.71(5.61)					
Number of episodes		7.82(5.11)					
Hospitalizations		7.53(9.50)					
DSM-IV diagnoses n(%)							
Schizophrenia		21 (61.8)					
Schizoaffective disorder		8 (23.5)					
Psychotic mood disorder		5 (14.7)					

P. = Patients; S. = Siblings; C. = Controls;  
\*Significance p<0.05

Table 2. Means and standard deviations (SD) of executive variables at baseline and follow-up. Repeated Measures ANOVA between patients, siblings and controls.

	Patients (n=34)		Siblings (n=34)		Controls (n=13)		ANOVA (Time effect)	ANOVA (Group effect)	ANOVA (Time x Group interaction)**
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up	F( p value) df=1	F( p value) df=2 <sup>b</sup>	F(p value) df=1 <sup>a</sup>
<b>TMT-B</b>	127.68 (49.47)	100.35(47.05)	77.29 (27.77)	65.32(22.94)	64.54 (18.6)	68.69(27.15)	0.65(p<0.421); df=1	<b>17.07(p&lt;0.001)*</b> ; df=2	<b>4.14(p=0.02)</b> ; df=1 <sup>a</sup>
<b>WCST: Categories</b>	2(1.44)	2.44(1.62)	2.79(1.7)	3.91(1.24)	3 (1.41)	3.31(1.18)	1.16(p<0.360); df=1	<b>8.85(p&lt;0.001)*</b> ; df=2 <sup>b</sup>	1.97(p=0.147); df=1
<b>WCST-Total errors</b>	23.15(8.99)	22.56(12.67)	19.35(11.28)	12.29(5.58)	17.92 (8.43)	18.23(8.3)	0.25(p=0.616); df=1	<b>7.88(p&lt;0.001)*</b> ; df=2 <sup>b</sup>	2.86(p=0.063); df=1
<b>WCST: Perseverative errors</b>	11.65(6.67)	11.53(7.02)	10.26(6.68)	6.21(2.84)	8.85 (3.6)	8.15(3.6)	0.65(p=0.424); df=1	<b>6.14(p=0.003)*</b> ; df=1 <sup>b</sup>	2.55(p=0.084); df=1
<b>WCST: Conceptual level responses</b>	32.71(13.24)	35.03(17.53)	38.71(16.18)	48.88(8.91)	40.62 (12.64)	41.08(11.63)	1.29(p=0.258); df=1	<b>6.91(p=0.002)*</b> ; df=2 <sup>b</sup>	2.39(p=0.099); df=1
<b>WCST: Failure to maintain set</b>	0.74(0.9)	0.47(0.66)	0.53(0.86)	0.44(0.66)	0.38 (0.65)	0.38(0.65)	0.23(p=0.633); df=1	1.99(p=0.144); df=2	0.38(p=0.686); df=1

TMT-B = Trail Making Test part B; WCST = Wisconsin Card Sorting Test; P = patients; S = siblings; C = controls

\*Significance p<0.05

\*\*Age at follow-up included as covariate.

<sup>a</sup> P<S and C (Time x group interaction)

<sup>b</sup> P<S and C (Group effect)

<sup>c</sup> P<S (Group effect)

Table 3. Cross tabulation of RCI against CSC by executive variables.

	RCI			CSC		RCI+CSC
	Worsening n(%)	Stable n(%)	Improvement n(%)	Worsening n(%)	Improvement n(%)	
<b>TMT-B: patients</b>	1(2.94)	26(76.47)	7(20.59)	0	5(14.71)	5(14.71) Improvement
<b>TMT-B: siblings</b>	0	32(94.12)	2(5.88)	0	0	0
<b>WCST-Categories: patients</b>	0	34(100)	0	3(8.82)	3(8.82)	0
<b>WCST-Categories: siblings</b>	0	34(100)	0	0	5(14.71)	0
<b>WCST-Total errors: patients</b>	3(8.82)	31(91.18)	0	11(32.35)	0	3(8.82) Worsening
<b>WCST-Total errors: siblings</b>	0	34(100)	0	0	6(17.65)	0
<b>WCST-Perseverative errors: patients</b>	0	33(97.06)	1(2.94)	5(14.71)	2(5.88)	1(2.94) Improvement
<b>WCST-Perseverative errors: siblings</b>	0	32(94.12)	2(5.88)	0	5(14.71)	2(5.88) Improvement
<b>WCST-Conceptual level responses: patients</b>	1(2.94)	33(97.06)	0	1(2.94)	0	1(2.94) Worsening
<b>WCST-Conceptual level responses: siblings</b>	0	33(97.06)	1(2.94)	0	1(2.94)	1(2.94) Improvement
<b>WCST-Failure to maintain set: patients</b>	0	34(100)	0	6(17.65)	13(38.24)	0
<b>WCST-Failure to maintain set: siblings</b>	0	33(97.06)	1(2.94)	8(23.53)	9(26.47)	1(2.94) Improvement

RCI = Reliable Change Index; CSC = Clinical Significant Change; TMT-B = Trail Making Test part B; WCST = Wisconsin Card Sorting Test

