Evaluation of a plasticity-based cognitive training program in schizophrenia: Results from the eCaesar trial

Henry W. Mahncke [a,*], Sarah-Jane Kim [a], Annika Rose [a], Catherine Stasio [a], Peter Buckley [b], Stanley Caroff [c], Erica Duncan [d], Sarah Yasmin [e], L. Fredrik Jarskog [f], J. Steven Lamberti [g], Keith Nuechterlein [h], Martin Strassnig [i], Dawn Velligan [j], Joseph Ventura [h], Trina Walker [k], T. Scott Stroup [l], Richard S.E. Keefe [k]

[a] Posit Science, United States of America
[b] Virginia Commonwealth University, United States of America
[c] Corporal Michael J. Crescenz VA Medical Center, The Perelman School of Medicine, University of Pennsylvania, United States of America
[d] Atlanta Veterans Affairs Medical Center, Emory University, United States of America
[e] Palo Alto Veterans Affairs Medical Center, United States of America
[f] University of North Carolina at Chapel Hill, United States of America
[g] University of Rochester, United States of America
[h] University of California, Los Angeles, United States of America
[i] University of Miami, United States of America
[j] University of Texas Health Science Center in Austin, United States of America
[k] Duke University, United States of America
[l] Columbia University, United States of America

ARTICLE INFO

ABSTRACT

Objective: Cognitive impairment in schizophrenia is a core feature of the disorder. Computerized cognitive training has shown promise in pilot studies. A 26-week randomized blinded placebo-controlled trial was conducted to investigate the effect of a novel computerized cognitive training program on cognitive and functional capacity outcomes.

Method: The study followed MATRICS guidelines for the evaluation of interventions designed to improve cognitive function in schizophrenia. Participants (N = 150) were randomized to experimental (computerized cognitive training in a game-like format) or active control (computer games) groups. Training was conducted in-clinic, with an intended training schedule of 5 days per week, 1 h per day, for 26 weeks. Co-primary outcome measures were the MATRICS Consensus Cognitive Battery (MCCB) composite score and the UCSD Performance-Based Skills Assessment (UPSA-2) total score, secondary outcome measures included the Cognitive Assessment Interview (CAI) and the Short-Form-12 Mental Composite Score (SF-12 MCS). Target engagement was assessed with task-learning based assessment.

Results: At baseline, the groups were well matched. No significant effect of the experimental treatment was seen on the primary or secondary outcome measures compared to the active control. Review of the task learning/target engagement data suggested inadequate target engagement.

Conclusions: Results do not support a cognitive or functional capacity benefit from this implementation of a computerized cognitive training program in people with schizophrenia. In future trials, careful consideration is merited of the assessment of task learning/target engagement, the effects of making the cognitive training game-like on motivation, and the implicit effects of trial requirements on participant selection.

1. Introduction

Cognitive impairment associated with schizophrenia is a well-recognized core symptom in schizophrenia that contributes to poor functional outcomes in patients and represents a significant unmet medical need (Kahn and Keefe, 2013). One approach to treating cognitive impairment has been through the development of cognitive training (also referred to as cognitive remediation or cognitive rehabilitation) programs. Such programs have been defined as behavioral training based interventions that aim to improve cognitive processes with the goal of durability and generalization, and can be

* Corresponding author.
E-mail address: henry.mahncke@positscience.com (H.W. Mahncke).

generally divided into compensatory strategies (to circumvent cognitive impairment) and restorative techniques (to improve cognitive function) (Medalia and Saperstein, 2013).

Restorative techniques have recently focused on computerized cognitive training programs that implement intensive "drill & practice" repetition of tasks derived from common neuropsychological assessments of cognitive function. The goal of such programs is to improve the neural systems underlying cognition, and drive generalization of task-based improvement to untrained measures of cognitive function and real-world skills. Several such programs have been developed, and have variously shown promising results (Bowie et al., 2012; Kurtz et al., 2007; Lindenmayer et al., 2013).

One specific restorative approach has been derived from brain plasticity experiments in animal and human models showing that it is possible to reorganize neural systems using intensive adaptive training programs. This approach is based on the viewpoint that a key contributor to poor cognitive function is an underlying deficit in the speed and accuracy of neural information processing coupled with relatively weakened neuromodulatory control over learning (Mahncke et al., 2006; Merzenich et al., 2015). A cognitive training program built specifically to remediate these deficits (Brain Fitness Program, Posit Science) has shown promising results in cognitive function in a several previous trials, including single-site(Ahmed et al., 2015; Fisher et al., 2010, 2009; Loewy et al., 2016; Surti et al., 2011) and multi-site trials(Fisher et al., 2014; Keefe et al., 2012). Notably, several trials have also shown negative results (Goff et al., 2007; Kantrowitz et al., 2016; Murthy et al., 2012; Rass et al., 2012).

Recent meta-analyses have supported the effectiveness of cognitive training (in general, including compensatory and restorative techniques) in schizophrenia(McGurk et al., 2007; Wykes et al., 2011). Nonetheless, at this time cognitive training is not part of schizophrenia treatment guidelines, nor is its provision typically reimbursed for health care providers. Several issues contribute to this situation. First, individual cognitive training trials have typically been small in size (~40 participants in the average study in Wykes et al. (Wykes et al., 2011)), generally with a single trial site. Replication of results in larger multi-site trials, as is typical for pharmaceutical trials, could make the results more compelling. Second, trials using cognitive training as stand-alone intervention have generally employed neuropsychological outcome measures, but not functional capacity measures. Functional capacity measures have recently been standardized and their inclusion is now recommended in clinical trials of putative cognitive enhancers (Green et al., 2011). Extending cognitive training results to include such measures would strengthen the case for the effectiveness of cognitive training. Third, no commercial cognitive training program has been cleared by the FDA as a medical device indicated for the treatment of cognitive symptoms of schizophrenia. Such a clearance, following an evidence-based review of trial data at the FDA, could contribute to the schizophrenia treatment guidelines and reimbursement.

The eCaesar (Evaluation of a Cognitive Adaptive E-treatment in Schizophrenia-diagnosed Adults) study reported here was designed to address these three specific issues by conducting a multi-site randomized controlled trial of a specific cognitive training intervention in schizophrenia using MATRICS standard design criteria and outcome measures.

2. Methods

2.1. Design

This was a multi-site randomized controlled prospective blinded trial. Eleven university-based trial sites were employed. Trial design and outcome measures followed the general principles specified in the MATRICS criteria for trials of cognition-enhancing treatments in schizophrenia (Buchanan et al., 2010, 2005).

2.2. Participants

Inclusion criteria were diagnosis of schizophrenia as defined by DSM-IV-TR criteria and confirmed by the Structured Clinical Interview for DSM-IV (SCID-P), score of 37 or greater on the reading sub-test of the Wide-Range Achievement Test (WRAT-3), clinically stable (non-acute) for ≥8 weeks prior to consent, maintained on stable doses of antipsychotics and/or other concomitant psychotropic treatment for at least 6 weeks prior to consent, no more than a moderate severity rating on hallucinations/unusual thought content items (≤4 on the Positive and Negative Symptoms Scale (PANSS)), minimal level of extrapyramidal symptoms (Simpson-Angus Scale total score ≤ 6), minimal level of depressive symptoms (Calgary Depression Scale(Addington et al., 1993) total score ≤ 10), and age ≥ 18 years. In addition, beyond the standard MATRICS criteria, participants were required to have learned English before age 12 years, and to have the visual, auditory, and motor capacity to use the computerized intervention.

Exclusion criteria were a psychiatric hospitalization within 8 weeks prior to consent, prescribed more than two anti-psychotics; and a history of mental retardation, pervasive developmental disorder, or other neurological disorder. Participants were also excluded for evidence of active suicidal ideation or behavior within one month of consent on the Columbia-Suicide Severity Rating Scale, (C-SSRS) (FDA, 2012), for use of a cognitive training program from Posit Science within 5 years of the date of consent, and for participation in a concurrent clinical trial that could affect the outcome of this one. Since previous studies had documented a negative interaction between serum anticholinergicity and cognitive training gains (Vinogradov et al., 2009), patients were excluded if treated with adjunctive medication with known significant anticholinergic side effects including diphenhydramine and/or benztropine, or prescription for an antipsychotic regimen considered to have high anticholinergicity including clozapine at any dose, olanzapine >20 mg, quetiapine ≥600 mg, chlorpromazine ≥600 mg, or thioridazine ≥600 mg (all doses per day). Participants who appeared to be intoxicated or under the influence of a controlled substance on any day of assessment were rescheduled or discontinued.

2.3. Procedures

Institutional review board approval was obtained at the coordinating center (Posit Science) and at each local site. Participants were reimbursed for their participation; those completing all training and assessment visits could earn $650. Written informed consent was obtained from all participants.

Participants were randomized following baseline assessment. Randomization employed a minimization procedure to stratify participants based on a baseline cut point of 27 on the MCCB composite score. Sites requested randomization allocation through e-mail, and a single staff member fulfilled requests through a concealed randomization allocation sequence.

Participants, neuropsychological testers, and clinician raters were blinded; and participants were reminded regularly not to discuss their training with such clinicians. To maintain the participant blind, consent forms described the study as comparing two distinct types of cognitive training and study staff described the hypothetical benefits of each type. Site staff directly interacting with participants during training were not blinded, but were instructed with scripts to describe each program's features as potentially beneficial.

The protocol specified 130 sessions of cognitive training, intended to be delivered in one-hour sessions, five days per week, for 26 weeks. If a session was missed, it could be made up by completing two sessions on the following day. Participants were given the opportunity (but not required) to extend their participation by up to 4 weeks at the end of the 26-week period to make up missed sessions. Participants came to the site each day for training, and an unblinded coach was available for technical assistance or to answer questions. A 26 week (6-month)

protocol was chosen based on preliminary data suggested that a longer and more comprehensive cognitive training program drove larger cognitive gains (Fisher et al., 2010).

Assessments were performed at baseline (before randomization), at mid-point (three months after randomization, to provide an assessment at a time-point matching the duration of previous trials), and after training program completion.

2.4. Cognitive training programs

Both the experimental treatment and the active control were software-implemented programs, designed to be self-administered by a participant using a personal computer connected to the internet. Participants trained in clinic, typically in a room with several computers, and a coach available to answer questions or provide technical help.

2.4.1. Experimental treatment (ET)

The experimental cognitive training program was custom-built for people with schizophrenia, with 16 cognitive exercises delivered as part of an overarching game wrapper. All exercises targeted speed and accuracy of information processing, with six focused on the auditory system (derived from the Brain Fitness Program, Posit Science), five focused on the visual system (InSight, from Posit Science), three novel exercises focused on executive function, and two novel exercises focused on social cognition. Each exercise employed an adaptive challenge algorithm, making the exercise more difficult as participants completed correct trials, and making the exercise less difficult following incorrect trials and in this way keeping users to ~80% of trials correct regardless of their absolute performance level. Each exercise was composed of numerous distinct stimulus sets that were used in the same exercise paradigm, participants moved through stimulus sets sequentially. Exercises were delivered on a specified schedule such that in each session four exercises were presented from an active set of six. After the content in a given exercise was completed (typically over multiple weeks), that exercise was removed from the active set and the next exercise was added to the active set. In this way, the auditory exercises were presented first, followed by the visual exercises, the executive function exercises, and the social cognition exercises. The game wrapper was designed following a set of in-person focus groups with people with schizophrenia at a local community mental health center that identified core game concepts appealing to this population. Aspirational consumerism and control over one's own personal environment were identified as compelling concepts; consequently, the program was designed such that participants earned virtual cash from completing trials in the cognitive exercises, and at the end of each day's training session, the participant could spend their virtual cash in a virtual store to decorate, customize, and expand their virtual apartment.

2.4.2. Active control (AC)

The active control program was designed to provide an experience that could be matched to the experimental treatment program in intensity and duration, while plausibly engaging cognitive systems to maintain the patient blind. Thirteen off-the-shelf computer games were selected (e.g., Solitaire, Checkers), and delivered with a schedule similar to the experimental treatment.

2.5. Outcome measures

The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) (Kern et al., 2008, p. 2; Nuechterlein et al., 2008) and the UCSD Performance-Based Skills Assessment (UPSA-2) (Green et al., 2011) were specified as co-primary measures; the Cognitive Assessment Interview (CAI) (Ventura et al., 2010) and the Short-Form-12 Health Survey Mental Component Score (SF-12, MCS) (Ware Jr et al., 1996) were secondary outcome measures. Two computerized assessments based on

cognitive exercises from the experimental treatment program were used as positive controls for task learning and target engagement, including an auditory time order judgment task where observers must correctly identify and sequence a pair of frequency-modulated auditory sweeps that can either ascend or descend in frequency and are separated by a brief inter-stimulus interval (Tallal and Piercy, 1973), and a useful field of view task where observers must identify the identity of a central target and the location of a peripheral target after both are briefly simultaneously presented (Ball et al., 1988).

2.6. Statistical analysis

A predefined analysis plan specified sample size, various study populations including the intent-to-treat (ITT) population, and the statistical approach. The trial was powered to test a clinically significant effect size of 0.5 (Cohen's d′) at 2-sided alpha level of 0.05 with a missing data rate of 25% at the final visit, requiring 72 participants per arm. Based on this, the enrollment goal was specified as 150 participants (slightly higher than statistically required). The co-primary outcome measures were the MCCB composite score and the UPSA-2 composite score, evaluated at the final assessment (V3) time point.

The ITT population included all randomized participants who attended their first training day. For confirmatory analysis, fully-evaluable at V2 (FE-V2) and V3 (FE-V3) populations were defined, as were fully-trained at V2 (FT-V2, completing 60+ sessions by V2) and V3 (FT-V3, completing 120+ sessions by V3).

Baseline data were compared with t-tests or chi-squares. Outcome measures were evaluated using linear mixed effects models. Missing data were accounted for using iterative full-information maximum likelihood estimation. Each model included treatment group and time as fixed factors and site as a random factor. An interaction term (training group × time) estimated the effect of cognitive training on outcome measure change. Confirmatory analysis was performed on the FE and FT populations using analysis of covariance (ANCOVA).

This trial was pre-registered at clinicaltrials.gov (NCT 01422902).

3. Results

3.1. Participants

Of 288 individuals contacted, 166 (43.8%) were eligible and consented, and 150 were randomized (Fig. 1). Seven participants dropped out of the ET group before their first training day, and two participants dropped out of the AC group before their first training day, yielding an ITT sample of 141 participants (ET 68; AC 73). Recruitment began in April 2012; the final participant completed post-training assessment in March 2015.

Pretraining demographic and baseline measures are shown in Table 1. There were no significant differences between the ET and AC groups.

At baseline, the population that did not complete the final assessment (the drop/withdraw, or D/W, population, N = 39) was significantly younger than the population that did go on to complete the final assessment (the fully-evaluable at V3, or FE-V3, population, N = 102), had a higher WRAT score, higher PANSS positive symptom score, higher SF-12 mental component score, and notably better performance on the two computerized assessments of auditory and visual speed (Table 2, all noted differences significant at the p < 0.05 level).

Following set-up, 44 participants completed the final assessment in the ET group and 24 dropped or withdrew, while 56 participants completed the final assessment in the AC group and 17 dropped or withdrew, which was not statistically significant (p = 0.12, chi-square). The drop/withdraw populations from the ET and AC groups were not significantly different in terms of demographic or outcome measures from each other. Table 2 (right columns) shows the comparisons between the ET D/W and the AC D/W populations. Each D/W group was also not statistically different from its respective completer group
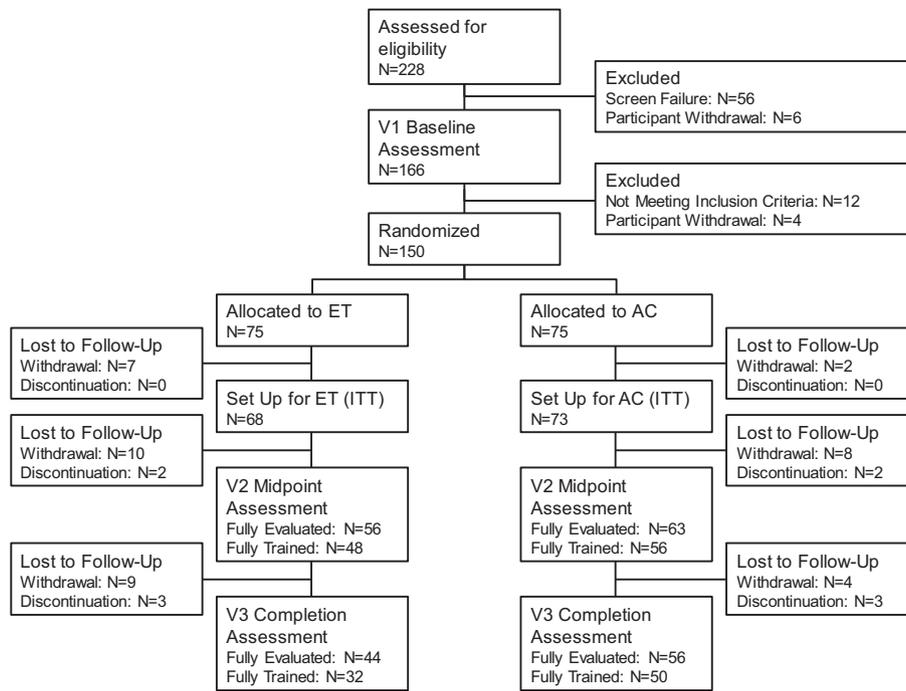
Fig. 1. CONSORT diagram.

(data not shown). Reasons for drop/withdraw were typically the time commitment of study participation, change in life circumstances (e.g., new job, family moving), or lost to follow-up.

Within the fully-evaluated at final assessment (FE-V3) group, 32 of 44 participants in the ET group were considered fully-trained (completing ≥120 of a potential 130 sessions), as were 50 of 58 participants in the AC group, indicating a non-significant trend towards lower treatment compliance (p = 0.09, chi-square). The not-fully trained populations from the ET and AC groups were not significantly different in terms of demographic or outcome measures from each other, and each drop/withdraw group was not significantly different from its respective completer group (data not shown).

3.2. Training effects on outcome measures

Within group change scores, between groups difference scores, and significance for ET and AC comparisons in the ITT group are reported in

Table 1
Baseline demographic and outcome measures.

	ITT group (N = 141)	ET group (N = 68)	AC group (N = 73)
Age (years)	42.9 ± 13.1	43.1 ± 12.4	42.8 ± 13.9
Education (years)	12.7 ± 2.1	12.5 ± 1.9	12.8 ± 2.2
WRAT (baseline IQ)	48.2 ± 5.6	47.5 ± 5.3	48.9 ± 5.8
CDSS (depression)	2.6 ± 2.3	2.5 ± 2.2	2.7 ± 2.4
PANSS positive	14.4 ± 4.4	14.6 ± 4.7	14.3 ± 4.2
PANSS negative	16.1 ± 5.5	16.1 ± 5.1	16.2 ± 5.9
Gender (% male)	81%	78%	84%
Ethnicity (% Caucasian)	85%	87%	84%
MCCB composite	29.3 ± 12.6	29.2 ± 12.6	29.4 ± 12.8
UPSA composite	87.1 ± 14.6	85.9 ± 13.9	88.2 ± 15.3
Auditory time order judgment (ms)	216 ± 227	235 ± 228	199 ± 225
Useful field of view (ms)	781 ± 702	872 ± 765	696 ± 632
CAI	60.5 ± 12.3	60.5 ± 12.6	60.6 ± 12.1
SF-12 (MCS)	46.5 ± 13.3	45.6 ± 12.6	47.5 ± 13.9

No significant differences between ET and AC groups (p < 0.05).
Mean ± 1 standard deviation or % of variable.

Table 3. Both the ET and AC groups showed numerical improvements in MCCB score, with no significant between group differences at the V2 (mid-point) or V3 (final) assessment. Both groups also showed numerical improvements in UPSA total score, with no significant between group differences at the V2 or V3 assessment. Secondary measures (CAI and SF-12 MCS) showed a similar pattern.

On the auditory time order judgment task, a train-to-the-task positive control for task learning, both groups showed improvements at the V2 time point with no significant difference between groups. At the V3 time point, the improvement (relative to baseline) in the ET group was statistically larger than that of the AC group.

On the visual useful field of view task, also a train-to-the-task positive control for task learning, both groups show improvements at the V2 and V3 time points, with the ET improvements statistically larger in both cases.

Primary outcome measures and train-to-the-task positive control data for the ITT population using the primary linear mixed model analysis approach are shown graphically in Fig. 2.

Analyses were repeated using ANCOVA with no missing data on the fully-evaluable populations (at V2 and V3) the fully-trained populations (V2, V3; all data not shown); both showed a pattern of results identical to the ITT population.

4. Discussion

The eCaesar study was designed to be a pivotal trial for this specific implementation of computerized brain-plasticity-based cognitive training in schizophrenia. Strengths of the study include its relatively large sample size and design consistent with MATRICS guidance. The overall drop/withdraw rate (29%) was slightly higher than planned (25%), but not inconsistent with meta-analyses of such rates in complex interventions in schizophrenia (Szymczynska et al., 2017). No significant effect of the program was seen in either of the co-primary endpoints (MCCB and UPSA), nor the secondary end-points (CAI and SF-12 MCS).

Current guidelines for clinical trial design from the National Institutes of Mental Health specify that trials should where possible include measures of target engagement, to facilitate the interpretation of both

Table 2
Drop/withdraw analysis.

	Comparison of fully-evaluated to drop/withdraws		Comparison of treatment group drop/withdraws	
	FE-V3 group (N = 100)	D/W-V3 group (N = 41)	D/W-V3 ET group (N = 24)	D/W-V3 AC group (N = 17)
Age (years)	45.7 ± 12.8[***]	36.3 ± 11.6[***]	38.8 ± 12.8	34.1 ± 9.7
Education (years)	12.7 ± 2.0	12.5 ± 2.1	12.3 ± 2.1	12.9 ± 2.1
WRAT (baseline IQ)	47.7 ± 5.5[*]	49.6 ± 5.5[*]	46.8 ± 5.1	48.5 ± 5.6
CDSS (depression)	2.7 ± 2.4	2.3 ± 2.1	2.1 ± 2.1	2.6 ± 1.1
PANSS positive	14.0 ± 4.5[*]	15.4 ± 4.2[*]	15.0 ± 4.6	15.8 ± 3.7
PANSS negative	15.9 ± 5.1	16.6 ± 6.4	16.1 ± 5.4	17.2 ± 7.7
Gender (% male)	79%	85%	88%	82%
Ethnicity (% Caucasian)	87%	80%	71%[*]	94%[*]
MCCB composite	28.2 ± 12.1	32.0 ± 13.6	32.0 ± 12.5	32.0 ± 15.3
UPSA composite	87.0 ± 15.3	87.3 ± 13.1	87.2 ± 12.9	87.4 ± 13.6
Auditory time order judgment (ms)	241 ± 247[**]	154 ± 152[**]	170 ± 160	131 ± 141
Useful field of view (ms)	852 ± 787[**]	606 ± 385[**]	672 ± 443	512 ± 270
CAI	60.2 ± 12.6	61.4 ± 11.8	60.1 ± 12.0	63.2 ± 11.6
SF-12 (MCS)	68.1 ± 22.1[**]	76.8 ± 21.7[**]	72.5 ± 23.1	82.9 ± 18.6

Mean ± 1 standard deviation or % of variable.
[***] p < 0.01.
[**] p < 0.05.
[*] p < 0.01.

positive and negative results (Insel, 2015). Conceptually, a trial that demonstrates target engagement but does not show generalized patient benefit could be interpreted as casting doubt on the value of the target, and a trial that demonstrates a lack of target engagement and lack of generalized patient benefit could be interpreted as casting doubt on the approach used to engage the target.

Table 3
Outcome measure analysis (ITT population).

	Experimental training within group differences			Active control within group differences			V2-V1 between groups difference		V3-V1 between groups difference	
	V1 baseline Mean ± SD (range)	V2-V1 change Mean (95% CI)	V3-V1 change Mean (95% CI)	V1 baseline Mean (95% CI)	V2-V1 change Mean (95% CI)	V3-V1 change Mean (95% CI)	Change difference[a] (95% CI) Effect size p value	F value (df)	Change difference (95% CI) Effect size p value	F value (df)
Primary measures										
MCCB composite	29.2 ± 12.6 (−1 to 65)	+0.9 (−0.7 to +2.4)	+1.0 (−0.8 to +2.8)	29.4 ± 12.8 (5 to 65)	+0.5 (−0.9 to +1.9)	+2.5 (0.9 to 4.1)	+0.4 (−1.7 to 2.5) d = 0.03 p = 0.72	F = 0.1274 (113.6)	−1.5 (−4.2 to 0.7) d = −0.11 p = 0.22	F = 1.5 (103.4)
UPSA composite	85.9 ± 13.9 (48–112)	+4.0 (1.3 to 6.7)	+4.0 (1.1 to 6.8)	88.2 ± 15.3 (48–117)	+3.6 (1.1 to 6.1)	+3.7 (1.1 to 6.2)	+0.4 (−3.3 to 4.0) d = 0.03 p = 0.84	F = 0.0429 (117.1)	+0.3 (−3.5 to 4.0) d = 0.02 p = 0.88	F = 0.0232 (109.2)
Positive controls										
Auditory time order judgment (milliseconds; lower is better)	235 ± 238 (21–989)	−98 (−159 to −38)	−114 (−173 to −54)	199 ± 225 (21–1069)	−55 (−111 to −1)	−32 (−87 to +23)	−43 (−125 to +38) d = 0.23 p = 0.30	F = 1.068 (129.6)	−82 (−162 to −2) d = 0.46 p = 0.047	F = 4.026 (124.3)
Useful field of view (milliseconds; lower is better)	872 ± 765 (106–4311)	−424 (−597 to −251)	−443 (−637 to −249)	696 ± 632 (34–5264)	−96 (−257 to +65)	−155 (−335 to +25)	−328 (−562 to −94) d = 0.52 p = 0.0069	F = 7.531 (127.8)	−288 (−549 to −25) d = 0.50 p = 0.033	F = 4.622 (135.4)
Secondary measures										
CAI	58.3 ± 12.4 (35–90)	+1.9 (−1.3 to +5.1)	+3.9 (−0.01 to +7.7)	55.6 ± 10.3 (29–80)	+5.4 (1.9 to 9.0)	+8.0 (3.4 to 12.6)	−3.5 (−8.2 to +1.1) d = −0.30 p = 0.14	F = 2.2 (54.3)	−4.2 (−10.0 to 1.7) d = −0.35 p = 0.17	F = 1.923 (57.7)
SF-12 (MCS)	45.6 ± 12.6 (6.5–64.1)	+1.1 (−2.6 to +4.8)	−3.3 (−6.1 to −0.6)	47.5 ± 13.9 (9.8–67.4)	+0.8 (−2.6 to +4.2)	−0.3 (−2.8 to 2.3)	+0.3 (−4.7 to +5.3) d = 0.02 p = 0.90	F = 0.0151 (124.5)	−3.0 (−6.7 to +0.7) d = −0.27 p = 0.11	F = 2.581 (116.5)

[a] Effect size signs oriented so positive numbers represent a greater change for the ET group.
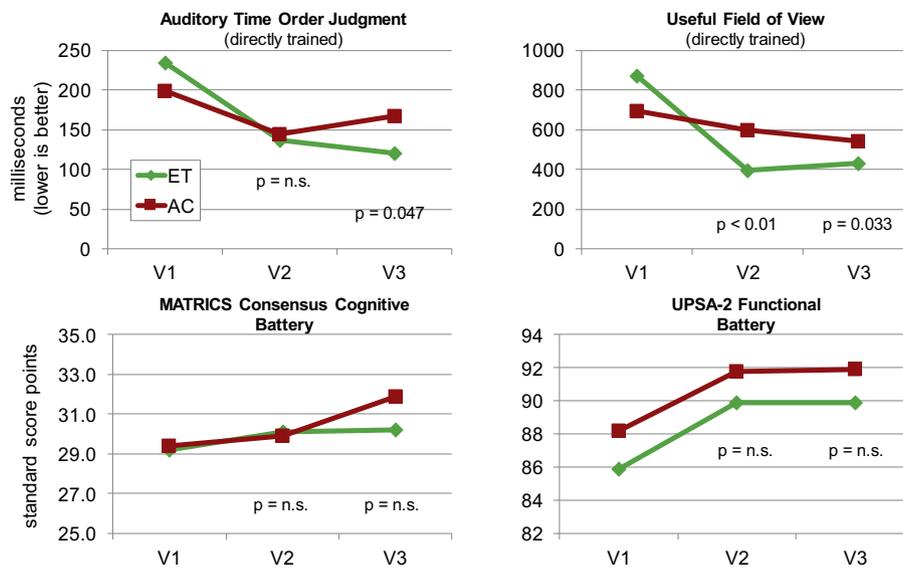
Fig. 2. ITT study population: train-to-the-task positive controls and primary outcome measures ET N = 68; AC N = 73.

The current trial included two measures of target engagement, auditory time order judgment (a measure of auditory processing speed) and visual useful field of view (a measure of visual processing speed). The study showed improvement in both the ET and AC groups in the auditory measure, with no significant between group difference at the mid-point assessment, when due to the ET exercise sequence all auditory training was complete. During the second half of the training period, a between group difference emerged; however, the absolute performance level of the ET group did not approach performance levels in previous trials that documented cognitive improvement. For example, the final performance scores in the ET group of 121 ms on the auditory time order judgment task were quantitatively worse than in two key previous pilot studies: CRIS score of 70 ms (Fisher et al., 2009) and CRSTN score of 71 ms (Keefe et al., 2012). A previous study quantitatively analyzed the relationship between gains in auditory speed on this measure and overall cognitive gains and concluded that the final level of auditory speed performance predicted the magnitude of cognitive gain, and participants who did not achieve a level faster than ~85 ms did not show generalized cognitive gains (Biagianti et al., 2016). Hence, one possible explanation for the lack of generalization to primary outcome measures in the current trial as they did in previous studies is that the program/protocol failed to engage the treatment target.

A significant between groups difference was seen in the visual measure for target engagement. This measure (useful field of view) is correlated with working memory performance in schizophrenia (Gray et al., 2014). This provides first evidence suggesting that engaging this target does not lead to improved cognitive function in schizophrenia, despite evidence for that effect in healthy aging (Ball et al., 2007).

A possible reason for the poor target engagement of the auditory measure is that substantial changes were made from the program as used in previous trials to the current program. The previous implementation of the cognitive training approach used in this study (Brain Fitness Program, Posit Science) was designed for healthy older adults and focused on motivating users with progress and performance metrics. Because of substantial interest in the use of gamification techniques to make cognitive training more engaging and compelling (Fleming et al., 2017; Lumsden et al., 2016), for the current study a "game wrapper" was developed specifically for people with schizophrenia. A key feature of this gamification was that each correct trial in a cognitive training exercise yielded a reward in virtual currency, and at the end of each day's training session the virtual currency could be used in an in-game shopping mall to select decorations for a virtual apartment.

Informal discussions suggested that users enjoyed the game wrapper and spent time decorating and customizing their virtual apartments. This approach may have caused a problem. Recent viewpoints have suggested that extrinsic motivation can have a negative impact on the amount of learning that takes place, in part by undermining intrinsic motivation (Saperstein and Medalia, 2015), and gamification elements specifically lowered learning rates in a recent trial of cognitive training designed to test this hypothesis, perhaps by distracting users from the cognitive tasks themselves (Katz et al., 2014). In particular, in the current study, an unforeseen consequence of the gamification approach was that users would receive the same virtual cash earnings regardless of their performance levels in the cognitive training tasks, because each cognitive training task adapted in difficulty to track the user to ~80% correct trials. This may have resulted in a lack of incentive for users to attend to the cognitive exercises and improve in-exercise performance. Future research in cognitive training program design should investigate the tradeoffs between gamification and cognitive gains, and implement designs that align those elements in the service of the primary goal of cognitive enhancement (for example, by providing in-game rewards on the basis of performance advancement relative to baseline on target engagement measures).

A second potential reason for the lack of efficacy could be that participants in the trial had a mean age of ~43, which was similar to an initial single-site study (~43) (Fisher et al., 2009) but older than the multi-site feasibility study (~39) (Keefe et al., 2012). Older patients appear to benefit from cognitive training less than younger participants (Kontis et al., 2013), and this may have contributed to the lack of efficacy.

A third potential reason might be that the trial design may have inadvertently selected a different population from that enrolled in previous studies. For example, in the current trial participants committed to traveling to their local clinic 5 days per week for 6 months, generally via public transportation, potentially biasing the study population towards individuals with better cognitive and functional capacity. Those who dropped out of the study were on average younger and had better baseline cognitive performance, suggesting that this relatively demanding schedule may have made study participation less appealing to participants whose motivation and cognitive/functional skills allowed them to choose other options. In addition, previous studies have generally been at VA sites, while the current study included a mix of 4 VA sites and 7 university sites, which could have resulted in a different mix of participant motivations, perhaps increasing the number of "professional subjects" (Devine et al., 2013; McCann et al., 2015). Recent studies have made use of home-based training models (Fisher et al., 2014; Loewy

et al., 2016), which may represent a design that attracts and retains a broader spectrum of participants.

Finally, it must be considered that the cognitive training approach is simply ineffective, and thus by implication, previous results from smaller studies were spurious. There has been one previous open label study showing negative results (Murthy et al., 2012), one randomized controlled trial showing negative results (Rass et al., 2012) and two studies involving a combination of this form of cognitive training with pharmaceutical augmentation showing negative results (Goff et al., 2007; Kantrowitz et al., 2016).

These issues could be resolved in a future trial, designed with the goals of the current trial - multi-site and involving standardized cognitive and functional capacity endpoints - while addressing the issues raised in this study, by redesigning the program to provide more continuous actionable feedback on target engagement, using a version of the cognitive training approach that relied on engagement through intrinsic motivation regarding performance improvement rather than extrinsic motivation around virtual rewards, and by restructuring the demands of the trial to a more conventional 8–12 week intervention that allowed retention of a more representative sample of participants. The unmet medical need of cognitive impairment associated with schizophrenia warrants further efforts to rigorously evaluate cognitive training techniques.

Sponsor's role & conflict of interest

5. Funding source

References

Addington, D., Addington, J., Maticka-Tyndale, E., 1993. Assessing depression in schizophrenia: the Calgary Depression Scale. Br. J. Psychiatry 163 (S22), 39–44.

Ahmed, A.O., Hunter, K.M., Goodrum, N.M., Batten, N.-J., Birgenheir, D., Hardison, E., Dixon, T., Buckley, P.F., 2015. A randomized study of cognitive remediation for forensic and mental health patients with schizophrenia. J. Psychiatr. Res. 68, 8–18. https://doi.org/10.1016/j.jpsychires.2015.05.013.

Ball, K.K., Beard, B.L., Roenker, D.L., Miller, R.L., Griggs, D.S., 1988. Age and visual search: expanding the useful field of view. J. Opt. Soc. Am. A 5, 2210–2219.

Ball, K., Edwards, J.D., Ross, L.A., 2007. The impact of speed of processing training on cognitive and everyday functions. J. Gerontol. B Psychol. Sci. Soc. Sci. 62 (1), 19–31.

Biagianti, B., Fisher, M., Neilands, T.B., Loewy, R., Vinogradov, S., 2016. Engagement with the auditory processing system during targeted auditory cognitive training mediates changes in cognitive outcomes in individuals with schizophrenia. Neuropsychology 30, 998–1008.

Bowie, C.R., McGurk, S.R., Mausbach, B., Patterson, T.L., Harvey, P.D., 2012. Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. Am. J. Psychiatry 169, 710–718.

Buchanan, R.W., Davis, M., Goff, D., Green, M.F., Keefe, R.S.E., Leon, A.C., Nuechterlein, K.H., Laughren, T., Levin, R., Stover, E., Fenton, W., Marder, S.R., 2005. A summary of the FDA-NIMH-MATRICS workshop on clinical trial design for neurocognitive drugs for schizophrenia. Schizophr. Bull. 31, 5–19. https://doi.org/10.1093/schbul/sbi020.

Buchanan, R.W., Keefe, R.S.E., Umbricht, D., Green, M.F., Laughren, T., Marder, S.R., 2010. The FDA-NIMH-MATRICS guidelines for clinical trial design of cognitive-enhancing drugs: what do we know 5 years later? Schizophr. Bull. https://doi.org/10.1093/schbul/sbq038.

Devine, E.G., Waters, M.E., Putnam, M., Surprise, C., O'Malley, K., Richambault, C., Fishman, R.L., Knapp, C.M., Patterson, E.H., Sarid-Segal, O., Streeter, C., Colanari, L., Ciraulo, D.A., 2013. Concealment and fabrication by experienced research subjects. Clin. Trials Lond. Engl. 10, 935–948. https://doi.org/10.1177/1740774513492917.

FDA, 2012. Guidance for industry suicidal ideation and behavior: Prospective assessment of occurrence in clinical trials (Draft Guidance). United States Food and Drug Administration.

Fisher, M., Holland, C., Merzenich, M.M., Vinogradov, S., 2009. Using neuroplasticity-based auditory training to improve verbal memory in schizophrenia. Am. J. Psychiatry 166, 805–811. https://doi.org/10.1176/appi.ajp.2009.08050757.

Fisher, M., Holland, C., Subramaniam, K., Vinogradov, S., 2010. Neuroplasticity-based cognitive training in schizophrenia: an interim report on the effects 6 months later. Schizophr. Bull. 36, 869–879. https://doi.org/10.1093/schbul/sbn170.

Fisher, M., Loewy, R., Carter, C., Lee, A., Ragland, J.D., Niendam, T., Schlosser, D., Pham, L., Miskovich, T., Vinogradov, S., 2014. Neuroplasticity-based auditory training via laptop computer improves cognition in young individuals with recent onset schizophrenia. Schizophr. Bull. 41, 250–258. https://doi.org/10.1093/schbul/sbt232.

Fleming, T.M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S.N., Cheek, C., Lucassen, M., Lau, H.M., Pollmuller, B., Hetrick, S., 2017. Serious games and gamification for mental health: current status and promising directions. Front. Psych. 7.

Goff, D.C., Lamberti, J.S., Leon, A.C., Green, M.F., Miller, A.L., Patel, J., Manschreck, T., Freudenreich, O., Johnson, S.A., 2007. A placebo-controlled add-on trial of the ampakine, CX516, for cognitive deficits in schizophrenia. Neuropsychopharmacology 33, 465–472.

Gray, B.E., Hahn, B., Robinson, B., Harvey, A., Leonard, C.J., Luck, S.J., Gold, J.M., 2014. Relationships between divided attention and working memory impairment in people with schizophrenia. Schizophr. Bull., sbu015 https://doi.org/10.1093/schbul/sbu015.

Green, M.F., Schooler, N.R., Kern, R.S., Frese, F.J., Granberry, W., Harvey, P.D., Karson, C.N., Peters, N., Stewart, M., Seidman, L.J., et al., 2011. Evaluation of functionally meaningful measures for clinical trials of cognition enhancement in schizophrenia. Am. J. Psychiatry 168, 400–407.

Insel, T.R., 2015. The NIMH experimental medicine initiative. World Psychiatry 14, 151–153. https://doi.org/10.1002/wps.20227.

Kahn, R.S., Keefe, R.S.E., 2013. Schizophrenia is a cognitive illness: time for a change in focus. JAMA Psychiat. 70, 1107–1112. https://doi.org/10.1001/jamapsychiatry.2013.155.

Kantrowitz, J.T., Sharif, Z., Medalia, A., Keefe, R.S.E., Harvey, P., Bruder, G., Barch, D.M., Choo, T., Lee, S., Lieberman, J.A., 2016. A multicenter, rater-blinded, randomized controlled study of auditory processing–focused cognitive remediation combined with open-label lurasidone in patients with schizophrenia and schizoaffective disorder. J. Clin. Psychiatry 77, 799–806. https://doi.org/10.4088/JCP.15m09998.

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., Shah, P., 2014. Differential effect of motivational features on training improvements in school-based cognitive training. Front. Hum. Neurosci. 8, 242.

Keefe, R.S.E., Vinogradov, S., Medalia, A., Buckley, P.F., Caroff, S.N., D'Souza, D.C., Harvey, P.D., Graham, K.A., Hamer, R.M., Marder, S.M., Miller, D.D., Olson, S.J., Patel, J.K., Velligan, D., Walker, T.M., Haim, A.J., Stroup, T.S., 2012. Feasibility and pilot efficacy results from the multisite cognitive remediation in the schizophrenia trials network (CRSTN) randomized controlled trial. J. Clin. Psychiatry 73, 1016–1022. https://doi.org/10.4088/JCP.11m07100.

Kern, R.S., Nuechterlein, K.H., Green, M.F., Baade, L.E., Fenton, W.S., Gold, J.M., Keefe, R.S.E., Mesholam-Gately, R., Mintz, J., Seidman, L.J., et al., 2008. The MATRICS consensus cognitive battery, part 2: co-norming and standardization. Am. J. Psychiatry 165, 214.

Kontis, D., Huddy, V., Reeder, C., Landau, S., Wykes, T., 2013. Effects of age and cognitive reserve on cognitive remediation therapy outcome in patients with schizophrenia. Am. J. Geriatr. Psychiatry 21, 218–230.

Kurtz, M.M., Seltzer, J.C., Shagan, D.S., Thime, W.R., Wexler, B.E., 2007. Computer-assisted cognitive remediation in schizophrenia: what is the active ingredient? Schizophr. Res. 89, 251–260.

Lindenmayer, J.-P., McGurk, S.R., Khan, A., Kaushik, S., Thanju, A., Hoffman, L., Valdez, G., Wance, D., Herrmann, E., 2013. Improving social cognition in schizophrenia: a pilot intervention combining computerized social cognition training with cognitive remediation. Schizophr. Bull. 39, 507–517. https://doi.org/10.1093/schbul/sbs120.

Loewy, R., Fisher, M., Schlosser, D.A., Biagianti, B., Stuart, B., Mathalon, D.H., Vinogradov, S., 2016. Intensive auditory cognitive training improves verbal memory in adolescents and young adults at clinical high risk for psychosis. Schizophr. Bull. 42, S118–S126. https://doi.org/10.1093/schbul/sbw009.

Lumsden, J., Edwards, E.A., Lawrence, N.S., Coyle, D., Munafò, M.R., 2016. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 4. https://doi.org/10.2196/games.5888.

Mahncke, H.W., Bronstone, A., Merzenich, M.M., 2006. Brain plasticity and functional losses in the aged: scientific bases for a novel intervention. Prog. Brain Res. 157, 81–109. https://doi.org/10.1016/S0079-6123(06)57006-2.

McCann, D.J., Petry, N.M., Bresell, A., Isacsson, E., Wilson, E., Alexander, R.C., 2015. Medication nonadherence, "professional subjects," and apparent placebo responders: overlapping challenges for medications development. J. Clin. Psychopharmacol. 35, 566–573. https://doi.org/10.1097/JCP.0000000000000372.

McGurk, S.R., Twamley, E.W., Sitzer, D.I., McHugo, G.J., Mueser, K.T., 2007. A meta-analysis of cognitive remediation in schizophrenia. Am. J. Psychiatry 164, 1791–1802.

Medalia, A., Saperstein, A.M., 2013. Does cognitive remediation for schizophrenia improve functional outcomes? Curr. Opin. Psychiatry 26, 151–157.

Merzenich, M.M., Van Vleet, T.M., Nahum, M., 2015. Brain plasticity-based therapeutics. Neuroplast. Neurorehabilitation 6.

Murthy, N.V., Mahncke, H., Wexler, B.E., Maruff, P., Inamdar, A., Zucchetto, M., Lund, J., Shabbir, S., Shergill, S., Keshavan, M., et al., 2012. Computerized cognitive remediation training for schizophrenia: an open label, multi-site, multinational methodology study. Schizophr. Res. 139, 87–91.

Nuechterlein, K.H., Green, M.F., Kern, R.S., Baade, L.E., Barch, D.M., Cohen, J.D., Essock, S., Fenton, W.S., Frese III, F.J., Gold, J.M., et al., 2008. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. Am. J. Psychiatry 165, 203.

Rass, O., Forsyth, J.K., Bolbecker, A.R., Hetrick, W.P., Breier, A., Lysaker, P.H., O'Donnell, B.F., 2012. Computer-assisted cognitive remediation for schizophrenia: a randomized single-blind pilot study. Schizophr. Res. 139, 92–98.

Saperstein, A.M., Medalia, A., 2015. The role of motivation in cognitive remediation for people with schizophrenia. In: Simpson, E.H., Balsam, P.D. (Eds.), Behavioral neuroscience of motivation, current topics in behavioral neurosciences. Springer International Publishing, pp. 533–546 https://doi.org/10.1007/7854_2015_373.

Surti, T.S., Corbera, S., Bell, M.D., Wexler, B.E., 2011. Successful computer-based visual training specifically predicts visual memory enhancement over verbal memory improvement in schizophrenia. Schizophr. Res. 132, 131–134.

Szymczynska, P., Walsh, S., Greenberg, L., Priebe, S., 2017. Attrition in trials evaluating complex interventions for schizophrenia: systematic review and meta-analysis. J. Psychiatr. Res. 90, 67–77.

Tallal, P., Piercy, M., 1973. Defects of non-verbal auditory perception in children with developmental aphasia. Nature 241, 468–469.

Ventura, J., Reise, S.P., Keefe, R.S., Baade, L.E., Gold, J.M., Green, M.F., Kern, R.S., Mesholam-Gately, R., Nuechterlein, K.H., Seidman, L.J., et al., 2010. The Cognitive Assessment Interview (CAI): development and validation of an empirically derived, brief interview-based measure of cognition. Schizophr. Res. 121, 24–31.

Vinogradov, S., Fisher, M., Warm, H., Holland, C., Kirshner, M.A., Pollock, B.G., 2009. The cognitive cost of anticholinergic burden: decreased response to cognitive training in schizophrenia. Am. J. Psychiatry 166, 1055–1062. https://doi.org/10.1176/appi.ajp.2009.09010017.

Ware Jr., J.E., Kosinski, M., Keller, S.D., 1996. A 12-Item Short-Form health survey: construction of scales and preliminary tests of reliability and validity. Med. Care 34, 220–233.

Wykes, T., Huddy, V., Cellard, C., McGurk, S.R., Czobor, P., 2011. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. Am. J. Psychiatry 168, 472–485.