Schizophrenia Research 158 (2014) 195–203

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Test–retest reliability of N400 event-related brain potential measures in a word-pair semantic priming paradigm in patients with schizophrenia

Jenna E. Boyd [a,b,c,*], Iulia Patriciu [b], Margaret C. McKinnon [a,b,c], Michael Kiang [a,b,d]

[a] Department of Psychiatry and Behavioural Neurosciences, McMaster University, Ontario L8S 4K1, Canada
[b] St. Joseph's Healthcare Hamilton, Hamilton, Ontario L8N 3K7, Canada
[c] Homewood Research Institute, Guelph, Ontario N1E 6K9, Canada
[d] Centre for Addiction and Mental Health, Toronto, Ontario M5T 1R8, Canada

ARTICLE INFO

Article history:
Received 26 March 2014
Received in revised form 11 June 2014
Accepted 16 June 2014
Available online 9 July 2014

Keywords:
Schizophrenia
Cognition
Semantics
Neurophysiology
Event-related potentials
Test–retest reliability

ABSTRACT

The N400 event-related brain potential (ERP), a negative voltage deflection occurring approximately 400 ms after onset of any meaningful stimulus, is reduced in amplitude when the stimulus is preceded by related context. Previous work has found this N400 semantic priming effect to be decreased in schizophrenia, suggesting impairment in using meaningful context to activate related concepts in semantic memory. Thus, N400 amplitude may be a useful biomarker of abnormal semantic processing and its response to treatment in schizophrenia. To help assess the validity of N400 amplitude as a longitudinal measure in schizophrenia, we evaluated its test–retest reliability. ERPs were recorded in sixteen schizophrenia patients who viewed prime words, each followed at 300- or 750-ms stimulus-onset asynchrony (SOA) by a target that was either a related or unrelated word, or non-word. Participants' task was to indicate whether or not the target was a real word. They were retested on the same procedure one week later. Test–retest reliability was assessed by calculating Pearson's *r* and intraclass correlation coefficients (ICCs) across timepoints for N400 amplitudes for related and unrelated targets, at each SOA. Consistent with previous results, there were no significant differences between patients' N400 amplitudes for related and unrelated targets, at any SOA/timepoint combination. Pearson's *r* and ICCs for N400 amplitudes at Fz across timepoints were significant for both target types at each SOA (ranges: *r* 0.52–0.64, ICC 0.52–0.63; all $p < .04$). The results suggest potential utility of N400 amplitude as a longitudinal neurophysiological biomarker of semantic processing abnormalities in schizophrenia.

1. Introduction

Meaningful stimuli — such as words, pictures, or environmental sounds — facilitate or prime the processing of related stimuli. For example, after seeing the word SHIP, people are faster to recognize the related word BOAT than the unrelated word BIRD. These *semantic priming effects* can be understood within a network model of long-term semantic memory, our store of knowledge about the world. In this type of model, meaningful concepts are represented as nodes in a neural network, with connections between nodes of related concepts (Collins and Loftus, 1975). When a particular node (in the example above, our mental concept of a ship) is activated (i.e., by reading the word SHIP), this activation is thought to spread to related concepts, thereby facilitating processing of their corresponding stimuli.

A neurophysiological index of semantic priming is the N400 scalp-recorded event-related brain potential (ERP). The N400 is a negative-going ERP waveform component peaking between 300 and 500 ms following the onset of any meaningful stimulus (Kutas and Federmeier, 2011). Its amplitude is reduced (less negative) when the eliciting stimulus is more related to preceding stimuli (Kutas and Hillyard, 1980, 1984, 1989; Chwilla et al., 1995). Within a network model of semantic memory, these *N400 semantic priming effects* (N400 amplitude difference between unrelated and related targets) reflect greater activation of concepts that are more related to preceding context. Thus, N400 semantic priming effects have been used as a neurophysiological probe of abnormal semantic activation in clinical populations (reviewed by Duncan et al., 2009).

In patients with schizophrenia, a large number of studies have provided evidence of larger than normal N400s in response to target stimuli that are related to preceding context, and/or smaller than normal N400 semantic priming effects (Bobes et al., 1996; Strandburg et al., 1997; Barch et al., 1996; Nestor et al., 1997; Ohta et al., 1999; Condray et al., 2003; Kostova et al., 2003; Iakimova et al., 2005; Kostova et al., 2005; Ditman and Kuperberg, 2007; Kiang et al., 2008; Salisbury, 2008;

* Corresponding author at: Room G137C, St. Joseph's Healthcare Hamilton, West 5th Campus, 100 West 5th St., Hamilton, ON L8N 3K7, Canada. Tel.: +1 905 522 1155x36326.
E-mail address: boydje2@mcmaster.ca (J.E. Boyd).

http://dx.doi.org/10.1016/j.schres.2014.06.018
0920-9964/

Guerra et al., 2009; Condray et al., 2010; Mathalon et al., 2010; Kiang et al., 2011, 2012, 2014). These results suggest that persons with schizophrenia are impaired in using meaningful contextual stimuli to activate related concepts. Moreover, several studies have found a correlation between these N400 semantic priming deficits and patients' psychotic symptom severity, raising the possibility that deficits in activating contextually related concepts may underlie the development and maintenance of delusions (Salisbury et al., 2000; Kiang et al., 2007, 2008).

In contrast, some other studies of schizophrenia patients have found smaller than normal N400 amplitudes to target stimuli related to preceding prime stimuli, and increased N400 semantic priming effects (Mathalon et al., 2002; Kreher et al., 2008; Salisbury, 2008; Kreher et al., 2009). Importantly, this pattern appears specific to short prime–target stimulus-onset asynchronies (SOAs) of less than approximately 300 ms, and patients with disorganized speech; and is thus thought to reflect an excess of rapid spread of activation in the semantic network of disorganized patients in particular (Ditman and Kuperberg, 2007; Salisbury, 2008; Kreher et al., 2009). Thus, this "hyperpriming" is not necessarily mutually exclusive with the presence of semantic priming deficits over longer time intervals in schizophrenia patients more generally.

Given that N400 semantic priming deficits (at least at SOAs of 300 ms or longer) have reliably been found in patients with schizophrenia, these deficits may be potentially useful as a biomarker of semantic processing abnormalities in schizophrenia, and of their response to treatment in clinical trials (Condray et al., 2003; Kiang et al., 2012). One necessary step in further validating putative biomarkers of schizophrenia is to establish their test–retest reliability both in normal individuals and in persons with schizophrenia (Cho et al., 2005). Previously, Kiang et al. (2013) found excellent test–retest reliability in healthy individuals for N400 amplitudes to both related and unrelated targets, and for N400 semantic priming effects, over a one-week period. In the present study we aimed to examine test–retest reliability of these N400 measures in patients with schizophrenia.

To this end, we recorded ERPs in schizophrenia patients while they viewed the same sequence of prime words, each followed by a target that was either a related or an unrelated word or a nonword, in two sessions one week apart. Although the procedure was otherwise similar to that used by Kiang et al. (2013) to examine N400 test–retest reliability in healthy individuals, that study used a 750-ms prime–target SOA, whereas in the present study we presented prime–target pairs at SOAs of both 300 and 750 ms. In light of previous work suggesting that the nature or magnitude of schizophrenia patients' N400 semantic priming abnormalities may depend on the interval elapsed since the prime stimulus, we added the shorter SOA because of the possibility that patients' N400 measures or their reliability might vary with SOA.

2. Methods

2.1. Participants

Participants included 16 outpatients with schizophrenia ($n = 11$) or schizoaffective disorder ($n = 5$). Patients were recruited in Hamilton, Ontario, Canada from two outpatient clinics specializing in treatment and rehabilitation of schizophrenia. All participants gave informed written consent. The protocol was approved by the St. Joseph's Healthcare Hamilton Research Ethics Board. Participants received cash compensation.

Participants were screened diagnostically with the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). Diagnostic and Statistical Manual of Mental Disorders (4th edition; DSM-IV) diagnoses were established using a best-estimate approach based on the MINI and information from medical records and clinician reports. Exclusion criteria included: current manic or depressive episode, lifetime substance dependence, or substance abuse in the past six months; exposure to a language other than English at home as a child; history of reading difficulties; and visual impairment. Table 1 shows group demographic characteristics, and total and factor scores (Miller et al., 1993) on the Scale for Assessment of Negative Symptoms (SANS; Andreasen, 1984a) and Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984b).

Table 1
Demographic and clinical characteristics of the study participants with mean and SD (range provided in brackets).

	Participants ($n = 16$)
Age, years	45.2 ± 8.1 (28–57)
Sex	5 females, 11 males
Handedness (Oldfield, 1971)	2 left, 13 right
Parental socioeconomic status (Blishen et al., 1987)	41.6 ± 4.8 (30.9–48.4)
Years of education	13.8 ± 2.0 (11–17)
National Adult Reading Test (O'Carroll et al., 1992) estimated IQ	107.7 ± 12.4 (87.2–124.7)
SANS total score	8.3 ± 4.7 (0–17)
SAPS total score	4.2 ± 3.8 (0–12)
SANS/SAPS factor scores	
Negative	5.5 ± 3.0 (0–11)
Positive	2.6 ± 2.8 (0–8)
Disorganized	1.6 ± 1.4 (0–4)

2.2. Stimuli

Stimuli included 80 related (e.g., METAL–STEEL) and 80 unrelated (DONKEY–PURSE) prime–target word pairs. For each related pair, the target was among the words most commonly given as associates to the prime in the University of South Florida word-association norms (Nelson et al., 1999); mean response probability of related targets (i.e. proportion of individuals producing that word in response to the prime) was 0.42 (SD = 0.23). For each unrelated pair, prime and target were not associates in the norms. Targets were matched across related and unrelated conditions for mean length and log-transformed frequency, as were primes (Francis and Kucera, 1982). Stimuli also included 160 word–nonword prime–target pairs (DRESS–ZORES), whose targets were pronounceable nonwords. No word occurred more than once among the stimuli. The stimulus list included all prime–target pairs in a fixed randomized order, in four blocks of 80 trials each. Two prime–target SOAs were used: 300 and 750 ms. There were two versions of the list such that the order of SOAs was counterbalanced across blocks. Specifically, in version A, SOA was 300 ms in blocks 1 and 2, and 750 ms in blocks 3 and 4; while in version B, the order was reversed. Each version was presented to half the participants.

2.3. Task

In an electrically shielded, sound-attenuated chamber, participants were seated 100 cm in front of a video monitor on which stimuli were visually presented, with each letter subtending on average 0.36° of the visual angle horizontally, and up to 0.55° vertically. Words were displayed in yellow letters on a black background.

Table 2
Mean percentage of correct lexical-decision responses ($n = 16$ participants) by experimental session, SOA, and target condition (SDs in parentheses).

	Time 1		Time 2	
	Short SOA	Long SOA	Short SOA	Long SOA
Related	97.4(4.3)	95.7(8.2)	98.9(2.8)	97.9(2.6)
Unrelated	96.8(5.5)	92.8(10.2)	95.4(8.9)	94.8(15.9)
Nonwords	96.9(4.2)	95.7(6.3)	98.4(1.8)	97.3(3.8)

Table 3
Mean percentage of accepted ERP trials ($n = 16$ participants) by experimental session, SOA, and target condition (SDs in parentheses).

	Time 1		Time 2	
	Short SOA	Long SOA	Short SOA	Long SOA
Related	88.9(8.5)	84.6(14.3)	84.0(13.7)	85.6(11.0)
Unrelated	90.3(7.4)	84.9(14.4)	85.4(16.1)	86.2(11.0)

Each participant was presented with the stimulus list, with short rest breaks between blocks. Each trial consisted of a: (a) row of preparatory fixation crosses (duration 500 ms); (b) blank screen (250 ms); (c) prime word (175 ms); (d) blank screen (125 ms or 575 ms, resulting in a 300- or 750-ms prime–target SOA, respectively); (e) target (250 ms); (f) blank screen (1250 ms); (g) the prompt *Yes or No?* until participants responded via a button-press; and (h) blank screen (3000 ms) until the onset of the next trial. All stimuli were centrally presented.

At the prompt, participants were required to press one of two buttons, positioned under their right and left thumbs, respectively, to indicate whether or not the target was a word. One button (labeled "Yes") signaled that it was a word, while the other button (labeled "No") signaled it was a nonword; assignment of buttons was counterbalanced across participants.

Each participant repeated the experimental procedure on two different occasions (Time 1 and Time 2) separated by approximately one week. The mean interval was 6.5 days (SD = 1.1, range 4–8).

2.4. Electroencephalographic data collection and analysis

During the experimental task, EEG was recorded using the ActiveTwo system (BioSemi BV, Amsterdam), from 32 sites approximately equally spaced across the scalp, positioned according to a modified International 10–20 System (Fp1-Fp2-AF3-AF4-F7-F3-Fz-F4-F8-FC5-FC1-FC2-FC6-T7-T8-C3-Cz-C4-CP5-CP1-CP2-CP6-P7-P3-Pz-P4-P8-PO3-PO4-O1-Oz-O2). The EEG was referenced to a left parietal Common Mode Sense (CMS) active electrode and a right parietal Driven Right Leg (DRL) passive electrode, which form a feedback loop driving the average potential across the montage as close as possible to the amplifier zero. The EEG was continuously digitized at 512 Hz and low-pass filtered at 128 Hz. Blinks and eye movements were monitored via electrodes on the supraorbital and infraorbital ridges and on the outer canthi of both eyes. Offline, the EEG was re-referenced to the algebraic mean of the mastoids, and bandpassed at 0.01–100 Hz. Continuous data was algorithmically corrected for eye-blink artifact (Jung et al., 2000). ERPs

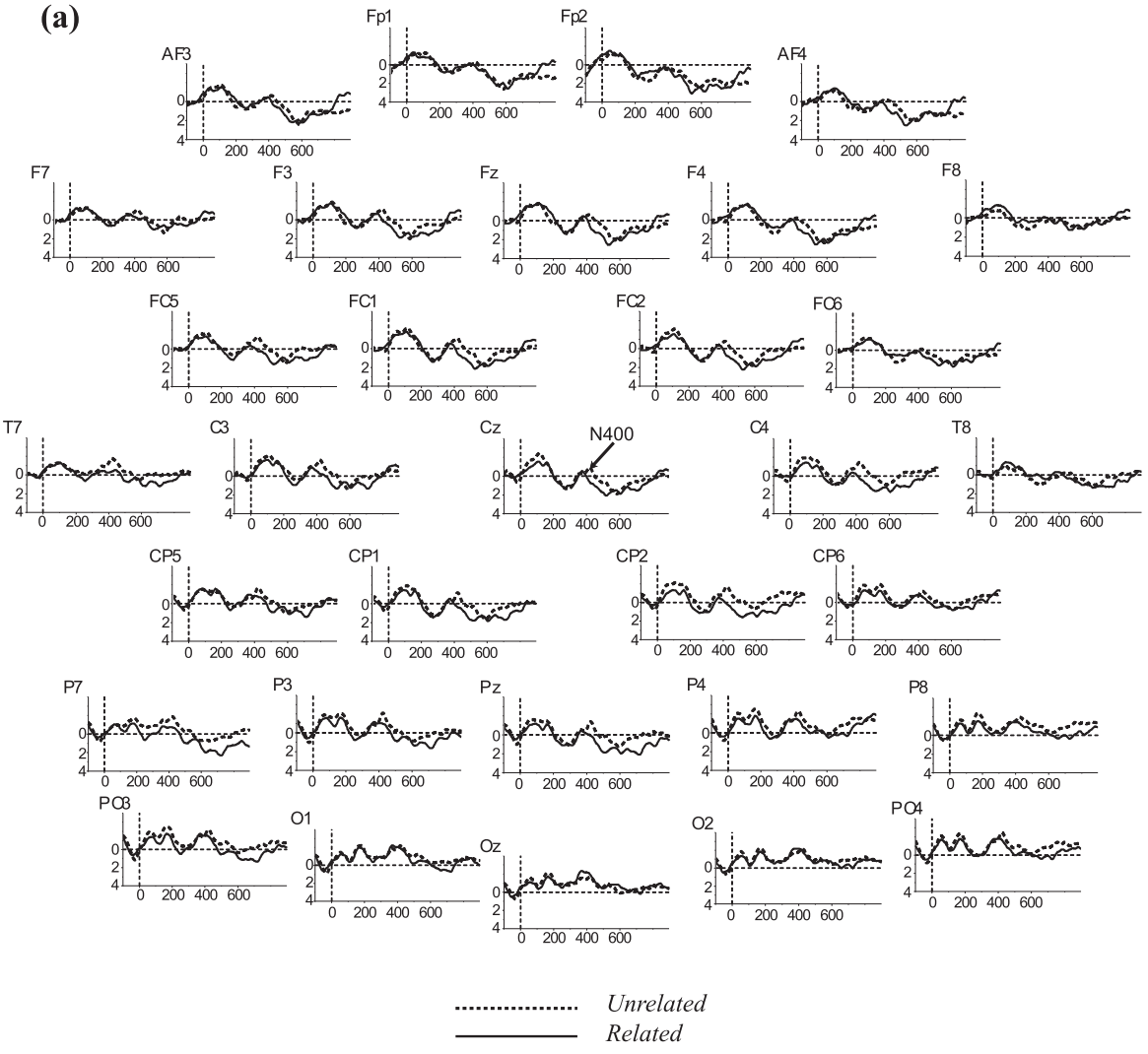

Fig. 1. Grand average ERPs to target words at all electrode sites, at (a) Time 1, short SOA, (b) Time 1, long SOA, (c) Time 2, short SOA, and (d) Time 2, long SOA. Negative voltage is plotted upward.

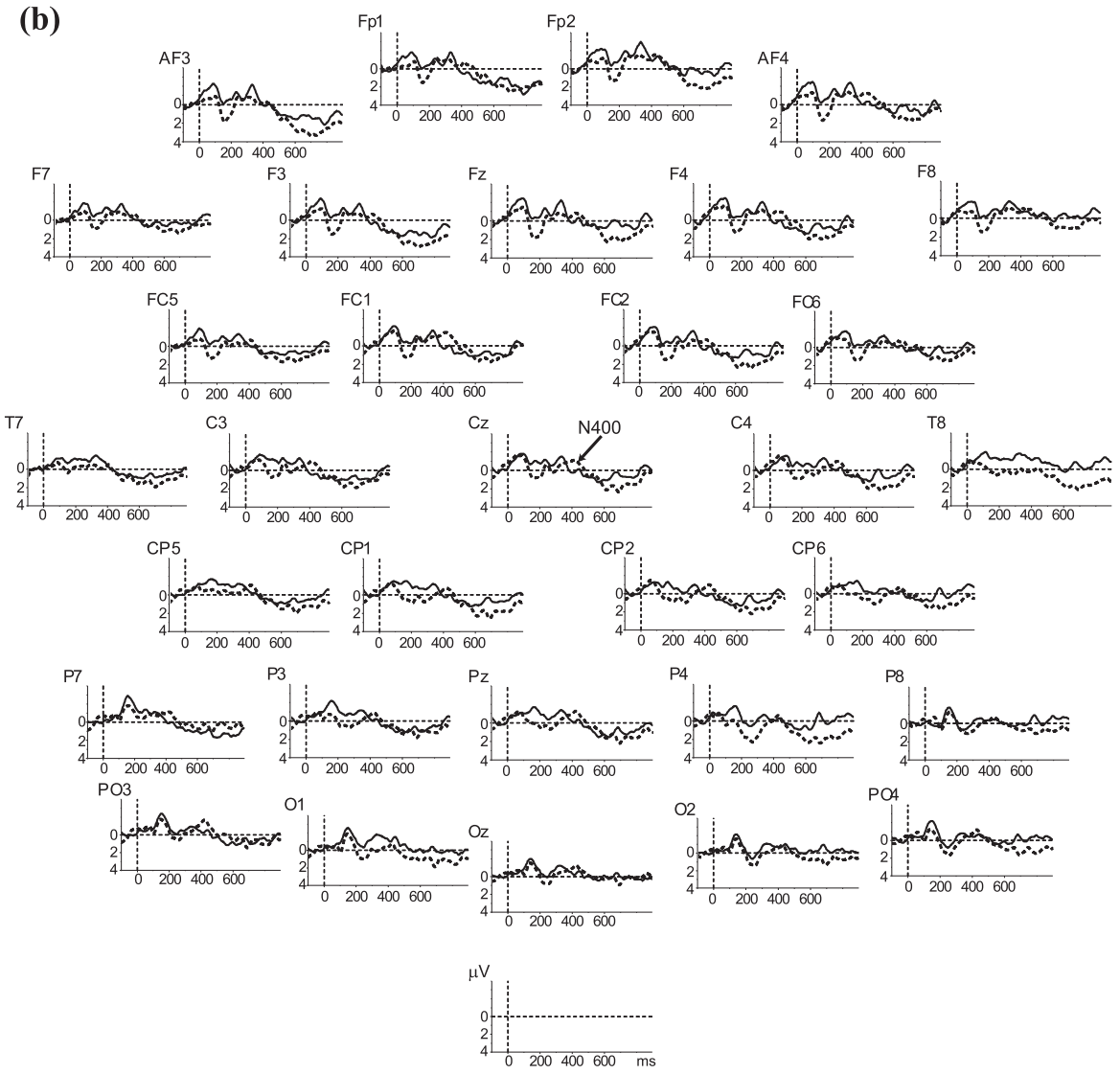

Fig. 1 (*continued*).

were computed for epochs from 100 ms pre-stimulus to 900 ms post-stimulus. Individual trials containing artifacts due to eye movement, excessive muscle activity or amplifier blocking were rejected off-line by visual inspection before time-domain averaging.

For each participant, separate ERP averages were obtained for trials with related and unrelated target words, after subtraction of the 100-ms prestimulus baseline. Consistent with established methods, N400 peak latency was defined as the interval between target onset and the largest negative peak from 200 to 600 ms post-stimulus and N400 amplitude was defined as the mean voltage from 200 to 500 ms post-stimulus (Kiang et al., 2012).

N400 semantic priming effects were defined as the mean voltage of the difference waveform obtained by subtracting the ERP average for related targets from the average for unrelated targets, from 200 to 500 ms post-stimulus.

2.5. Statistical analyses

p-Values in analyses of variance (ANOVAs) with within-subject factors are reported after Greenhouse–Geisser Epsilon correction. Pairwise comparisons of factor-level means were made with Tukey simultaneous comparisons, with a family confidence coefficient of 0.95. All *p*-values are two-tailed.

The percentage of correct responses was analyzed by repeated-measures ANOVA, with Target (related vs. unrelated vs. nonword), SOA (300 vs. 750 ms), and Time (1 vs. 2) as within-subject factors.

The percentage of accepted ERP trials was analyzed by repeated-measures ANOVA, with Target (related vs. unrelated), SOA (300 vs. 750 ms), and Time (1 vs. 2) as within-subjects factors.

N400 peak latency and mean amplitude were analyzed by repeated-measures ANOVA with Target (related vs. unrelated), SOA (300 vs. 750 ms), Electrode (13 levels: F3/FC1/C3/CP1/P3/Pz/P4/CP2/C4/FC2/F4/Fz/Cz, corresponding to a contiguous array of medial central sites where N400 effects are typically most prominent (Kutas and Federmeier, 2000), and Time (1 vs. 2) as within-subject factors. N400 semantic priming effects were analyzed by repeated-measures ANOVA with Electrode (13 levels, corresponding to the sites listed above), SOA (300 vs. 750 ms), and Time (1 vs. 2) as within-subject factors.

Across participants, for each SOA, pairwise correlations between values at Time 1 and Time 2 were calculated for N400 amplitudes for related targets, N400 amplitudes for unrelated targets, and N400 semantic priming effects; for each of the midline sites Fz (frontal), Cz (central)

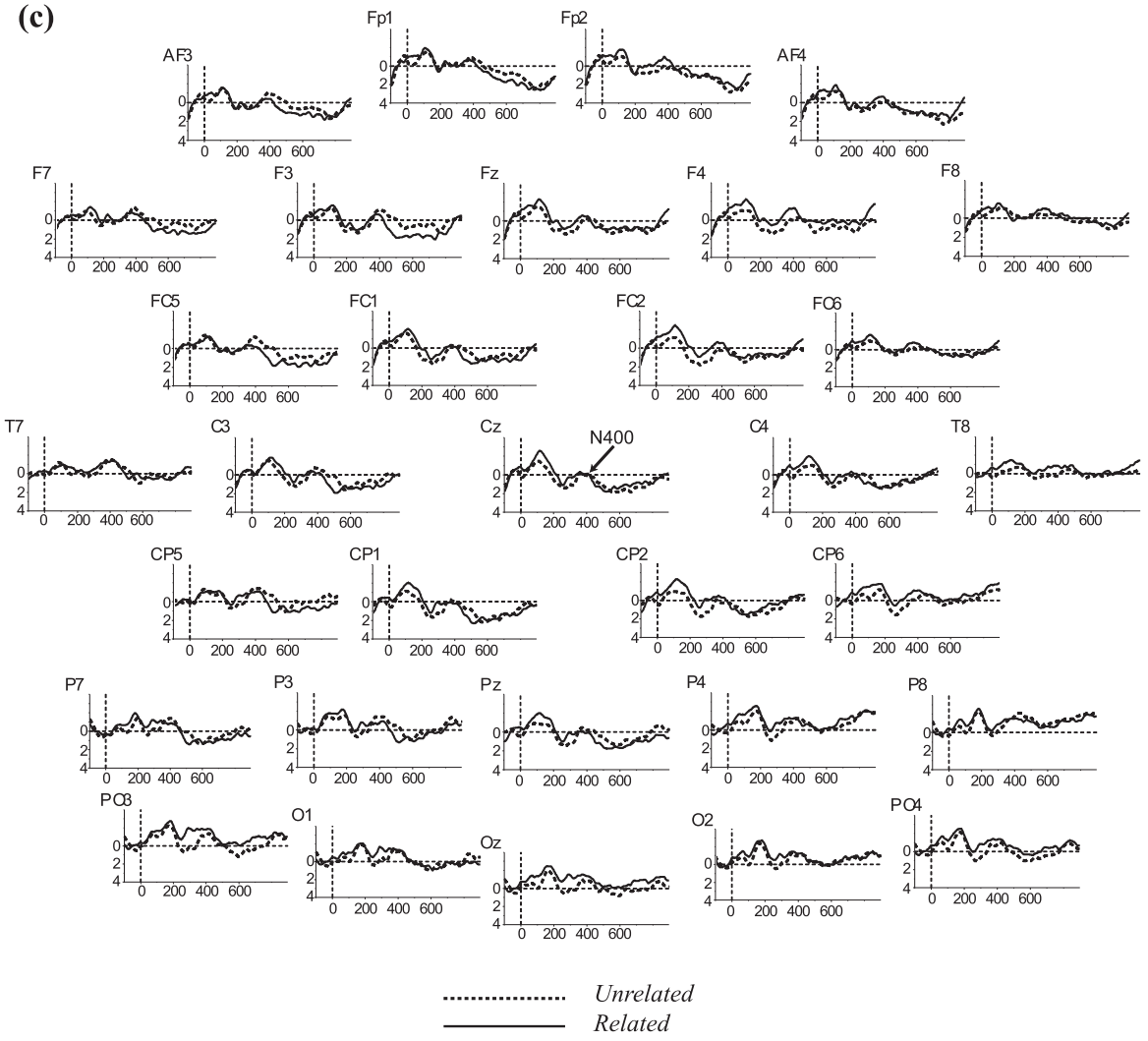

Fig. 1. (*continued*).

and Pz (parietal). As convergent measures of test–retest reliability, Pearson's correlation coefficients *r* and intraclass correlation coefficients (ICCs) were calculated.

In addition, to explore associations between N400 measures and different symptoms, we calculated pairwise correlations (Pearson's *r*) for each SOA between N400 amplitudes for related targets, N400 amplitudes for unrelated targets, and N400 semantic priming effects vs. SANS/SAPS factor scores. N400 measures were taken as the mean value over Time 1 and Time 2, at sites Fz and Cz, as these sites were found to have the highest test–retest reliability of those tested.

3. Results

3.1. Correct response rates

Participants' mean correct response rates are shown in Table 2. Overall, the high correct-response rates indicate that participants were attending to the stimuli and task. No significant effects of Target, SOA, or Time were found. There was a significant interaction of SOA × Time [$F(1, 15) = 5.05$, $p = 0.04$]; indicating that correct response percentage was lower for long-SOA trials at Time 1 than for short-SOA trials at Time 1, or short- or long-SOA trials at Time 2, all three of which did not differ significantly from one another.

3.2. Percentage of accepted ERP trials

The percentage of accepted ERP trials is shown in Table 3. ANOVA did not show any significant effects of Target, SOA, or Time, or interactions of these factors (all $p > 0.19$).

3.3. ERPs

Grand average ERPs for all electrodes are shown for each SOA at Time 1 and Time 2 in Fig. 1.

In the ANOVA of N400 peak latency, there were no significant effects of Target, SOA, Time, Electrode or significant interactions of these factors (all $p > 0.11$). Grand mean N400 peak latency was 358 ms (SD = 102, range: 203–594).

N400 mean amplitudes are shown in Fig. 2 and in Table 6. There were no significant effects of Target, SOA, Time, Electrode or significant interactions of these factors, on N400 mean amplitude (all $p > 0.19$). Values for N400 semantic priming effects are shown in Table 6. In the ANOVA of N400 semantic priming effects, there were no significant

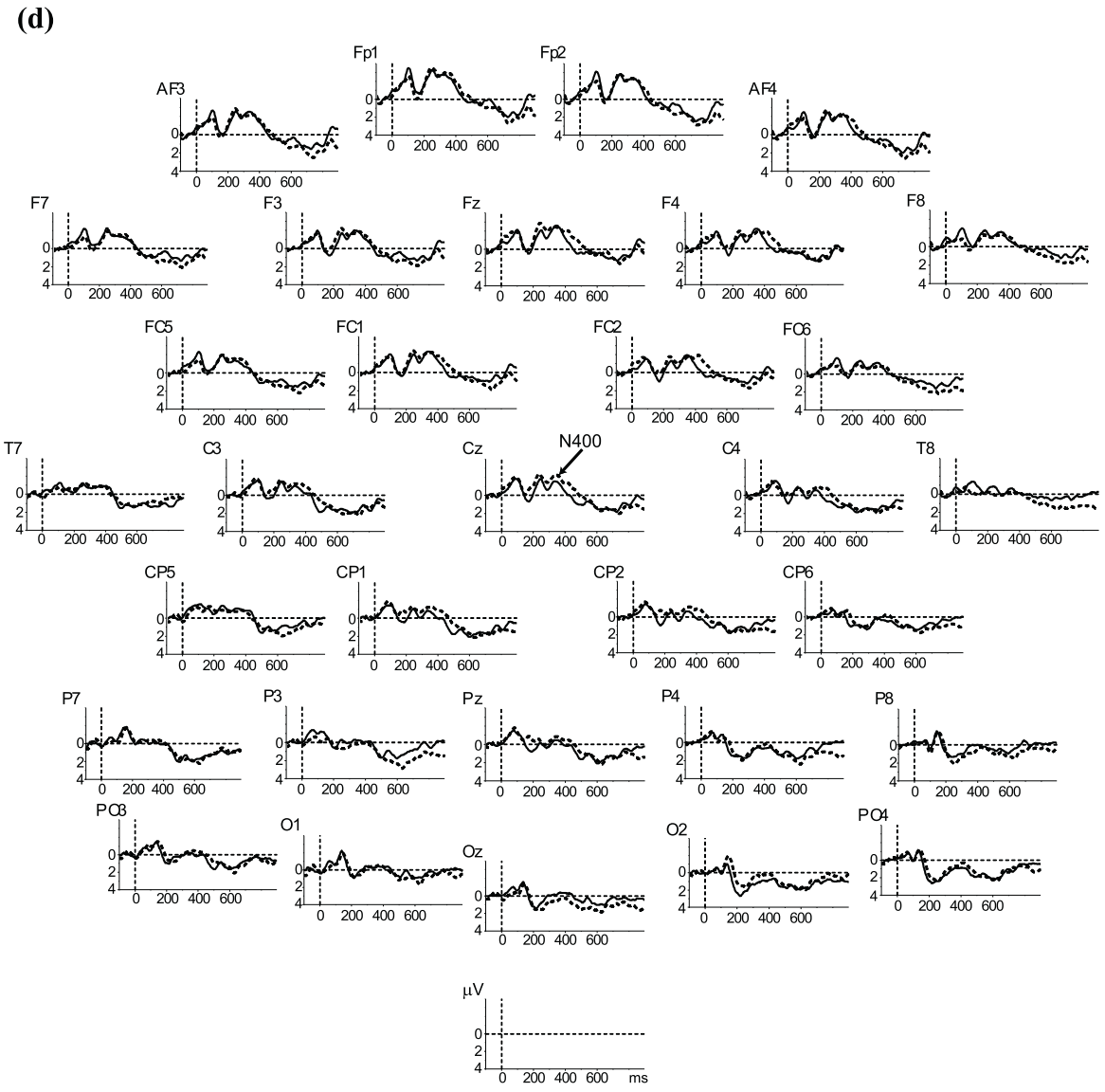

Fig. 1. (*continued*).

effects of SOA, Time or Electrode, or significant interactions of these factors (all $p > 0.20$).

3.4. Correlations between Time 1 and Time 2 for N400 measures

Pearson's *r* and ICCs for correlations between Time 1 and Time 2 for N400 measures are shown in Table 5. Scatterplots of these correlations at Fz are shown in Fig. 3. Significant correlations between Time 1 and Time 2 were found at Fz for all target type/SOA combinations; at Cz for all combinations except unrelated targets at the short SOA; and at Pz for unrelated targets at the short SOA and related targets at the long SOA.

3.5. Correlations between N400 measures and SANS/SAPS factor scores

Pearson's *r* for correlations between N400 measures (averaged across electrodes Fz and Cz and Time 1 and Time 2) and SANS/SAPS factor scores are shown in Table 4. Smaller N400 semantic priming effects were significantly correlated with SAPS/SANS Positive factor scores at the long SOA ($r = -0.66$, $p = 0.005$), indicating that patients with more severe psychotic symptoms had more similar N400 amplitudes to related and unrelated targets. No other correlations were significant (all $p > 0.10$).

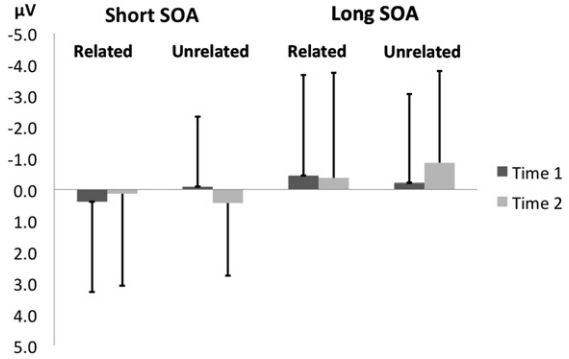

Fig. 2. Mean N400 amplitudes for related and unrelated targets and short and long SOAs (averaged across 13 medial central sites: F3, FC1, C3, CP1, P3, Pz, P4, CP2, C4, FC2, F4, Fz, Cz) at Time 1 and Time 2. Error bars indicate the standard deviation.

Table 4
Pairwise correlations (Pearson's *r*) for each SOA, averaged across Time 1 and Time 2 for electrodes Fz and Cz between N400 amplitudes for related targets and unrelated targets, and N400 semantic priming effects vs. SANS/SAPS factor scores.

	N400 mean amplitudes				Semantic priming effects	
	Related targets, short SOA	Related targets, long SOA	Unrelated targets, short SOA	Unrelated targets, long SOA	Short SOA	Long SOA
SANS/SAPS negative	−0.25	−0.41	−0.36	−0.22	−0.02	0.39
SANS/SAPS positive	0.31	0.39	0.41	0.02	−0.41	−0.66**
SANS/SAPS disorganized	0.08	−0.06	−0.10	−0.24	−0.23	−0.25

** $p < 0.01$.

Table 5
Pearson's *r* and intraclass correlation coefficients (ICC) between mean N400 amplitudes at Time 1 and Time 2 for each SOA/target type combination at electrode sites Fz, Cz, and Pz ($n = 16$ participants).

Electrode	N400 amplitudes, related targets, short SOA		N400 amplitudes, related targets, long SOA		N400 amplitudes, unrelated targets, short SOA		N400 amplitudes, unrelated targets, long SOA		N400 priming effects, short SOA		N400 priming effects, long SOA	
	r	ICC	*r*	ICC	*r*	ICC	*r*	ICC	*r*	ICC	*r*	ICC
Cz	0.60*	0.60**	0.67**	0.66**	0.40	0.39	0.54*	0.53*	0.02	0.02	−0.12	−0.10
Fz	0.52*	0.52*	0.64**	0.63**	0.61*	0.61**	0.61*	0.59**	0.08	0.08	0.28	0.25
Pz	0.30	0.30	0.50*	0.50*	0.43	0.43*	0.36	0.35	−0.041	−0.04	0.10	0.07

* $p < 0.05$.
** $p < 0.01$.

Table 6
Mean N400 mean amplitudes and semantic priming effects ($n = 16$) for electrodes Fz, Cz, and Pz for each SOA/target combinations at Time 1 and Time 2 (SDs in parentheses). Values are in μV.

	N400 amplitudes, related targets, short SOA		N400 amplitudes, related targets, long SOA		N400 amplitudes, unrelated targets, short SOA		N400 amplitudes, unrelated targets, long SOA		N400 priming effects, short SOA		N400 priming effects, long SOA	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Fz	0.62(2.89)	0.18(2.93)	−0.71(3.96)	−1.29(4.28)	0.29(2.33)	0.60(2.27)	−0.40(2.96)	−2.03(3.83)	−0.33(2.39)	0.41(2.53)	0.32(2.40)	−0.74(3.83)
Cz	0.69(2.55)	−0.45(3.55)	0.27(2.49)	0.69(2.55)	−0.50(4.04)	−0.63(3.56)	−0.45(3.56)	−1.50(2.98)	−0.41(2.74)	0.24(2.86)	0.04(1.65)	−0.88(3.14)
Pz	0.62(3.23)	0.45(3.14)	−0.54(2.64)	0.21(2.87)	−0.01(2.39)	0.50(2.32)	−0.27(2.64)	−0.29(2.23)	−0.62(2.59)	0.05(2.90)	0.28(1.29)	−0.50(3.31)

4. Discussion

This study aimed to assess test–retest reliability of N400 ERP measures in patients with schizophrenia in a word-pair semantic priming paradigm. We recorded ERPs in schizophrenia patients while they viewed the same series of prime–target word-pairs in two sessions approximately one week apart. Prime words were followed at either a short (300-ms) or a long (750-ms) SOA by either a related or an unrelated target word, or a pronounceable nonword. Participants' task was to indicate via button-press whether or not the target was a real word. Across both SOAs, our patient sample exhibited no significant difference in N400 amplitudes elicited by related vs. unrelated target words, consistent with previous findings indicating smaller than normal N400 semantic priming effects (amplitude differences between unrelated and related targets) in schizophrenia over comparable SOAs (Condray et al., 2003; Kiang et al., 2008, 2014).

In the current study, schizophrenia patients demonstrated good test–retest reliability for N400 amplitudes for both related and unrelated targets, at both SOAs. For example, ICCs for these measures at Fz ranged from 0.52 to 0.63, which is considered good in the context of clinical trials (Fleiss, 1999; Manoach et al., 2001; Schmidt et al., 2012).

In contrast, patients did not exhibit significant test–retest correlations for N400 semantic priming effects between Time 1 and Time 2. This finding diverged from results obtained in healthy individuals (Kiang et al., 2013), who exhibited excellent test–retest reliability for N400 semantic priming effects as well as for N400 amplitudes to related and unrelated targets. Low test–retest reliability for schizophrenia patients' N400 semantic priming effects is likely attributable to the fact that these effects were not significantly different from zero in this group — i.e., there was no difference between N400 amplitudes for related and unrelated targets, suggesting that patients were processing these targets similarly at the neural level. In that case, in contrast to the situation in healthy individuals, the magnitude of a patient's N400 semantic priming effect at Time 2 would depend less on the magnitude of this effect at Time 1, and more on other factors that are random with respect to prime–target relatedness, reducing test–retest reliability.

Notably, in the present study, smaller N400 semantic priming effects at the long SOA were significantly correlated with higher SAPS/SANS Positive factor scores (see Table 4). This result corroborates previous reports that found such a relationship in patients with schizophrenia (Salisbury et al., 2000; Kiang et al., 2007, 2008) and thus raises the possibility that semantic priming deficits in schizophrenia may be related to the development of delusions.

In a recent longitudinal study, Besche-Richard et al. (2014) reported that N400 semantic priming deficits in patients with schizophrenia were not stable, but improved, over one-year follow-up. Unlike controls, who exhibited similar N400 semantic priming effects at baseline and one year, the patients exhibited such effects only at one year. The patient group's improvement in N400 priming occurred in parallel with a mean reduction in clinical symptoms (although the authors did not test for associations between improvements in N400 priming and symptoms). These results suggest that, in addition to having good reliability, N400 amplitude measures may be a biomarker of clinical improvements and treatment effectiveness.

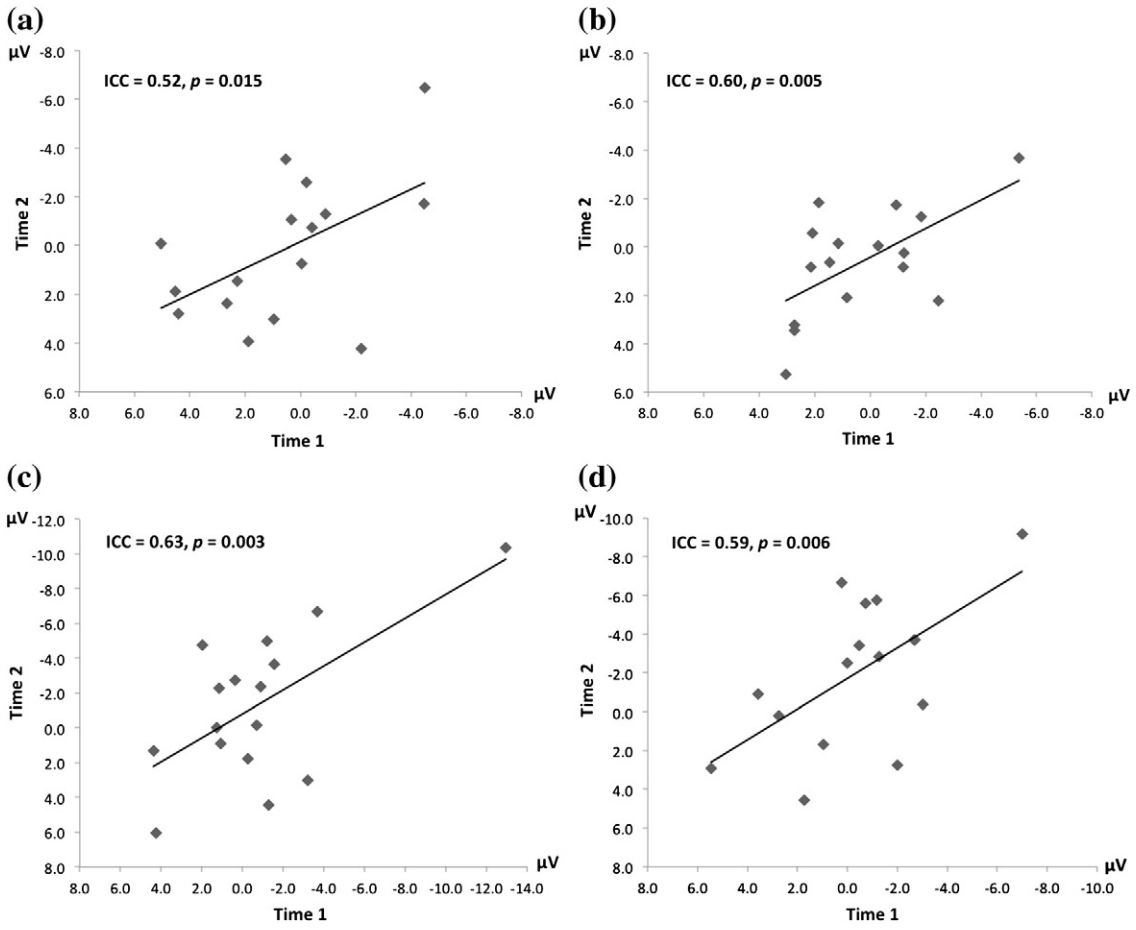

Fig. 3. Scatterplots of values at Time 2 vs. values at Time 1 at electrode Fz for: (a) N400 amplitudes for related targets at the short SOA; (b) N400 amplitudes for unrelated targets at the short SOA; (c) N400 amplitudes for related targets at the long SOA; and (d) N400 amplitudes for unrelated targets at the long SOA.

A limitation of this study is that the patient sample was not directly compared with healthy controls on N400 semantic priming effects. Although the literature includes substantial evidence for smaller than normal N400 semantic priming effects in schizophrenia patients compared to controls (Bobes et al., 1996; Strandburg et al., 1997; Ohta et al., 1999; Condray et al., 2003; Kostova et al., 2003; Iakimova et al., 2005; Kostova et al., 2005; Ditman and Kuperberg, 2007; Kiang et al., 2008; Salisbury, 2008; Guerra et al., 2009; Condray et al., 2010; Mathalon et al., 2010; Kiang et al., 2011, 2012, 2014), our findings do not allow for a definitive conclusion that the current sample had smaller than normal N400 semantic priming effects (as suggested by an absence of significant differences between N400 mean amplitude for contextually related and unrelated target words).

The good test–retest reliability of N400 mean amplitudes in a word-pair priming paradigm in this study supports their potential clinical utility as longitudinal biomarkers of semantic processing dysfunction in patients with schizophrenia. In particular, test–retest reliabilities of N400 amplitudes at midline frontal and central sites appeared to be within the range of values that have been obtained in healthy individuals for other proposed cognitive ERP biomarkers for schizophrenia, such as the amplitudes of the mismatch negativity (Kathmann et al., 1999; Jemel et al., 2002; Hall et al., 2006; Lew et al., 2007) and P300 (Fabiani et al., 1987; Kinoshita et al., 1996; Hall et al., 2006; Lew et al., 2007). Although N400 amplitude test–retest reliabilities in the current study appear to fall in the lower part of that range, those other studies examined healthy individuals whereas, in general, reliabilities for patient populations may be lower due to noisier ERP data (Luck, 2005). Contrary to this conjecture, however, test–retest ICCs for P300 amplitude over a mean of 11 days ranged from 0.77 to 0.85 (depending on electrode site) in patients with primary psychotic disorders, compared to 0.74 to 0.77 in healthy control participants (Simons et al., 2011). There have been few studies of short-term (days to weeks) test–retest reliability of cognitive ERP components in psychotic disorders, even though characterizing reliability has been identified as a research priority in developing appropriate ERP and other biomarkers for schizophrenia (Cho et al., 2005; Barch and Mathalon, 2011; Light and Braff, 2005). Further research on test–retest reliabilities of different ERP components in schizophrenia, encompassing different age groups, stimulus parameters, and test–retest intervals, would improve our understanding of how these values compare with each other and with their counterparts in healthy individuals.

Role of the funding source

The sponsor (the Canadian Institutes for Health Research) had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Contributors

Michael Kiang and Iulia Patriciu contributed to designing the study. Iulia Patriciu managed the recruitment of participants and collection of data. Jenna Boyd and Michael Kiang managed the literature searches, and Jenna Boyd performed the data analyses and wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflicts of interest

Michael Kiang has served as a collaborator on studies funded by Roche. The other authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Canadian Institutes for Health Research (#112359). The authors thank the staff of the Community Schizophrenia Service and Hamilton Program for Schizophrenia for the assistance with participant recruitment.

References

Andreasen, N.C., 1984a. Scale for the Assessment of Negative Symptoms. University of Iowa, Iowa City, IA.

Andreasen, N.C., 1984b. Scale for the Assessment of Positive Symptoms. University of Iowa, Iowa City, IA.

Barch, D.M., Mathalon, D.H., 2011. Using brain imaging measures in studies of procognitive pharmacologic agents in schizophrenia: psychometric and quality assurance considerations. Biol. Psychiatry 70 (1), 13–18.

Barch, D.M., Cohen, J.D., Servan-Schreiber, D., Steingard, S., Steinhauer, S.S., van Kammen, D.P., 1996. Semantic priming in schizophrenia: an examination of spreading activation using word pronunciation and multiple SOAs. J. Abnorm. Psychol. 105 (4), 592–601.

Besche-Richard, C., Iakimova, G., Hardy-Bayle, M.C., Passerieux, C., 2014. Behavioral and brain measures (N400) of semantic priming in patients with schizophrenia: test–retest effect in a longitudinal study. Psychiatry Clin. Neurosci. 68, 365–373.

Blishen, B.R., Carroll, W.K., Moore, C., 1987. The 1981 socioeconomic index for occupations in Canada. Can. Rev. Sociol. Anthropol. 24, 465–488.

Bobes, M.A., Lei, Z.X., Ibanez, S., Yi, H., Valdes-Sosa, M., 1996. Semantic matching of pictures in schizophrenia: a cross-cultural ERP study. Biol. Psychiatry 40 (3), 189–202.

Cho, R.Y., Ford, J.M., Krystal, J.H., Laruelle, M., Cuthbert, B., Carter, C.S., 2005. Functional neuroimaging and electrophysiology biomarkers for clinical trials for cognition in schizophrenia. Schizophr. Bull. 31 (4), 865–869.

Chwilla, D.J., Brown, C.M., Hagoort, P., 1995. The N400 as a function of the level of processing. Psychophysiology 32, 274–285.

Collins, A.M., Loftus, E.F., 1975. A spreading-activation theory of semantic processing. Psychol. Rev. 82 (6), 407–428.

Condray, R., Siegle, G.J., Cohen, J.D., van Kammen, D.P., Steinhauer, S.R., 2003. Automatic activation of the semantic network in schizophrenia: evidence from event-related brain potentials. Biol. Psychiatry 54 (11), 1134–1148.

Condray, R., Siegle, G.J., Keshavan, M.S., Steinhauer, S.R., 2010. Effects of word frequency on semantic memory in schizophrenia: electrophysiological evidence for a deficit in linguistic access. Int. J. Psychophysiol. 75 (2), 141–156.

Ditman, T., Kuperberg, G.R., 2007. The time course of building discourse coherence in schizophrenia: an ERP investigation. Psychophysiology 44 (6), 991–1001.

Duncan, C.C., Barry, R.J., Connolly, J.F., Fischer, C., Michie, P.T., Naatanen, R., et al., 2009. Event related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clin. Neurophysiol. 120, 1883–1908.

Fabiani, M., Gratton, G., Karis, D., Donchin, E., 1987. Definition, identification and reliability of measurement of the P300 component of the event related brain potential. In: Ackles, P.K., Jennings, J.R., Coles, M.G.H. (Eds.), Advances in Psychophysiology, vol. 2. JAI Press, Greenwich CT, pp. 1–78.

Fleiss, J.L., 1999. Design and Analysis of Clinical Experiments. Wiley and Sons, Hoboken, NJ.

Francis, W.N., Kucera, H., 1982. Frequency of Analysis of English Usage. Houghton Mifflin, Boston MA.

Guerra, S., Ibanez, A., Martin, M., Bobes, M.A., Reyes, A., Mendoza, R., Bravo, T., Dominguez, M., Sosa, M.V., 2009. N400 deficits from semantic matching of pictures in probands and first-degree relatives from multiplex schizophrenia families. Brain Cogn. 70 (2), 221–230.

Hall, M.H., Schulze, K., Rijsdijk, F., Picchioni, M., Ettinger, U., Bramon, E., et al., 2006. Heritability and reliability of P300, P50, and duration mismatch negativity. Behav. Genet. 36 (6), 845–857.

Iakimova, G., Passerieux, C., Laurent, J.P., Hardy-Bayle, M.C., 2005. ERPs of metaphoric, literal, and incongruous semantic processing in schizophrenia. Psychophysiology 42 (4), 380–390.

Jemel, B., Achenbach, C., Muller, B., Ropcke, B., Oades, R.D., 2002. Mismatch negativity results from bilateral asymmetric dipole sources in the frontal and temporal lobes. Brain Topogr. 15 (1), 13–27.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.K., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37 (2), 163–178.

Kathmann, N., Frodl-Bauch, T., Hegerl, U., 1999. Stability of the mismatch negativity under different stimulus and attention conditions. Clin. Neurophysiol. 110 (2), 317–323.

Kiang, M., Kutas, M., Light, G.A., Braff, D.L., 2007. Electrophysiological insights into conceptual disorganization in schizophrenia. Schizophr. Res. 92 (1–3), 225–236.

Kiang, M., Kutas, M., Light, G.A., Braff, D.L., 2008. An event-related brain potential study of direct and indirect semantic priming in schizophrenia. Am. J. Psychiatry 165 (1), 74–81.

Kiang, M., Christensen, B.K., Zipursky, R.B., 2011. Depth-of-processing effects on semantic activation deficits in schizophrenia: an electrophysiological investigation. Schizophr. Res. 133 (1–3), 91–98.

Kiang, M., Christensen, B.K., Kutas, M., Zipursky, R.B., 2012. Electrophysiological evidence for primary semantic memory functional organization deficits in schizophrenia. Psychiatry Res. 196 (2–3), 171–180.

Kiang, M., Patriciu, I., Roy, C., Christensen, B.K., Zipursky, R.B., 2013. Test–retest reliability and stability of N400 effects in a word-pair semantic priming paradigm. Clin. Neurophysiol. 124, 667–674.

Kiang, M., Christenson, B.K., Zipursky, R.B., 2014. Event-related brain potential study of semantic priming in unaffected first-degree relatives of schizophrenia patients. Schizophr. Res. 153 (1–3), 78–86.

Kinoshita, S., Inoue, M., Maeda, H., Nakamura, J., Morita, K., 1996. Long-term patterns of change in ERPs across repeated measurements. Physiol. Behav. 60 (4), 1087–1092.

Kostova, M., Passerieux, C., Laurent, J.P., Saint-Georges, C., Hardy-Bayle, M.C., 2003. Functional analysis of the deficit in semantic context processes in schizophrenic patients: an event-related potentials study. Neurophysiol. Clin. 33 (1), 11–22.

Kostova, M., Passerieux, C., Laurent, J.P., Hardy-Bayle, M.C., 2005. N400 anomalies in schizophrenia are correlated with the severity of formal thought disorder. Schizophr. Res. 78 (2–3), 285–291.

Kreher, D.A., Holcomb, P.J., Goff, D., Kuperberg, G.R., 2008. Neural evidence for faster and further automatic spreading activation in schizophrenic thought disorder. Schizophr. Bull. 34 (3), 473–482.

Kreher, D.A., Goff, D., Kuperberg, G.R., 2009. Why all the confusion? Experimental task explains discrepant semantic priming effects in schizophrenia under "automatic" conditions: evidence from event-related potentials. Schizophr. Res. 111, 174–181.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4 (12), 463–470.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Ann. Rev. Psychol. 62, 621–647.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205.

Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.

Kutas, M., Hillyard, S.A., 1989. An electrophysiological probe of incidental semantic association. J. Cogn. Neurosci. 1, 38–49.

Lew, H.L., Gray, M., Poole, J.H., 2007. Temporal stability of auditory event-related potentials in healthy individuals and patients with traumatic brain injury. J. Clin. Neurophysiol. 24 (5), 392–397.

Light, G.A., Braff, M.D., 2005. Stability of mismatch negativity deficits and their relationship to functional impairments in chronic schizophrenia. Am. J. Psychiatry 162, 1741–1743.

Luck, S., 2005. An Introduction to the Event-related Potential Technique. MIT Press, Cambridge MA.

Manoach, D.S., Halpern, E.F., Kramer, T.S., Chang, Y., Goff, D.C., Rauch, S.L., et al., 2001. Test–retest reliability of a functional MRI working memory paradigm in normal and schizophrenic subjects. Am. J. Psychiatry 158 (6), 955–958.

Mathalon, D.H., Faustman, W.O., Ford, J.M., 2002. N400 and automatic semantic processing abnormalities in patients with schizophrenia. Arch. Gen. Psychiatry 59 (7), 641–648.

Mathalon, D.H., Roach, B.J., Ford, J.M., 2010. Automatic semantic priming abnormalities in schizophrenia. Int. J. Psychophysiol. 75 (2), 157–166.

Miller, D.D., Arndt, S., Andreasen, N.C., 1993. Alogia, attentional impairment, and inappropriate affect: their status in the dimensions of Schizophrenia. Compr. Psychiatry 34 (4), 221–226.

Nelson, D.L., McEvoy, C.L., Schreiber, T.A., 1999. The University of South Florida word association norms. www.usf.edu/FreeAssociation/.

Nestor, P.G., Kimble, M.O., O'Donnell, B.F., Smith, L., Niznikiewicz, M., Shenton, M.E., et al., 1997. Aberrant semantic activation in schizophrenia: a neurophysiological study. Am. J. Psychiatry 154 (5), 640–646.

O'Carroll, R., Walker, M., Dunan, J., Murray, C., Blackwood, C., Ebmeier, K.P., et al., 1992. Selecting controls for schizophrenia research studies: the use of the National Adult Reading Test (NART) is a measure of premorbid ability. Schizophr. Res. 8 (2), 137–141.

Ohta, K., Uchiyama, M., Matsushima, E., Toru, M., 1999. An event-related potential study in schizophrenia using Japanese sentences. Schizophr. Res. 40 (2), 159–170.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113.

Salisbury, D.F., 2008. Semantic activation and verbal working memory maintenance in schizophrenic thought disorder: insights from electrophysiology and lexical ambiguity. Clin. EEG Neurosci. 39 (2), 103–107.

Salisbury, D.F., O'Donnell, B.F., McCarley, R.W., Nestor, P.G., Shenton, M.E., 2000. Event related potentials elicited during a context-free homograph task in normal versus schizophrenic subjects. Psychophysiology 37 (4), 456–463.

Schmidt, L.A., Santesso, D.L., Miskovic, V., Mathewson, K.J., McCabe, R.E., Antony, M.M., et al., 2012. Test–retest reliability of regional electroencephalogram (EEG) and cardiovascular measures in social anxiety disorder (SAD). Int. J. Psychophysiol. 84 (1), 65–73.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (Suppl. 20), 22–33.

Simons, C.J.P., Sambeth, A., Krabbendam, L., Pfleifer, S., van Os, J., Riedel, W.J., 2011. Auditory P300 and N100 components as intermediate phenotypes for psychotic disorder: familial liability and reliability. Clin. Neurophysiol. 122 (10), 1984–1990.

Strandburg, R.J., Marsh, J.T., Brown, W.S., Asarnow, R.F., Guthrie, D., Harper, R., Yee, C.M., Nuechterlein, K.H., 1997. Event-related potential correlates of linguistic information processing in schizophrenics. Biol. Psychiatry 42 (7), 596–608.