We both say tomato: Intact lexical alignment in schizophrenia and bipolar disorder

Victoria Sharpe[1], Lotte Schoot[1], Kathryn Eve Lewandowski[2,3], Dost Ongur[2,3], Halide Bilge Türközer[2,3,4], Tuna Hasoglu[2,3,5], & Gina R. Kuperberg[1,3,6]

[1]Department of Psychology, Tufts University, Medford, MA,

[2]McLean Hospital, Harvard Medical School, Belmont, MA

[3]Department of Psychiatry, Harvard Medical School, Boston, MA

[4]The University of Texas Southwestern Medical Center

[5]Department of Psychiatry and Behavioral Health, Penn State College of Medicine, Hershey, PA

[6]Massachusetts General Hospital, Harvard Medical School, Boston, MA

Corresponding Author:

Victoria Sharpe, M.S.

Department of Psychology

Tufts University

490 Boston Avenue

Medford, MA 02155

e-mail: Victoria.sharpe@tufts.edu

Abstract

In people with schizophrenia and related disorders, impairments in communication and social functioning can negatively impact social interactions and quality of life. In the present study, we investigated the cognitive basis of a specific aspect of linguistic communication—*lexical alignment*— in people with schizophrenia and bipolar disorder. We probed lexical alignment as participants played a collaborative picture-naming game with the experimenter, in which the two players alternated between naming a dual-name picture (e.g., *rabbit/bunny*) and listening to their partner name a picture. We found evidence of lexical alignment in all three groups, with no differences between the patient groups and the controls. We argue that these typical patterns of lexical alignment in patients were supported by preserved—and in some cases increased—bottom-up mechanisms, which balanced out impairments in top-down perspective-taking.

Keywords: linguistic alignment; psychosis; social cognition; mentalizing; lexical priming; language production

For most healthy adults, conversation is a powerful means of communication. In people with schizophrenia and related disorders, however, communication impairments can negatively impact social interactions and quality of life (e.g., Bowie & Harvey, 2008; Bowie et al., 2011; Buck & Penn, 2015) . The present study investigated the cognitive basis of a specific aspect of linguistic communication—*lexical alignment*—in people with schizophrenia and bipolar disorder.

Lexical alignment plays a crucial role in facilitating successful, rewarding communication (Pickering & Garrod, 2006; Fusaroli et al., 2012; Reitter & Moore, 2014). It describes the tendency of speakers to use the same words as their conversational partners to refer to objects in their common environment (Garrod & Anderson, 1987; Brennan & Clark, 1996; Branigan et al., 2011) . For example, if during a conversation between Tori and Gina, Tori points out a cute "bunny", it is more likely that Gina will later refer to this animal as a "bunny" than as a "rabbit".

Lexical alignment is thought to be largely mediated by bottom-up priming mechanisms in which a "prime" word directly activates its internal lexical representation, facilitating subsequent processing of this word. That is, hearing a particular word spoken by a conversational partner would pre-activate its lexical representation, and so a speaker would subsequently find it easier to retrieve/access this word than if it had not been encountered (Pickering & Garrod, 2004) . Importantly, however, as for other forms of lexical priming (e.g. Posner & Snyder, 1975; Tweedy et al., 1977; Neely, 1991; Lau et al., 2013), this type of bottom-up lexical facilitation can be modulated by higher-level beliefs or goals in a top-down fashion (see Supplementary Materials for a more in-depth discussion of bottom-up versus top-down processes in priming and lexical alignment). In the case of lexical alignment, these beliefs include the speaker's knowledge about her conversational partner, and her implicit goal of ensuring effective communication (Branigan et al., 2011) . Thus, successful alignment is closely related to perspective-taking or *mentalizing*[1] (Giles et al., 1991) and to *audience design (*the adaptation of our speech in accordance with what we believe our conversational partner can understand; Clark & Murphy, 1982; Bell, 1984). These principles are thought to play an important role in allowing speakers to align to their partners' use of "dispreferred" names (Brennan and Clark, 1996). For example, returning to the conversation above, Gina's belief that Tori would understand her better by referring to the animal as a "bunny" may help override her default tendency to use the more frequent name ("rabbit").

Given the importance of lexical alignment in communication, it is important to understand how it is affected in people with severe mental illness. Schizophrenia, for example, is characterized by impairments across multiple cognitive and social domains (see Green et al. 2019 for review), which can affect communication and impair psychosocial functioning (e.g., Hogarty & Flesher, 1999; Patterson et al., 2001; Bowie & Harvey, 2008; Buck & Penn, 2015; Vaskinn & Horan, 2020) . However, most studies of lexical processing in schizophrenia have been carried out in non-communicative environments (for reviews, see Kuperberg 2010a, b; Brown & Kuperberg, 2015) .

It is unclear from these previous studies whether one might expect lexical alignment to be spared or impaired in schizophrenia; either would have interesting theoretical implications. On the one hand, studies of lexico-semantic priming suggest that bottom-up influences on priming are preserved (e.g., Vinogradov et al., 1992; Barch et al., 1996) , or even enhanced (e.g., Spitzer et al., 1994; Kreher et al., 2008) in this population (although these effects have been

[1] Throughout the paper, we use the terms *mentalizing* and *perspective-taking* interchangeably, and avoid using the term, *theory of mind*. For discussion of this choice, and how these abilities relate to language, see Supplementary Materials.

studied in comprehension rather than production). This might lead one to predict that lexical alignment should be intact. On the other hand, people with schizophrenia tend to perform poorly at mentalizing (Brune, 2005). Moreover, lexical priming in schizophrenia is usually reduced (relative to healthy controls) when experimental conditions encourage top-down influence of higher-level goals (Barch et al., 1996; Kreher et al., 2009; Sharpe et al., 2020) . These findings might lead one to predict that lexical alignment should be impaired in schizophrenia.

The study of lexical alignment therefore provides a unique opportunity to understand how two fundamental constructs—bottom-up priming and top-down mentalizing—contribute to social-communicative interactions in schizophrenia. It also offers the opportunity to understand how these constructs vary dimensionally *across* diagnostic categories (Insel et al., 2010). For example, there is some evidence that patients with bipolar disorder also exhibit impairments in social and communicative abilities, albeit to a lesser degree than those seen in schizophrenia (Lee et al., 2013; Lee et al., 2017; see also Bora & Pantelis, 2016 and Samame, 2013 for review) . However, there has been little work on linguistic communication in bipolar disorder, and *no* studies of lexical alignment in bipolar.

<u>The present study</u>

The first goal of this study was to determine whether people with schizophrenia and bipolar disorder can lexically align to their conversational partner to the same degree as healthy adults. To this end, we probed lexical alignment as participants played a collaborative picture-naming game with the experimenter. This game was based on a paradigm developed by Branigan and colleagues (Branigan et al., 2016) and was designed to mimic the cooperative nature of conversation while eliciting speech in a controlled fashion. The participant and the experimenter alternated in naming pictures that appeared in front of them on separate screens (see Fig. 1). Critically, a subset of pictures had two names—a preferred name (e.g., *rabbit*) and a dispreferred name (e.g., *bunny*). Two trials before it was the participant's turn to name a dual-name picture, the experimenter named the same picture, either using its dispreferred name (50% of trials) or its preferred name (50% of trials). We operationalized lexical alignment as the odds of a participant using a name that matched the name previously used by the experimenter and asked whether this differed between patients and controls. We also asked whether alignment was more effortful in patients, as indexed by naming response times.

Our second goal was to understand the relationship between lexical alignment in the patient groups and the two major constructs that have been linked to alignment — lexical priming and mentalizing/perspective-taking. To index *lexical priming* in language production, we embedded another subset of "single-name" trials within the picture naming game and measured the difference in the time it took for participants to name repeated pictures (named first by the experimenter) versus unrepeated (named first by the participant). To index *perspective-taking*, participants carried out a standardized test of mentalizing—Section 3 of The Awareness of Social Inference Test (TASIT; McDonald et al., 2003)—in which they watched videos of characters using lies and/or sarcasm and answered yes/no questions about these characters' intentions. We selected this measure because it is commonly used to assess mentalizing in schizophrenia (e.g., Roberts & Penn, 2009; Sparks et al., 2010) , and is psychometrically sound (McDonald et al., 2006; Pinkham et al., 2016), with direct relevance to how people make inferences about others' mental states in conversation. For each of these constructs, we first asked whether there were differences in performance between patients and controls. Based on the previous literature, we predicted that, while lexical production priming would be spared (or even increased) in patients, mentalizing would be impaired. We then asked whether, *within* the patient groups, individual differences in each measure predicted individual differences in lexical alignment. We predicted that, while impaired mentalizing would be linked to decreased lexical alignment, greater lexical priming would be linked to increased lexical alignment.

Methods

Participants

Sixty-four outpatients were recruited from McLean Hospital—32 with schizophrenia/schizoaffective disorder, and 32 with bipolar disorder (DSM-IV). Thirty-one control participants with no history of neuropsychiatric disorders, who were not taking psychoactive medication, were recruited by advertisement. All were native American English speakers with no (self-reported) history of learning or language disability. All provided informed consent following the guidelines of the Partners Institutional Review Board and were compensated for their time. Patients were assessed within eight weeks of the experimental session using several standardized scales (see Table 1 and Supplementary Materials for further details about clinical characterization and recruitment).

	Control	*Schizophrenia*	*Bipolar*	*Schizophrenia v. Control*	*Bipolar v. Control*
N	31	32	32	-	-
Gender (M | F)[a]	17 | 14	22 | 10	16 | 16	χ^2= 0.77, p = 0.38	χ^2= 0.02, p = 0.89
Race (AI | A | AA | W | MR | NR)[b, c]	2 | 0 | 4 | 23 | 1 | 1	0 | 0 | 7 | 23 | 2 | 0	0 | 1 | 3 | 28 | 0 | 0	χ^2= 4.14, p = 0.39	χ^2= 5.62, p = 0.35
Age	38.52 (13.8)	44.19 (11.1)	33.53 (9.8)	t = 1.79, p =0 .08	t= -1.65, p= 0.11
Parental SES[d]	49.96 (12.3)	42.88 (15.8)	53.82 (9.2)	t = -1.74, p = 0.09	t = 1.28, p = 0.21
Premorbid Verbal IQ[e]	110.21 (8.2)	100.58 (10.9)	109.04 (9.2)	t = -3.98, p = 0.00	t = -0.53, p = 0.60
TASIT[f]	55.16 (4.5)	46.50 (9.6)	52.0 (6.1)	t = -4.62, p = 0.00	t = -1.64, p = 0.11
Duration of Illness (years)	-	21.39 (11.4)	15.19 (11.2)		
PANSS: General[g]	-	32.16 (7.5)	28.84 (6.0)		
PANSS: Positive[g]	-	16.25 (5.6)	11.48 (3.8)		
PANSS: Negative[g]	-	15.59 (5.4)	11.32 (3.8)		
YMRS: Total[h]	-	11.31 (5.7)	8.87 (8.3)		
Antipsychotics (T | AT)[i]	-	8 | 21	7 | 16		
CPZ Equivalent (mg)[j]	-	433.48 (340.8)	402.33 (301.1)		

Table 1. Means are indicated, with standard deviations in parentheses. [a]M: Male, F: Female. [b]AI: American Indian/Alaska Native, A: Asian, , AA : African American/Black, W: White, MR: More than one race, NR: Unknown or not reported. [c]Hawaiian/Pacific Islander was also considered, but no participants identified with this group. [d]SES: Parental Socioeconomic Status, assessed using the Hollingshead Index (Hollingshead, 1965). [f]The Awareness of Social Inference Test (McDonald et al., 2003). [g]PANSS: Positive and Negative Syndrome Scale (Kay et al., 1987. [h]Young Mania Rating Scale (Young et al., 1978) . [i]T: Typical, AT: Atypical. [j]CPZ: Chlorpromazine equivalent, see Supplementary Materials for details of calculation. The three groups were matched on gender, $X^2(2, 95)=2.484$, $p = .289$, but it was not possible to match all groups on age, parental socioeconomic status, or premorbid verbal IQ. To account for potential confounds, these three variables were covariates in all analyses in which Group was a predictor of interest.

The collaborative picture naming game

The game was set up as shown in Figure 1. On each trial, a picture of an object, animal, or person appeared on each screen. The two players alternated between naming the picture (a "production trial") and listening to their partner name the picture, then stating whether their picture was the "same" or "different" (a "comprehension trial"). Below we describe the participants' production trials used in our analyses. See Supplementary Materials for a full description of the comprehension trials and development of stimuli.

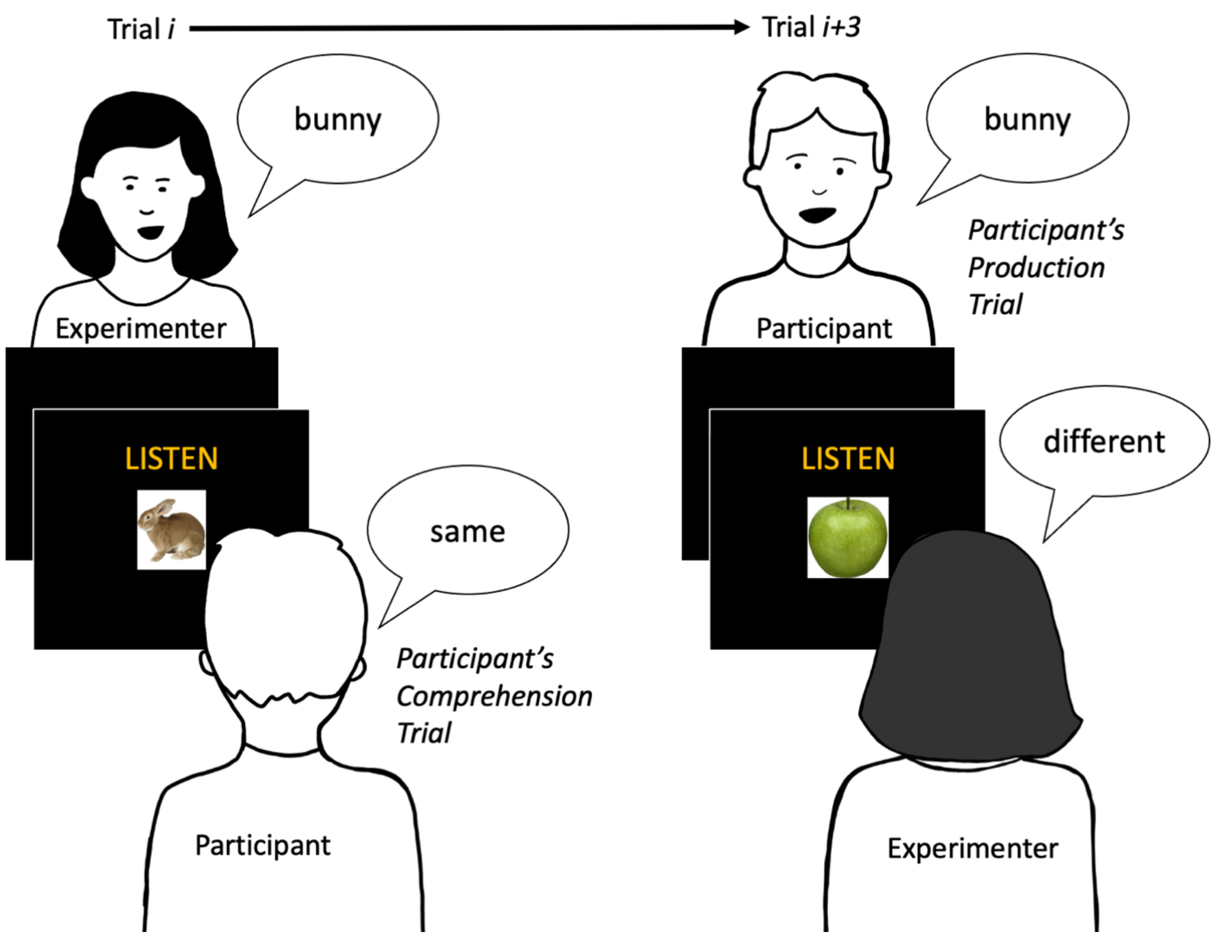

Figure 1. Setup of the cooperative picture naming game. The participant and experimenter faced one another, each looking at their own computer screen, which was not visible to the other player. On the i*th trial, the experimenter named a dual-name*

picture, and the participant, looking at his own screen, then stated whether the picture that he was looking at was the "same" or "different." After two intervening trials (the i+3th trial), the participant named the same picture (here, using the same name the experimenter used), and the experimenter said "same" or "different."

Of these 110 production trials, 20 were used to probe alignment. Each used a picture that had two possible names: a *preferred* name (e.g., "rabbit"), which was more frequent (SUBTL-US corpus; Brysbaert & New, 2009), and a *dispreferred* name (e.g., "bunny") that was nonetheless judged to be highly acceptable (mean acceptability rating of 6.5 on a 1-7 scale). These dual-name stimuli were developed by carrying out two norming studies, described in Supplementary Materials.

During the experiment itself, each of the 20 dual-name pictures was initially named by the experimenter and, two trials later, by the participant. In 10 trials, the experimenter used the picture's dispreferred name; in the other 10 trials, the experimenter used the picture's preferred name. These conditions (preferred, dispreferred) were counterbalanced across lists so that no participant named the same picture twice, but across participants, each dual-name picture was preceded half the time by its dispreferred name and half by its preferred name.

Twenty additional production trials were used to probe lexical production priming. In these trials, the pictures had a single name that was matched in frequency to the dispreferred name of one of the dual-name pictures. In 10 of these trials, the picture was repeated; that is, two trials back, the participant first heard its name spoken by the experimenter. In the other 10 trials, the picture was unrepeated: the participant named the picture first. Again, these conditions (repeated, unrepeated) were counterbalanced across lists. The remaining 70 production trials were filler trials with high frequency single-name pictures. All were named for the first time by the participant.

Before the experiment itself, participants received instructions stressing the cooperative aspect of the game and asking them to describe the pictures as specifically as possible, using a single word (e.g., “rabbit” rather than “animal” or “small grey rabbit”). Participants then completed a practice phase. Trials were self-paced, and the experiment was split into five blocks. Between blocks, the percentage of trials that were correctly identified as “same” or “different” appeared on the screen, providing feedback on joint performance to encourage cooperation. The session was recorded, and the names and naming times were extracted following procedures described in Supplementary Materials.

TASIT

After the picture naming game, participants completed Section 3 of TASIT (McDonald et al., 2003). The score was calculated as the number of correct responses (maximum: 68).

Statistics

All analyses were conducted using mixed-effects regression (see Supplementary Materials for details). In all analyses, we included the maximal identifiable random effects structure, which generally consisted of random intercepts at the subject and item levels, and random slopes for all predictors of interest that varied within subjects and items, respectively. Thus, we accounted for by-item and by-subject variance, while controlling Type I error (Barr et al., 2013).

Results

Is lexical alignment spared or impaired in people with schizophrenia and bipolar disorder?

To assess the ability of participants to align to the experimenter, we created a binary variable, *Alignment,* which specified whether, for dual-name pictures, participants produced names that matched or mismatched the names produced by the experimenter. We then used logistic regression, with Alignment as the dependent variable, to determine whether the log odds of alignment exceeded chance levels within each group. Given our design, the probability of match by chance was 0.5 (or equivalently, log odds = 0). In contrast, a participant who lexically aligned with the experimenter would be more likely to use the dispreferred name after hearing the experimenter use the dispreferred name, increasing the probability that the name they used matched that of the experimenter (log odds > 0). We found that, in all three groups, the log odds of producing a name that matched the name used by the experimenter were significantly greater than zero (Table 2).

To determine whether this alignment effect differed between the three groups, we carried out a between-group analysis with Alignment as the dependent variable and with Group (schizophrenia vs. bipolar vs. controls) as the between-subjects predictor of interest. Three additional subject-level variables were included as “control” regressors: Premorbid Verbal IQ (Blair & Spreen, 1989) , Age, and Parental SES (Hollingshead, 1965). We found no evidence of any between-groups differences in the log odds of alignment (Table 2, Fig. 2A). However, because a non-significant effect does not imply evidence *for* the null hypothesis, we then confirmed the absence of alignment differences between groups with non-inferiority testing (Lakens et al., 2018) and computed Bayes Factors, reported in Supplementary Materials.

Log Odds of Lexical Alignment

	Predictor	**Estimate (log odds)**	**SE**	**Wald's z**	**p**	**Sig.**
Within Controls	Intercept	1.08	0.17	6.31	0.00	***
Within Schizophrenia	Intercept	1.01	0.13	7.58	0.00	***
Within Bipolar	Intercept	1.39	0.17	8.44	0.00	***
Between Groups	Intercept	1.02	0.17	6.12	0.00	***
	Group (Schiz.)	0.09	0.19	0.47	0.64	
	Group (Bipolar)	0.34	0.18	1.94	0.05	^

Table 2. ^Marginally significant with slightly greater alignment in the bipolar group. Logistic regression analyses examining the log odds of lexical alignment within and between groups. Effects of predictors of interest are shown. Note that alignment within each group is given by the Intercept of each model, which describes the log odds of the participant's word choice matching that of the experimenter. See Supplementary Materials for full results.

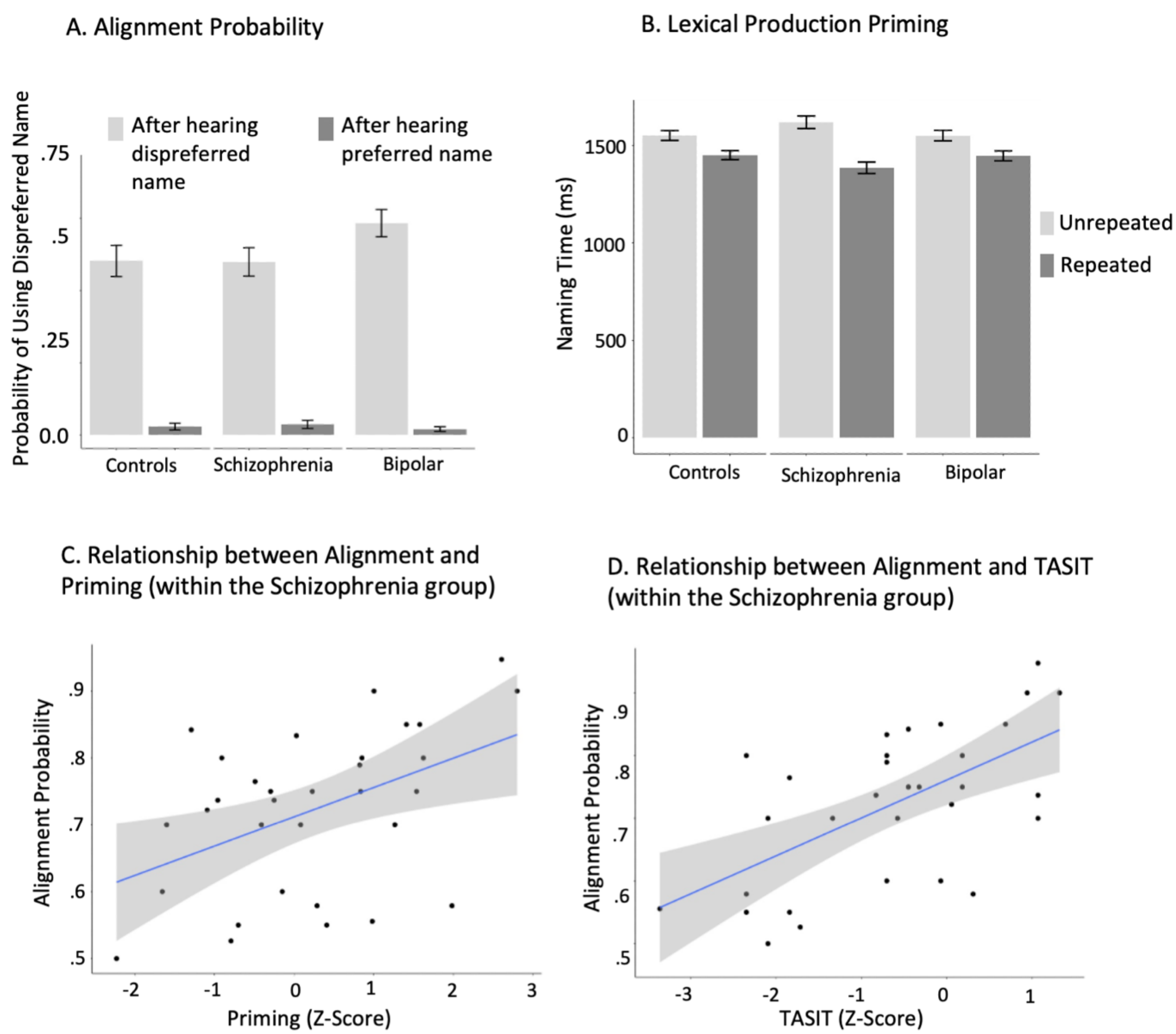

Figure 2. (A) Probability of using the dispreferred name after hearing the experimenter use the dispreferred name versus the preferred name. Error bars represent one standard error. (B) Lexical production priming: predicted naming times for repeated versus unrepeated low frequency single-name pictures, after controlling for demographic and lexical factors and for mean response time. Error bars represent one standard error. (C) Positive correlation between lexical production priming effects and alignment probabilities within the schizophrenia group. Points represent actual data, whereas the line represents the estimated regression coefficient (Est. = 0.23; SE = 0.09; p =0.01) and the shaded region around it represents the predicted standard error. (D) Positive correlation between mentalizing (TASIT scores) and alignment probabilities within the schizophrenia group. Points represent actual data, whereas the line represents the estimated regression coefficient (Est. = 0.32; SE = 0.08; p =0 .00) and the shaded region around it represents the predicted standard error.

We then addressed the possibility that alignment was more effortful in the patient groups than in the control group. To do this, we extracted naming times for a subset of dual-name targets in which the experimenter had first produced the dispreferred name (e.g., “bunny”), and carried out linear mixed effects regression to compare participants’ times to name pictures on which they did align (e.g., produced “bunny”; 52% of trials) and did not align (e.g., produced “rabbit”; 48% of trials) to the experimenter. Log-transformed naming times served as the dependent variable, and we also included the average naming time for each participant as a subject-level predictor, to account for the fact that patients’ longer overall naming times can lead to spurious inflations of naming time differences (Chapman et al., 1994) . We found no significant difference in naming times aligned and non-aligned trials in any group, nor was the interaction between Group and Alignment significant (see Table 3).

Effect of Alignment on Naming Times of Dual-Name Trials

	Predictor	Estimate	SE	t	p	Sig.
Within Controls	Alignment	-0.01	0.03	-0.20	0.84	
Within Schizophrenia	Alignment	0.00	0.04	0.03	0.97	
Within Bipolar	Alignment	-0.05	0.04	-1.23	0.24	
Between Groups	Alignment	0.02	0.04	-0.20	0.64	
	Group (Schiz.)	0.20	0.07	3.01	0.00	**
	Group (Bipolar)	0.12	0.06	2.04	0.05	.
	Alignment*Group (Schiz.)	0.02	0.07	2.52	0.81	
	Alignment*Group(Bipolar)	-0.07	0.05	-1.01	0.21	

Table 3. Effect of Alignment on log-transformed naming times within and between groups. Effects of predictors of interest are shown. See Supplementary Materials for full results. Statistical significance was assessed using a type-III sums of squares estimation, with p-values estimated using the Satterthwaite approximation (Satterthwaite, 1946).

What are the relationships between lexical alignment, lexical production priming and mentalizing ability within the patient groups?

Our second goal was to understand how individual differences in lexical production priming and mentalizing ability might drive individual differences in lexical alignment. We began by asking whether these two constructs differed between groups.

Lexical Production Priming. We used mixed effects linear regression to determine whether there were differences in production priming between groups (again, log-transformed and with by-subject mean naming times as a covariate). We found a significant main effect of Repetition, with faster naming times to repeated than non-repeated single-name trials (see Table 4, Fig. 2B). We also found a Group by Repetition interaction, driven by a larger production priming effect in the schizophrenia group than the control group, which persisted above and beyond the variance accounted for by Premorbid Verbal IQ, Parental Socioeconomic Status, and Age. There was no difference in production priming between the bipolar group and the controls.

Effect of Repetition on Naming Times

	Predictor	Estimate (ms)	SE (ms)	t	p	Sig.
Within Controls	Repetition	-0.08	0.02	-3.75	0.00	**
Within Schizophrenia	Repetition	-0.13	0.03	-4.39	0.00	***
Within Bipolar	Repetition	-0.08	0.02	-4.31	0.00	***
Between Groups	Repetition	-0.07	0.03	-2.57	0.01	*
	Group (Schiz.)	0.04	0.02	1.53	0.13	
	Group (Bipolar)	0.00	0.02	0.14	0.89	
	Repetition*Group (Schiz.)	-0.07	0.03	-2.39	0.02	*
	Repetition*Group (Bipolar)	-0.01	0.03	-0.30	0.76	

Table 4. Effect of Repetition (repeated vs. non-repeated) on log-transformed naming times within and between groups. Effects of predictors of interest are shown. See Supplementary Materials for full results. Statistical significance was assessed using a type-III sums of squares estimation, with p-values estimated using the Satterthwaite approximation (Satterthwaite, 1946).

TASIT. As expected, TASIT scores were significantly lower in the schizophrenia group (mean = 46.50, sd = 9.61) than in controls (mean = 55.16, sd = 4.50; Est. = -3.57, SE = 1.62, $p < 0.05$). The difference in scores between the bipolar group (mean = 52.90, sd = 6.12) and the controls was marginally significant (Est. = -2.59, SE = 1.48, $p = 0.08$).

Predictors of Alignment within patient groups. Above, we showed that both lexical production priming and mentalizing were abnormal in patients with schizophrenia. To test our hypothesis that both constructs would predict individual differences in lexical alignment amongst patients, we carried out regressions *within* the schizophrenia group. We found that better TASIT scores and larger lexical production priming effects both predicted a greater probability of alignment (Table 5, Fig. 2C, D). TASIT and lexical production priming were uncorrelated with one another ($r = -0.03$, $p = 0.19$), despite adequate variance in the distribution of these scores, and predictors remained significant when we included them in the same regression model, suggesting that they each accounted for unique sources of variance in alignment probability.

In the bipolar group, production priming again predicted alignment (with a larger priming effect predicting greater alignment), but the effect of TASIT on alignment was non-significant (Table 5). The effect of production priming remained significant when TASIT was included in the same model. Note that the distribution of TASIT scores in the bipolar group was heavily right-skewed, suggesting a ceiling effect may have accounted for the absence of an effect of TASIT on alignment.

For exploratory analyses of the relationship between alignment and symptoms, see Supplementary Materials.

Effect of Mentalizing and Lexical Production Priming on Log Odds of Alignment

	Predictor	Estimate (log odds)	SE	z	p	Sig.
Within the Schizophrenia Group						
Model 1: TASIT Only	TASIT	0.32	0.08	4.00	0.00	***
Model 2: Priming Only	Priming	0.23	0.09	2.65	0.01	**
Model 3: TASIT + Priming	TASIT	0.30	0.08	3.57	0.00	***
	Priming	0.18	0.08	2.20	0.03	*
Within the Bipolar Group						
Model 1: TASIT Only	TASIT	-0.20	0.19	-1.04	0.30	
Model 2: Priming Only	Priming	0.42	0.15	2.85	0.00	**
Model 3: TASIT + Priming	TASIT	-0.29	0.18	-1.63	0.10	
	Priming	0.47	0.15	3.19	0.00	**

Table 5. Logistic regression analyses examining the effect of mentalizing (TASIT) and lexical production priming (Priming) on the log odds of alignment in patients. Effects of predictors of interest are shown. See Supplementary Materials for full results.

Discussion

We used a collaborative picture-naming task to investigate lexical alignment in people with schizophrenia, bipolar disorder, and healthy controls. We found evidence of lexical alignment in all three groups, with no differences between the patient groups and the controls. In our analyses of response times, we also found no evidence that alignment to the dispreferred names was more effortful in either of the patient groups than the controls. Interestingly, however, the schizophrenia group showed significant differences from controls in two constructs that have been linked to lexical alignment: *reduced* mentalizing, operationalized by performance on the TASIT, and *increased* production priming, indexed as faster times to name repeated *versus* unrepeated single-name pictures. Moreover, within the schizophrenia group, individual differences in both these constructs predicted variability in lexical alignment such that better TASIT scores and increased production priming predicted an increased probability of lexical alignment.

Given that schizophrenia and bipolar disorder have both been linked to impairments in multiple aspects of social cognition and communication (e.g., Bora & Pantelis, 2016), the preservation of lexical alignment in both groups is striking. It is, however, consistent with a small body of work showing that other types of alignment—on "politeness" (Stewart et al., 2008) and syntactic structure (Dwyer et al., 2020)—are relatively spared in schizophrenia. It is also consistent with work showing that, despite clear social atypicalities, people with autism spectrum disorder exhibit typical alignment across multiple levels of linguistic representation (e.g., Allen et al., 2011; Slocombe et al., 2013; Nadig et al., 2015; Branigan et al., 2016). Taken together, these findings suggest that, rather than being globally impaired, people with these neuropsychiatric disorders have windows of spared social and communicative functioning.

Given the known link between lexical alignment and perspective-taking in healthy adults

(Branigan et al., 2011), it is particularly interesting that lexical alignment in schizophrenia was spared, despite patients' poor performance on the TASIT. We suggest two possible reasons for this dissociation. The first is that the TASIT tapped into a form of perspective-taking that was distinct from the more implicit "audience design" function that has been proposed to influence lexical alignment. In the TASIT, participants watch videos and make judgments about the intentions of the depicted characters who use non-literal language (e.g., sarcasm). Making explicit judgments requires offline reasoning and introspection that may not be engaged during everyday interaction (see Frith, 2004 for discussion).

On the other hand, the TASIT also taps into implicit top-down perspective-taking mechanisms that may be shared with lexical alignment. For example, selecting a dispreferred (non-literal) meaning over a preferred (literal) meaning requires the ability to infer another person's communicative intent. Analogously, in the present alignment task, the *selection* of a dispreferred name (*bunny*) over a preferred name (*rabbit*) requires participants to represent the experimenter's mind and infer that alignment would facilitate communicative success (Branigan et al., 2011) . The implicit online use of top-down pragmatic information is impaired in schizophrenia (Rabagliati et al., 2018), and such impairments have long been linked to communicative difficulties (e.g., Harrow et al., 1989; Pawelczyk et al., 2017). We therefore favor another explanation for why lexical alignment was spared in the schizophrenia group — that patients *compensated* for their impaired top-down mentalizing with enhanced bottom-up priming mechanisms.

In support of this account, the schizophrenia group showed a significantly larger lexical production priming effect than controls. This is consistent with some previous studies reporting enhanced lexical priming in schizophrenia during both comprehension (e.g., Spitzer et al., 1994;

Kreher et al., 2008)) and production (Kuperberg et al., 2018). In this previous work, however, "hyperpriming" was observed only under highly automatic conditions in which the prime preceded the target by less than 400ms within a single trial. While there are some studies showing that, when targets are separated from primes by several trials (analogous to the production lexical priming paradigm used here), repetition priming during comprehension is preserved (e.g., Clare et al., 1993; Sponheim, 2004; Doniger et al., 2001) , the present findings provide the first evidence that, during language production, lexical repetition priming is enhanced in schizophrenia.

As noted in the Introduction, lexical alignment in healthy adults is thought to be mediated by both bottom-up priming (Pickering & Garrod, 2004) and top-down mentalizing (Branigan et al., 2011) mechanisms. For example, after hearing the experimenter produce a dispreferred name (*bunny*), bottom-up influence may increase the level of activation of this relatively infrequent lexical representation so that it is equal to that of the more frequent lexical representation (*rabbit*). Therefore, when the participant names the picture, top-down perspective-taking may play an additional role in selecting the dispreferred name to achieve alignment, as described above. We suggest that, in schizophrenia, enhanced bottom-up priming raised the activation level of the repeated dispreferred lexical representation such that it was even higher than that of the preferred lexical representation, allowing patients to produce the dispreferred name and align with the experimenter, without relying on top-down mentalizing.

Further evidence that enhanced bottom-up priming might have compensated for reduced top-down mentalizing during lexical alignment in the patient groups comes from our examination of individual differences. In the schizophrenia group, we found that patients with worse mentalizing performance were less likely to align with the experimenter, while patients with larger-than-average lexical production priming effects showed greater alignment. We also showed that lexical production priming predicted lexical alignment within the bipolar group, although in this group, there was less variability in TASIT scores and no evidence of a relationship between TASIT and alignment.

Limitations and Open Questions

The interpretation offered above—that lexical alignment was preserved in schizophrenia because increased bottom-up priming compensated for reduced top-down mentalizing—leads to several testable predictions.

First, it predicts that lexical alignment may not *always* be spared in schizophrenia. More complex social situations are likely to impose greater demands on top-down mentalizing (Reitter & Moore, 2014), calling on producers to dynamically up- or down-regulate alignment based on their beliefs about different partners (cf. Branigan et al., 2011) . Using a non-social lexical priming paradigm, we recently showed that, in contrast to healthy controls, people with schizophrenia failed to implicitly adapt to changes in linguistic environments by engaging in top-down prediction (Sharpe et al., 2020). This raises the possibility that, even though patients' greater reliance on bottom-up priming allowed them to achieve similar *baseline* levels of alignment to controls, they may be less able to flexibly *adapt* their alignment in more dynamic social contexts (Nadig et al., 2015; Branigan et al., 2016). Thus, understanding how patients use top-down beliefs to facilitate alignment has implications for predicting psychosocial outcomes, as patients who are unable to flexibly adapt their speech to their partners' are likely to have difficulties in social communication.

A limitation of the present study is that we were unable to link heterogeneity of sociocognitive impairments/mechanisms of lexical alignment with heterogeneity in specific symptom profiles amongst patients because our sample size was too small with too little variance in psychotic symptoms. Of particular interest is the relationship between lexical alignment and positive thought disorder, which is the most obvious clinical manifestation of linguistic communicative difficulties in schizophrenia (American Psychiatric Association, 2013). Some previous studies have linked thought disorder to impairments in taking the listeners' perspective into account (Rochester et al., 1977; Hoffman et al., 1982; Hoffman, 1986; see also Rutter, 1985; Harrow et al., 1989), and to more severe impairments in mentalizing ability (Frith and Corcoran, 1996). There is also preliminary evidence that thought-disordered patients exhibit impaired alignment at other levels of linguistic representation (syntax & description types; Dwyer et al., 2020). It will therefore be important for future studies using larger samples to investigate how individual symptoms predict individual differences in lexical alignment.

Conclusion

To conclude, our results suggest that, at least in structured communicative environments like the present experiment, lexical alignment is unlikely to contribute to communication difficulties in chronic schizophrenia or bipolar disorder. We argue that these typical patterns of lexical alignment were supported by preserved—and in some cases increased—bottom-up mechanisms, which balanced out deficits in top-down perspective-taking. This interpretation raises important questions about how imbalances in bottom-up and top-down mechanisms of alignment may impact communication during more complex social interactions.

Acknowledgments

This work was funded by the National Institute of Child Health and Human Development (R01HD082527 to G.R.K.). We are also grateful to the Sidney R. Baer, Jr. Foundation for their support of undergraduate students including Rebecca Schaub, Sayaka Koga, Emma Hayden, and Emma Lampropoulos. In particular, we thank Rebecca Schaub and Sayaka Koga for their assistance in developing stimuli, and Sayaka Koga, Emma Lampropoulos, Jaisal Merchant, and Grace Konstantin for their critical role in data collection and recruitment. In addition, we thank Nate Delaney-Busch and Meredith Brown for their advice during development of the study.

References

Allen, M. L., Haywood, S., Rajendran, G., & Branigan, H. (2011). Evidence for syntactic alignment in children with autism. *Dev Sci, 14*(3), 540-548.

American Psychiatric Association. (2013). *DSM-V: Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Barch, D. M., Cohen, J. D., Servan-Schreiber, D., Steingard, S., Steinhauer, S., & van Kammen, D. (1996). Semantic priming in schizophrenia: An examination of spreading activation using word pronunciation and multiple SOAs. *Journal of Abnormal Psychology, 105*(4), 592-601.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.

Bell, A. (2008). Language style as audience design. *Language in Society, 13*(2), 145-204.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *Clinical Neuropsychologist, 3*(2), 129-136.

Bora, E., & Pantelis, C. (2016). Social cognition in schizophrenia in comparison to bipolar disorder: A meta-analysis. *Schizophr Res, 175*(1-3), 72-78.

Bowie, C. R., Gupta, M., & Holshausen, K. (2011). Disconnected and underproductive speech in schizophrenia: unique relationships across multiple indicators of social functioning. *Schizophr Res, 131*(1-3), 152-156.

Bowie, C. R., & Harvey, P. D. (2008). Communication abnormalities predict functional outcomes in chronic schizophrenia: Differential associations with social and adaptive functions. *Schizophrenia Research, 103*(1-3), 240-247.

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition, 121*(1), 41-57.

Branigan, H. P., Tosi, A., & Gillespie-Smith, K. (2016). Spontaneous lexical alignment in children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(11), 1821.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482-1493.

Brown, M., & Kuperberg, G. R. (2015). A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia. *Frontiers in Human*

Neuroscience, 9, 643.
Brune, M. (2005). "Theory of mind" in schizophrenia: A review of the literature. *Schizophr Bull, 31*(1), 21-42.
Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977-990.
Buck, B., & Penn, D. L. (2015). Lexical Characteristics of Emotional Narratives in Schizophrenia: Relationships With Symptoms, Functioning, and Social Cognition. *J Nerv Ment Dis, 203*(9), 702-708.
Chapman, L. J., Chapman, J. P., Curran, T. E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. *Developmental Review, 14*, 159-185.
Clare, L., McKenna, P. J., Mortimer, A. M., & Baddeley, A. D. (1993). Memory in schizophrenia: What is impaired and what is preserved? *Neuropsychologia, 31*(11), 1225-1241.
Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in psychology, 9*, 287-299.
Doniger, G. M., Silipo, G., Rabinowicz, E. F., Snodgrass, J. G., & Javitt, D. C. (2001). Impaired sensory processing as a basis for object-recognition deficits in schizophrenia. *Am J Psychiatry, 158*(11), 1818-1826.
Dwyer, K., David, A. S., McCarthy, R., McKenna, P., & Peters, E. (2020). Linguistic alignment and theory of mind impairments in schizophrenia patients' dialogic interactions. *Psychol Med, 50*(13), 2194-2202.
Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological medicine, 34*(03), 385-389.
Frith, C. D., & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychol Med, 26*(3), 521-530.
Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylen, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychol Sci, 23*(8), 931-939.
Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27*(2), 181-218.
Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics* (pp. 1-68). Cambridge, UK: Cambridge University Press.
Green, M. F., Horan, W. P., & Lee, J. (2019). Nonsocial and social cognition in schizophrenia: current evidence and future directions. *World Psychiatry, 18*(2), 146-161.
Harrow, M., Lanin-Kettering, I., & Miller, J. G. (1989). Impaired perspective and thought pathology in schizophrenic and psychotic disorders. *Schizophrenia Bulletin, 15*(4), 605-623.
Hoffman, R. E. (1986). Tree structures, the work of listening, and schizophrenic discourse: A reply to Beveridge and Brown. *Brain and Language, 27*(2), 385-392.
Hoffman, R. E., Kirstein, L., Stopek, S., & Cicchetti, D. V. (1982). Apprehending schizophrenic discourse: A structural analysis of the listener's task. *Brain and Language, 15*(2), 207-233.
Hogarty, G. E., & Flesher, S. (1999). Practice principles of cognitive enhancement therapy for schizophrenia. *Schizophrenia Bulletin, 25*(4), 693-708.
Hollingshead, A. B. (1965). *Two factor index of social position*. New Haven, CT: Yale University Press.
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry, 167*(7), 748-751.
Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin, 13*(2), 261-276.
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. *Schizophrenia Research, 111*(1-3), 174-181.
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. *Schizophrenia Bulletin, 34*(3), 473-482.
Kuperberg, G. R. (2010). Language in schizophrenia Part 1: an Introduction. *Language and Linguistics Compass, 4*(8), 576-589.
Kuperberg, G. R. (2010). Language in schizophrenia Part 2: What can psycholinguistics bring to the study of schizophrenia...and vice versa? *Language and Linguistics Compass, 4*(8), 590-604.
Kuperberg, G. R., Delaney-Busch, N., Fanucci, K., & Blackford, T. (2018). Priming Production: Neural evidence for

enhanced automatic semantic activity immediately preceding language production in schizophrenia. *NeuroImage: Clinical, 18*, 74-85.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science, 1*(2), 259-269.

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience, 25*(3), 484-502.

Lee, J., Altshuler, L., Glahn, D. C., Miklowitz, D. J., Ochsner, K., & Green, M. F. (2013). Social and nonsocial cognition in bipolar disorder and schizophrenia: relative levels of impairment. *Am J Psychiatry, 170*(3), 334-341.

Lee, J., Rizzo, S., Altshuler, L., Glahn, D. C., Miklowitz, D. J., Sugar, C. A., . . . Green, M. F. (2017). Deconstructing Bipolar Disorder and Schizophrenia: A cross-diagnostic cluster analysis of cognitive phenotypes. *J Affect Disord, 209*, 71-79.

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): a clinical test of social perception. *Disabil Rehabil, 28*(24), 1529-1542.

Nadig, A., Seth, S., & Sasson, M. (2015). Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Front Psychol, 6*, 1888.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading and Visual Word Recognition* (pp. 264-333). Hillsdale, NJ: Erlbaum.

Patterson, T. L., Moscona, S., McKibbin, C. L., Davidson, K., & Jeste, D. V. (2001). Social skills performance assessment among older patients with schizophrenia. *Schizophrenia Research, 48*(2), 351-360.

Pawełczyk, A., Kotlicka-Antczak, M., Łojek, E., Ruszpel, A., & Pawełczyk, T. (2017). Schizophrenia patients have higher-order language and extralinguistic impairments. *Schizophrenia Research*.

Pickering, M. J., & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behav Brain Sci, 27*(2), 212-225.

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation, 4*(2-3), 203-228.

Pinkham, A. E., Penn, D. L., Green, M. F., & Harvey, P. D. (2016). Social cognition psychometric evaluation: Results of the initial psychometric study. *Schizophr Bull, 42*(2), 494-504.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition: the Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rabagliati, H., Delaney-Busch, N., Snedeker, J., & Kuperberg, G. R. (2018). Spared bottom-up but impaired top-down interactive effects during naturalistic language processing in schizophrenia: Evidence from the visual world paradigm. *Psychological Medicine*, 1-11.

Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language, 76*, 29-46.

Roberts, D. L., & Penn, D. L. (2009). Social cognition and interaction training (SCIT) for outpatients with schizophrenia: a preliminary study. *Psychiatry Research, 166*(2), 141-147.

Rochester, S., Martin, J. R., & Thurston, S. (1977). Thought-process disorder in schizophrenia: The listener's task. *Brain and Language, 4*(1), 95-114.

Rutter, D. R. (1985). Language in schizophrenia. The structure of monologues and conversations. *British Journal of Psychiatry, 146*, 399-404.

S., M., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after tramatic brain injury. *J Head Trauma Rehabil, 18*(3), 219-238.

Samame, C. (2013). Social cognition throughout the three phases of bipolar disorder: A state-of-the-art overview. *Psychiatry Res, 210*(3), 1275-1286.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin, 2*(6), 110-114.

Sharpe, V., Weber, K., & Kuperberg, G. R. (2020). Impairments in probabilistic prediction and Bayesian learning can explain reduced neural semantic priming in schizophrenia. *Schizophr Bull, 46*(6), 1558-1566.

Slocombe, K. E., Alvarez, I., Branigan, H. P., Jellema, T., Burnett, H. G., Fischer, A., . . . Levita, L. (2013). Linguistic alignment in adults with and without Asperger's syndrome. *J Autism Dev Disord, 43*(6), 1423-1436.

Sparks, A., McDonald, S., Lino, B., O'Donnell, M., & Green, M. J. (2010). Social cognition, empathy and functional outcome in schizophrenia. *Schizophrenia Research, 122*(1-3), 172-178.

Spitzer, M., Beuckers, J., Beyer, S., Majer, S., & Hermle, L. (1994). Contextual insensitivity in thought-disordered schizophrenic patients: Evidence from pauses in spontaneous speech. *Language and Speech, 37*(2), 171-185.

Sponheim, S. R., Steele, V. R., & McGuire, K. A. (2004). Verbal memory processes in schizophrenia patients and biological relatives of schizophrenia patients: intact implicit memory, impaired explicit recollection. *Schizophr Res, 71*(2-3), 339-348.

Stewart, S. L., Corcoran, R., & Drake, R. J. (2008). Alignment and theory of mind in schizophrenia. *Cognitive neuropsychiatry, 13*(5), 431-448.

Tweedy, J. R., Lapinski, R. H., & Schvaneveldt, R. W. (1977). Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. *Memory & Cognition, 5*(1), 84-89.

Vaskinn, A., & Horan, W. P. (2020). Social Cognition and Schizophrenia: Unresolved Issues and New Challenges in a Maturing Field of Research. *Schizophr Bull, 46*(3), 464-470.

Vinogradov, S., Ober, B. A., & Shenaut, G. K. (1992). Semantic priming of word pronunciation and lexical decision in schizophrenia. *Schizophrenia Research, 8*, 171-181.

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry, 133*(5), 429-435.