ADDRESSING FORMAL THOUGHT DISORDER IN PSYCHOSIS THROUGH NOVEL ASSESSMENT AND TARGETED INTERVENTION

by

Matthew P. Marggraf

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Psychology at IUPUI

Indianapolis, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. Kyle S. Minor, Chair

Department of Psychology

Dr. Michelle P. Salyers

Department of Psychology

Dr. Paul H. Lysaker

Department of Psychology

Dr. Jesse C. Stewart

Department of Psychology

Approved by:

Dr. Jesse C. Stewart

To my Mom and Dad

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my mentor, Dr. Kyle Minor, for his guidance and unwavering support throughout my graduate training and during this project. I would also like to thank my committee members for providing their expertise and support. I would like to acknowledge the Indiana Clinical and Translational Science Institute, from whom I was awarded a pre-doctoral training grant which helped fund this project.

I would like to especially thank my graduate cohort: Yue, Jay, Loretta, Shaun, Devin, Brittanny, and Danielle, for their love, support, and encouragement; for celebrating the little things along the way and helping me persevere during challenging times. I would also like to thank my parents for their unconditional love, understanding, and endless encouragement throughout my life, and especially during my pursuit of this degree. Lastly I would like to thank the rest of my family and my friends for their love and support over the years.

TABLE OF CONTENTS

ABSTRACT 6
INTRODUCTION 7
Study Aims 10
METHODS 11
Participants and Recruitment 11
Procedures 12
Measures 13
Data Analyses 16
Power Analyses 17
RESULTS 18
Sample Characteristics 18
Aim 1 18
Aim 2 19
DISCUSSION 20
Strengths and Limitations 24
Conclusion 24
REFERENCES 26
APPENDIX A. MERIT CORE ELEMENTS 33
APPENDIX B. AUTOMATED FTD DESCRIPTIONS 35
APPENDIX C. FIGURES 36
APPENDIX D. TABLES 37

ABSTRACT

Formal thought disorder (FTD) is a debilitating symptom of psychosis. It is linked to functional deficits and generally demonstrates poor response to interventions. Metacognition has emerged as a potential therapeutic target that may be effective in reducing FTD, as metacognitive deficits and FTD both arise from disruptions in associative thought processes. This study's primary aim was to determine whether FTD could be reduced with metacognitive therapy. Pre-post changes in FTD severity were assessed using clinician-rated and automated measures in 20 individuals with psychotic disorders who received 12 sessions of evidence-based metacognitive therapy. We also examined whether reductions in FTD were larger when assessed with automated instruments versus clinician-rated measures. Aim two compared associations between FTD and three outcome variables (social functioning, role functioning, metacognition) across FTD-measurement approach. Results indicated that automated FTD, but not clinician-rated FTD, was significantly reduced post-intervention. This effect was more robust within a subsample exhibiting greater levels of FTD. Strength of associations between FTD and outcome variables did not differ across FTD measurement approach. These findings provide initial evidence that a targeted metacognitive intervention can reduce FTD. Effects were strongest for automated instruments, which may be more sensitive to detecting change; however, differences in measurement type did not extend to associations with selected outcome variables. This study provides preliminary support for future efforts to reduce FTD. Large-scale studies with longer intervention periods may further our understanding of the effectiveness of metacognitive intervention on FTD.

INTRODUCTION

Formal thought disorder (FTD) has been recognized as a hallmark symptom of psychotic disorders for over a century (Bleuler, 1911/1950). It represents a fundamental disruption in the organization and maintenance of goal-directed thought processes, which can result in disjointed, disconnected patterns of thinking and communication. FTD manifests clinically as disorganized speech, which has long been considered a core clinical symptom in diagnostic classification systems (e.g., Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; APA, 2013). The presence of FTD can result in speech that is characterized as tangential, circumstantial, incoherent, and, in general, difficult for the listener to understand (Andreasen, 1979a). FTD is estimated to affect approximately half of those with schizophrenia (Breier & Berg, 1999) and is linked to poor social functioning (e.g. reductions in the quality and quantity of interpersonal relationships; Bowie, Gupta, & Holshausen, 2011; De Sousa et al., 2015; Marggraf, Lysaker, Salyers, & Minor, 2020) and role functioning (e.g. impaired performance in occupational or academic responsibilities; Holshausen, Harvey, Elvevag, Foltz, & Bowie, 2014; Marengo & Harrow, 1997; Racenstein, Penn, Harrow, & Schleser, 1999). These social and role functioning deficits contribute to the $150 billion annual economic burden of psychotic disorders (Cloutier et al., 2016). FTD has been observed to persist even during periods of relative stability in other symptoms (Marengo & Harrow, 1997; Yalincetin et al., 2016).

Few studies have addressed how to treat FTD across the psychopharmacological and psychosocial intervention literature. With regard to medication trials, some have reported a limited treatment response (Harrow & Marengo, 1986; Harvey, Docherty, Septer, & Rassmussen, 1990; Remberk, Namyslowska, & Rybakowski, 2012). Regarding psychosocial intervention, relatively little is known about the effectiveness of psychotherapy on FTD as none have directly focused on reducing FTD, nor do studies typically include FTD as an outcome variable (Kircher, Brohl, Meier, & Englen, 2018). Existing interventions focus broadly on symptoms or specifically on improving skill deficits or changing delusional thought content (Beck, Rector, Stolar, & Grant, 2009). Thus, they may not offer adequate guidelines for a

clinician working with patients experiencing FTD. Indeed, Beck and colleagues (2011) suggest that FTD may be the least explored and treated symptom in psychotherapy. It is possible that an intervention focusing on the organization of thought processes could yield reduced FTD.

A second possible reason our knowledge about the effectiveness of therapy is limited is that poor therapeutic alliance is associated with FTD (Cavelti et al., 2016). Given the importance of verbal communication during psychotherapy, the disorganized speech central to FTD may represent a barrier to building a therapeutic alliance. This is detrimental given the crucial impact therapeutic alliance has on therapy outcome in psychosis populations (Farrelly et al., 2014; Priebe, Richardson, Cooney, Adedeji, & McCabe, 2011); this may lead some clinicians to view FTD as an insurmountable barrier to effective engagement in therapy. However, recent therapy frameworks have been developed that indicate FTD can be understood in the context of, and not prohibitive to, effective therapeutic intervention (e.g. Lysaker & Lysaker, 2006; Hamm & Firmin, 2016).

For severely disorganized individuals, Lysaker and Lysaker (2006) assert that key elements in the treatment and recovery process include becoming more integrated, making sense of their experiences, and relating to others. This is based on a line of work that has examined metacognition – a domain that is profoundly impaired in those with psychotic disorders and consists of mental activities ranging from discrete abilities (e.g. identifying thoughts and emotions) to synthetic processes (e.g. forming and integrating a complex representation of self and others) (see: Semerari et al., 2003; Lysaker et al., 2012; Lysaker et al., 2015). A recent line of research has identified metacognitive deficits as a potential target for recovery-oriented psychotherapy (Lysaker, Glynn, Wilkniss, & Silverstein, 2010; Lysaker & Dimaggio, 2014; Lysaker et al., 2014). Critically, metacognition has demonstrated strong links to FTD, with both appearing to arise from related disordered thought processes (Minor et al., 2015; Minor & Lysaker, 2014). That these disordered processes are related is consistent with Bleuler's (1911/1950) early conceptualization that disruption in associative thought processes is the fundamental cognitive impairment in those with psychotic disorders. A reduced capacity to link related ideas together compromises a person's ability to communicate thoughts and ideas cohesively (i.e., FTD) and form coherent, complex ideas about the self and world (i.e., metacognition). Therefore, a metacognitive therapy focused on organizing one's thoughts about themselves and others into a cohesive, meaningful narrative understanding will likely result in

(and possibly require) an organization of underlying thought processes. Thus, improved metacognitive capacity may lead to specific improvements in FTD. To date, no published study has examined whether FTD improves following metacognitive therapy.

Metacognitive Reflection and Insight Therapy (MERIT; Lysaker et al., 2014) is a leading metacognitive therapy that has demonstrated feasibility and preliminary effectiveness for improving metacognition in psychosis populations (de Jong et al., 2016a; Van Dokersgoed et al., 2014). Unlike other existing treatments for psychosis (e.g. skill-focused, CBT), MERIT is a patient-driven therapy, centered in reflection on the therapeutic process, rather than on the instruction of skills or efforts to correct the content of beliefs. Metacognition is viewed as a hierarchical capacity, and thus, interventions are tailored to "meet" the client at the appropriate metacognitive level. This could be useful for assisting a client to understand that their mental content and communications are their own and progress toward organization and integration of thoughts and—together with the therapist—establish consensual communication patterns. This may also foster therapeutic alliance given that interventions are appropriate for an individual's cognitive abilities. Further, MERIT requires the establishment of a non-hierarchical, consultative therapy alliance, wherein both the therapist and client join together to make sense of the client's life. To date, no study has examined the effectiveness of MERIT on FTD; however, support for this outcome has been demonstrated in several published case studies (Lysaker & Lysaker 2006; Hamm & Firmin, 2016; De Jong, Van Donkersgoed, Pijnenborg, & Lysaker, 2016b). Thus, MERIT may be a uniquely suited intervention to improve FTD, with the potential to overcome obstacles present in other psychosocial interventions (see Figure A.1).

Even if FTD could be improved following targeted treatment, the ability to successfully detect a change in FTD remains a critical issue. Although clinician-rated measures are the most commonly used approach for the assessment of FTD, disadvantages with this approach may obscure the ability to detect improved FTD. These disadvantages include the inherent subjectivity accompanying clinician ratings and the use of ordinal scales. Ordinal scales yield a limited response set (e.g., 1 to 7 scale), which may be to be insensitive to all but gross changes over time or between individuals (Cohen & Elvevag, 2014; Cohen et al., 2008). Thus, in therapy studies, it is possible that improvements in FTD may not be captured by these measures. To address the limitations imposed by traditional clinician-rated assessments, an automated measurement of FTD – which holds the potential to be more objective and sensitive – will also

be used in the current study. Over the last two decades, technological advancements have led to the emergence of automated, computational programs that can rapidly assess numerous linguistic features of speech and writing samples (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998; McNamara, Louwerse, McCarthy, & Graesser, 2010). These instruments are capable of measuring deeper-level, multi-dimensional linguistic features by comparing an individual's communication to large text corpora to assess the degree of FTD and incorporate information at varying levels of discourse (e.g. within and across sentences and paragraphs) – beyond the scope of information able to be processed by humans. An emerging line of research has examined how these automated instruments could be applied to psychosis-spectrum populations to assess FTD (Elvevag et al., 2007; Elvevag et al., 2010; Minor, Willits, Marggraf, Jones, & Lysaker, 2018; Moe et al., 2016; Nicodemus et al., 2014; Willits, Rubin, Jones, Minor, & Lysaker, 2018). Automated analysis has been used to examine how speech content changes in response to therapy (Arvidsson, Sikstrom, Werbart, 2011) and to successfully differentiate the speech of individuals with psychotic disorders from controls and unaffected first-degree relatives (Elvevag et al., 2010; Nicodemus et al., 2014), as well as within-group differences (i.e. high-FTD vs. low-FTD; Elvevag et al., 2007). Direct comparisons between automated FTD and clinician-rated FTD have yielded small to medium convergence (i.e., medium effect size, Elvevag et al., 2007; effect sizes ranging from small to medium, Minor et al., 2018), suggesting that automated analysis may be capturing related, yet distinct aspects of FTD. To my knowledge, no existing studies have used automated instruments to examine whether FTD improves in response to therapy.

Study Aims

The primary aims of this study were to: 1a) test if FTD can be reduced after 12 sessions of an evidence-based, targeted psychosocial intervention (i.e. MERIT) by comparing pre-post treatment data in a pilot study; 1b) determine if automated assessments produced larger effect sizes compared to clinician-rated FTD when examining pre-post changes due to potentially increased sensitivity; and 2) examine whether associations between FTD and social functioning, role functioning, and metacognition differ based on measurement approach.

METHODS

Participants and Recruitment

Data for this study was collected as part of a larger randomized controlled trial examining the effectiveness of traditional- versus individually-tailored (i.e., integrating real-world social interactions) therapy across 24 sessions of MERIT. Real-world social interactions were recorded on a device with the Electronically Activated Recorder (EAR) and all participants wore EAR devices for two days after each therapy session, regardless of MERIT treatment condition. For the current study, the midpoint of the larger study (i.e., 12 sessions) was used as the post-intervention point. Participants were recruited from local community mental health centers (CMHCs) in a large Midwestern city from July 2016 through February 2019. To facilitate recruitment, study personnel visited local CMHCs and discussed the study with mental health providers and care coordinators during weekly staff meetings, who disseminated information about the study to patients who met study criteria. Study personnel also distributed study brochures to interested participants in drop-in/waiting room areas of a local CMHC.

Participants were eligible to participate in the study if they: 1) had a psychotic disorder (e.g. schizophrenia, schizoaffective disorder, psychosis NOS) confirmed by a Mini-International Neuropsychiatric Interview (M.I.N.I.; Sheehan et al., 1998); 2) were between ages 18-60; 3) were proficient in English; 4) had no inpatient hospitalization or changes in medications within 30 days of the initial assessment; 5) were not currently receiving metacognitive therapy (i.e., MERIT or any metacognitive intervention); and 6) were able to provide informed consent. If a participant was currently receiving psychotherapy that was not metacognitive-based (e.g., group therapy, case management, supportive therapy), they were still eligible to participate and were allowed to continue with their current treatment regimen. Exclusion criteria included having a documented history of intellectual disability or neurological condition, experiencing head injury resulting in loss of consciousness greater than five minutes, or meeting criteria for current alcohol or substance dependence.

In total, 35 participants enrolled in the study and completed at least a portion of the baseline assessment. To be included in the Aim 2 sample, participants had to have complete baseline data which included: demographic data, clinician-rated FTD, automated-FTD, social

functioning, role functioning, and metacognition. Four participants were excluded from all analyses due to incomplete data. Thus, the final Aim 2 sample included 31 participants. From this pool, participants were included in the Aim 1 sample if they had complete baseline data and had complete post-intervention FTD data (i.e., clinician-rated FTD, automated FTD) [additional excluded n = 11]). Of the 11 participants not included in the Aim 1 sample, three were randomized to a non-treatment condition, one never attended their first MERIT session, and seven dropped out prior to the 12-session mark. This resulted in completed data for 20 participants in the Aim 1 sample. Our attrition rate of 28.6% (8 out of 28) is consistent with other pilot studies of similar duration using MERIT (De Jong et al., 2016a: 12 session, 25% attrition) and other metacognitive interventions (Favrod et al., 2011: 8 sessions, 28% attrition rate).

Procedures

To assess eligibility, participants provided relevant demographic and medical information and completed the MINI administered by a trained research assistant. Participants who did not meet study criteria were deemed ineligible and were compensated $10 for their time. Eligible participants completed a baseline assessment, which included psychiatric symptoms, FTD, social functioning, role functioning, and metacognition. After the baseline assessment, participants were randomized to one of the two treatment conditions or a treatment-as-usual condition that did not participate in either therapy condition. Participants in treatment conditions attended 12 weekly, in-person, individual sessions of MERIT. Each session typically lasted 45-50 minutes.

During sessions, the therapist and patient engaged in joint thinking and reflection about the patient's life with the goal of organizing thoughts toward an integrated sense of understanding of the patient's life and how they fit into the broader world and society – based on eight core elements of MERIT (see Appendix A for descriptions of core elements of a MERIT session). Although MERIT is considered to be a long-term therapy, evidence from a pilot study of an RCT found that clinical gains in metacognition were observed after 12 sessions (De Jong et al., 2016a). Given the time constraints and pilot nature of the current study, 12 sessions of MERIT was used as the intervention duration to maximize the amount of data that could be collected during the project window. After 12 sessions of MERIT, participants

completed a post-intervention assessment consisting of the same battery administered at baseline. Participants were paid $30 for each assessment session and $10 at each therapy session for returning the EAR device (as part of larger, parent study). Those who did not have access to transportation were provided transportation via local taxi services at no cost to the participant.

Therapy sessions were conducted by myself and other master's-level graduate students with several years of experience administering MERIT to individuals with psychotic disorders. To ensure fidelity to MERIT, all study therapists received weekly supervision from a licensed clinician, who has expertise in treating individuals with psychosis and co-developed MERIT (Paul Lysaker). To further monitor adherence to MERIT, study therapists rated themselves on the Therapist Metacognitive Adherence Scale (T-MAS) after every session. On the T-MAS, therapists use a five-point scale to rate the extent to which each of the eight core elements of MERIT was present in therapy. Informed consent was obtained before performing any study procedures. As part of the consent process, participants reviewed the consent form with a research assistant and were given an opportunity to ask questions about the study procedures and their participation. Study procedures were approved by the local institutional review board.

Measures

Clinician-rated FTD was measured using the Conceptual Disorganization item on the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). The Conceptual Disorganization item measures disorganized thought processes characterized by a disruption of goal-directed sequencing (e.g. circumstantiality, tangentiality, non-sequiturs) Ratings are made based on the severity of FTD observed while administering the PANSS interview. Ratings range from 1 (absent) to 7 (extreme). The PANSS is considered a "gold standard" measure for assessing symptomatology in psychosis (Bell et al., 1992; Peralta & Cuesta, 1994), and the conceptual disorganization item has been used extensively to measure FTD in psychosis (De Sousa et al., 2015; Minor et al., 2015; Suttajit et al., 2015; Tan et al., 2014). All study personnel who administered the PANSS received training and demonstrated acceptable inter-rater reliability (i.e., inter-rater class correlation coefficients $\geq .80$ on training videos).

Automated FTD was measured using Coh-Metrix, a software program designed to compute various language-use statistics from transcribed written or spoken language documents

(McNamara, Graesser, McCarthy, & Cai, 2014). The Coh-Metrix contains 108 indices across 11 categories and provides a comprehensive assessment of discourse characteristics, from unidimensional descriptors (e.g. word count, pronoun usage) to multidimensional cohesion metrics that assess discourse cohesion within sentences, between sentences, between paragraphs and throughout entire text samples. The Coh-Metrix was chosen as an automated measure because of its capability to measure a variety of indices related to semantic cohesion – a core feature of FTD. Specifically, five indices were used to assess FTD: Narrativity, Syntactic Simplicity, Word Concreteness, Referential Cohesion, and Deep Cohesion (see Appendix B for descriptions of automated indices). These indices were chosen based on the results of a principal components analysis of 54 indices related to text cohesion, of which these 5 indices accounted for 54% of the variance (Grasser, McNamara, & Kulikowich, 2011). Higher indices score reflect lower levels of FTD (i.e., greater text "easability"). These five cohesion indices were aggregated and averaged (using z-scores) to compute a composite automated-FTD variable to be used in statistical analyses. Automated instruments have been validated in psychosis samples, with automated FTD successfully predicting clinicians' ratings of FTD (Elvevag et al., 2007), demonstrating associations with recognized neurobiological markers of FTD (Nicodemus et al., 2014; Tagaments, Cortes, Griego, & Elvevag, 2014), and predicting conversion to psychosis in those who are at clinical high-risk for psychotic disorders (Bedi et al., 2015). Coh-Metrix, specifically, has been used to assess FTD using the IPII in individuals with psychotic disorders in previous studies (Minor et al., 2018; Willits et al., 2018).

Social Functioning was assessed using the Global Functioning: Social Scale (GF: Social; Auther, Smith & Cornblatt, 2006), a clinician-rated, interview-based measure of social functioning. Specific questions are asked during the interview to collect information about intimate relationships, close and casual friendships, and conflict within social relationships. The information is used to make an overall rating on the scale, which consists of ten anchor points ranging from 1 (extreme social isolation) to 10 (superior functioning in a wide range of social and interpersonal activities). The GF: Social has demonstrated strong psychometric properties (Cornblatt et al., 2007) and has previously been used to assess social functioning in psychotic disorder samples (Fulford et al., 2013; Minor et al., 2016; Perez, Schafer, & Cadenhead, 2014; Piskulic, Addington, Auther, & Cornblatt, 2011).

Role Functioning was assessed using the Global Functioning: Role Scale (GF: Role; Niendam et al., 2006), a clinician-rated, interview-based measure of occupational and/or academic role functioning. Specific questions are asked during the interview to elicit information regarding the quantity, degree of difficulty, and quality of performance in academic and/or occupational roles (Niendam et al., 2006). This information is used to make an overall rating on the scale, which consists of ten anchor points ranging from 1 (extreme role dysfunction) to 10 (superior role functioning). The GF: Role Scale has demonstrated strong psychometric properties (Cornblatt et al., 2007) and has been used to assess role functioning in individuals with psychotic disorders in previous studies (Fulford et al., 2013; Minor et al., 2016; Piskulic, et al., 2011; Tully, Lincoln, Liyanage-Don, & Hooker, 2013).

Metacognition was assessed with the Metacognition Assessment Scale–Abbreviated (MAS-A; Semerari et al., 2003) – a rating scale designed to measure an individual's metacognitive capacity. The MAS-A consists of four subscales rated in a Likert-style format: 1) "self-reflectivity", which is the ability to identify and understand one's own mental states; 2) "awareness of the other's mind", defined as the ability to think about another person's mental state; 3) "decentration", which is the ability to see the world as existing with others having independent motives; and 4) "mastery", which refers to the ability to use metacognitive knowledge about one's self to effectively implement coping strategies to deal with psychological problems and the distress related to such problems. Total scores on the MAS-A range from 0 – 28 (self reflectivity and mastery are rated on a 0-9 point scale; awareness of the other's mind is rated on a 0-7 point scale; decentration is rated on a 0-3 point scale). Higher scores reflect the presence of greater metacognitive capacity. The MAS-A has consistently demonstrated good psychometric properties (Lysaker et al., 2010, Lysaker et al., 2012) and has been used to assess metacognitive capacity in numerous studies of individuals with psychotic disorders (Lysaker et al., 2005; Lysaker & DiMaggio, 2014; Lysaker et al., 2014; Lysaker et al., 2015). For this project, MAS-A ratings were made by a consensus group of trained raters who received extensive training in the MAS-A and demonstrated acceptable inter-rater reliability. Study therapists were not part of the MAS-A rating consensus group and raters were blind to the assessment time point (i.e., baseline versus post-intervention).

Speech Sample. The Indiana Psychiatric Illness Interview (IPII; Lysaker et al., 2002) served as the speech sample on which metacognition (full interview) and automated FTD (first

section only) were rated. The IPII is a semi-structured interview that assesses perceptions of one's life and mental illness using an open-ended question format. The IPII was chosen over other clinical interviews because of its open-ended nature, which allowed participants to have freedom in what they discussed and how long they spoke on a given topic. Another unique aspect of this interview is that interviewers do not introduce content; thus, if a participant does not bring up hallucinations, for example, hallucinations are not discussed. Research assistants administering the IPII attended a two-day workshop on IPII administration. The IPIIs were audio-recorded and transcribed (removing interviewer speech) for automated analysis.

Data Analyses

Aim 1a: To assess whether FTD will improve after 12-weeks of MERIT, a series of paired-samples t-tests were calculated to examine differences in FTD variables at baseline versus post-intervention assessment. Specifically, there were two main dependent-sample tests conducted, where time (baseline, post-intervention) was the within-subjects independent variable (IVs) and FTD as measured by: 1) clinician-rated scale (conceptual disorganization on PANSS); and 2) automated analysis (Coh-Metrix composite variable) were dependent variables (DVs), respectively. Additional t-tests were also conducted to assess pre-post changes for each of the five individual automated variables (Narrativity, Familiarity, Syntactic Simplicity, Referential Cohesion, Deep Cohesion). Given that FTD is not present in all patients experiencing psychosis (Breier & Berg, 1999), we expected that some individuals may have baseline clinician-rated FTD scores at floor level (i.e., rated "1" on PANSS: Conceptual Disorganization) and thus, reduction in clinician-rated FTD would not be possible for those participants; therefore, a sensitivity analysis was conducted with the analyses outlined above using data only from participants with ratings of ≥ 2 on PANSS: Conceptual Disorganization. *Aim 1b*: To examine whether automated FTD differs from clinician-rated FTD in the magnitude of pre-post change, effect sizes (calculated in Aim 1a) were compared. Effect sizes were based on Cohen's d (1992) where 0.20 is small, 0.50 is medium, and 0.80 is large.

Aim 2: To explore whether associations between FTD and three outcome variables (social functioning, role functioning, and metacognition) differed depending on measurement approach, bivariate correlations were calculated between clinician-rated FTD and the three outcome variables; as well as the composite automated-FTD variable and the three outcome variables.

Fisher's *r*-to-Z transformations were computed to determine whether associations between FTD measurement approach (clinician-rated, automated) and the matched outcome variables were significantly different (i.e., $p < 0.05$). Prior to testing study aims, all variables of interest were screened for normality violations and outliers ($z > 3.5$). Of note, data transformation applied to only one data point (Baseline Word Concreteness) in the entire data set. The raw score was reduced to the equivalent of the outlier cutoff and this value was used in subsequent analyses.

Power Analyses

An *a priori* power analysis was conducted for Aim 1 using G*Power 3.1 calculator (α = 0.05, power = 0.80). The analysis indicated that with a sample of 20 participants, I was powered to detect medium-large effects ($d = 0.67$) according to Cohen's (1992) guidelines. Although sufficient for this dissertation, ideally I would want to be powered to detect small effects (i.e., $d = 0.20$), which would require a sample of 199 participants. Because this dissertation served primarily as a pilot study, and given the limitations of available resources (e.g. compensation costs, study therapists), recruiting the number required to adequately test for small effects was not feasible. A power analysis was conducted for Aim 2 ($\alpha = 0.05$, power = 0.80), and indicated that with a sample size of 31, I was powered to detect only large effects ($r = 0.48$) for bivariate correlations. Again, while it would be ideal to be able to detect medium (i.e., $r = 0.30$) or small effects (i.e., $r = 0.10$), this would require sample sizes of 84 and 782 participants, respectively.

RESULTS

Sample Characteristics

Participants in the Aim 1 sample (n = 20; see Table 1) were mostly non-Hispanic (n = 18, 90%), African-American/Black (n = 15, 75%), and female (n =11, 55%). Participants had a mean age of 44.25 (SD = 10.95), most had completed high school or received a GED (n = 14, 70%), were unemployed (n = 15, 75%), and had participated in psychotherapy in the past (n = 17, 85%). The larger Aim 2 sample (n = 31) was quite similar in demographic characteristics (see Table 1). Participants were mostly non-Hispanic (n = 27, 87%), African-American/Black (n = 19, 61%), and female (n = 17, 55%). Participants had a mean age of 45.97 (SD = 9.80), most had completed high school/GED (n = 20, 68%), were unemployed (n =25, 81%), and had participated in psychotherapy in the past (n = 27, 87%). Study completers (n = 20) and non-completers (n = 11) did not significantly differ on any demographic variable or outcome measure (see Table 2).

Aim 1

To determine whether FTD could be reduced after 12 sessions of MERIT, pre-post FTD ratings using both novel (automated) and traditional (clinician-rated) assessment were analyzed. Contrary to my hypothesis, clinician-rated FTD was not significantly different at post-intervention ($t(19) = -0.25$, $p = 0.804$, $d = 0.06$). Consistent with my hypothesis, FTD was significantly reduced at post-intervention for the composite automated-FTD variable ($t(19) = -2.55$, $p = 0.019$; $d = 0.57$) and the effect size was medium (see Table 3). When pre-post changes for the five automated-variables were examined separately, only Deep Cohesion ($t(19) = -2.84$, $p = 0.010$, $d = 0.64$) demonstrated significant improvement, and the effect size was medium (all other p's > 0.05). Thus, my hypothesis that FTD would be reduced at post-intervention was partially supported.

Follow-up sensitivity analyses were conducted using data from a subset of participants (n = 8) for whom at least some degree of clinician-rated FTD was present at baseline assessment (i.e., PANSS: Conceptual Disorganization ≥ 2). Consistent with the previous analyses, clinician-rated FTD was not significantly reduced post-intervention ($t(7) = 0.61$, p =

0.563, $d = 0.21$). The composite automated-FTD variable exhibited significant improvement at post-intervention, ($t(7) = -2.53$, $p = 0.039$, $d = 0.90$) and the effect size was large (see Table 4). When the five automated-variables were analyzed separately, only Deep Cohesion ($t(7) = -4.40$, $p = 0.003$, $d = 1.57$) demonstrated significant improvement (all other p's > 0.05).

Aim 2

Relationships between FTD measurement approach (clinician-rated, automated) and three outcome variables (social functioning, role functioning, metacognition) were compared. This resulted in a set of 6 primary comparisons (clinician-rated, composite automated FTD x 3 outcome variables; see Table 5). Clinician-rated FTD was not significantly associated with any of the three outcome variables (all p's > 0.05). The composite automated FTD variable was inversely related to role functioning, such that reduced FTD was associated with worse role functioning ($r = -0.45$, $p = 0.011$); it was not significantly related to social functioning or metacognition. When Fisher's r-to-Z transformations were calculated to compare the strength of associations with outcome variables, clinician-rated FTD and automated-FTD were not significantly different (all p's > 0.05; see Table 4). Thus, our exploratory secondary aim was not supported.

Post hoc analyses of associations between FTD measurement in the Aim 2 sample showed small convergence between the clinician-rated scale with the automated composite ($r = 0.12$), Narrativity ($r = 0.12$), Syntactic Simplicity ($r = -0.21$), Word Concreteness ($r = -0.11$), Referential Cohesion ($r = 0.29$), and Deep Cohesion ($r = 0.18$); none were statistically significant (all p's > 0.05).

DISCUSSION

FTD is a core symptom of psychosis that is associated with numerous adverse outcomes. The primary aim of this study was to examine whether improvements in FTD could be observed after twelve sessions of MERIT. Other aims were to examine whether pre-post effect sizes and baseline associations between FTD and outcome variables (i.e. social functioning, role functioning, metacognition) differed with measurement approach (i.e. automated versus clinician-rated). Three key findings emerged. First, post-intervention reductions in FTD were observed for automated measures. Second, pre-post FTD effect sizes appear larger for automated variables than clinician-rated FTD. Third, associations between FTD and social functioning, role functioning, and metacognition did not significantly differ based on measurement approach.

The most noteworthy finding from this study was that FTD was reduced post-intervention when automated instruments were used. This pattern of results was maintained when examining a subsample who had at least some degree of baseline FTD. Although this is initial evidence, these findings are encouraging as very few studies have reported ways to address FTD. With regard to psychosocial interventions, FTD has largely been underexplored (Beck et al., 2011). Existing intervention studies and randomized controlled trials have primarily focused on reducing positive and negative symptoms, and typically do not report FTD outcomes (see reviews by Hazell et al., 2016; Sivec & Montesano, 2012). In their recent review of FTD phenomenology, Kircher and colleagues (2018) noted that they were unable to identify any psychotherapy study where FTD was a primary or even secondary outcome measure (Roche et al., 2015 described similar concerns). Empirical evidence for addressing FTD has been restricted to medication trials. Although there has been some recent progress (Park et al., 2019), psychopharmacological interventions have also had limited effects on reducing FTD (Harrow & Marengo, 1986; Harvey et al., 1990). Identifying ways to treat this symptom is critical as FTD has been linked to longer duration of hospitalization (Andreou et al., 2008; Breier & Berg, 1999; Lenz, Katschnig, & David, 1986) and poor insight (Baier et al., 2000; Smith et al., 2004; but not Barrera et al., 2009). Thus, our finding of reduced FTD using automated instruments provides an encouraging first step in identifying a potentially effective, evidence-based treatment approach.

In addition to improvements in the automated-FTD composite variable, Deep Cohesion was significantly improved at post-intervention assessment. This finding also became more robust in the sensitivity analysis. The Deep Cohesion index reflects the degree to which a speaker may introduce relationships or topics, but includes explicit connective language features, allowing the listener to develop a more coherent understanding of the causal events, process, or actions being communicated (McNamara et al., 2014). Although the other automated indices were not statistically significant, these variables generally produced small effect sizes (*d*'s ranging from 0.10 – 0.29) in the direction of improvement. This suggests that while small gains were observed across all levels of discourse, the most pronounced effect of MERIT occurred at the within- and between-paragraph level of discourse as opposed to lexical (Narrativity, Word Concreteness), syntactic (Syntactic Simplicity), or sentence-to-sentence levels (Referential Cohesion). Of the automated variables, Deep Cohesion also appears to be the most closely related to the goals of MERIT. Deep cohesion measures how strongly multiple topics or relationships are connected to a central idea, which is linked with MERIT's goal of increasing one's capacity to perform synthetic metacognitive acts wherein thoughts, feelings, and connections between events are integrated into more complex representations of self, others, and the world (Lysaker et al., 2010; Semerari et al., 2003). As one develops a more coherent narrative of who they are and how they fit in the world, it appears that their ability to communicate these ideas with others also improves.

One of the more encouraging aspects of our study was that MERIT was effective in reducing FTD despite the relatively brief intervention period (i.e., 12 therapy sessions). Although a previous pilot study observed clinical improvements in metacognition after 12 sessions (de Jong et al., 2016a), MERIT is designed to be a longer-term therapy. Indeed, case studies that have specifically described gross improvements in FTD implemented MERIT over much longer periods (e.g., 18 months, Hamm & Firmin, 2016; 24 months, Lysaker & Lysaker, 2006). Thus, observing reductions in FTD after just 12 sessions highlights the potential efficiency and impact of this therapeutic approach for targeting FTD. Additionally, these reductions were observed without significant improvements in metacognition (post hoc analyses indicated all p's >0.05 for pre-post comparisons of metacognition scores), which may support an iterative process by which FTD and metacognition improve. It is possible that early therapeutic gains after just 12 sessions of therapy only manifested in speech with deeper

semantic connections. Once these connections are strengthened, it may facilitate participation in more synthetic metacognitive actions, such as identifying links between thoughts and behaviors across one's life, which could result in additional organization of thoughts. Future studies using longer intervention periods will likely further our understanding of how aspects of FTD and metacognitive processes inter-relate, potentially characterize a dose-response relationship, and discern whether significant FTD reductions can be observed with clinician-rated measures along with automated instruments.

Although improvements in automated FTD were observed post-MERIT, significant reductions in clinician-rated FTD were not. There are several potential explanations for this discrepancy. First, the aspects of FTD that were most improved may not be as discernable to the human ear and, thus, not as easily rated by a clinician. For example, Deep Cohesion is a multidimensional index, computed in part based on how often words in sentences appear together in large sets of common text databases across different genres (e.g., scientific texts, language arts) while simultaneously comparing connective markers of speech (e.g., referential pronouns) across sentences (McNamara et al., 2014). Human raters are unable to process this amount of information "in the moment" when considering the speech of the individual talking to them. Another potential explanation is the structure of the clinician-rated scale itself. The measure used in this study was: a) based on a single item; and b) used a seven-point ordinal scale, where each point was grounded in qualitative descriptions of the scale anchor points. For an individual's rating to drop from a "4" to a "3", for example, several conditions of the individual's speech must be judged to have been met (e.g., speech is not loose or irrelevant when dealing with complex material or when under minimal pressure). Thus, when the rating options are limited to ordinal scales and a single item, a "higher bar" likely needs to be cleared for ratings to change. Other ordinal clinician-rated FTD scales exist (e.g., Thought, Language, Communication Index; Andreasen 1979a) that assess aspects of FTD separately (e.g., individual ratings for circumstantiality, tangentiality) which could have detected whether different aspects of clinician-rated FTD change in response to treatment. However, as noted by Cohen and Elvevag (2014), ordinal scales are likely less sensitive in detecting mild to moderate changes compared to automated indices.

Measurement type accounted for MERIT's impact on FTD; however, neither type helped explain the relationship between FTD and other outcome variables. Using data from the

larger baseline sample, we found that the strength of associations between FTD and social functioning, role functioning, and metacognition did not significantly differ between automated and clinician-rated measures. This suggests that while automated FTD may be more sensitive to detecting change in FTD, it is comparable to clinician-rated measures in its association with outcome measures. This was somewhat surprising as the clinician-rated FTD and automated-FTD appeared to measure distinct, yet-related, aspects of FTD given the relatively modest convergence between the FTD measures observed in our study (though convergence was generally in line with a previous study using Coh-Metrix variables: Minor et al., 2018).

The implementation of automated instruments has been a relatively recent addition to FTD measurement and little is known about automated FTD's association with outcome measures. One study from our lab demonstrated that automated FTD indices accounted for substantial variance above and beyond clinician-rated disorganized symptoms in predicting metacognitive capacity (Minor et al., 2018); however, associations with measurement type were not directly compared and we were unable to identify previous studies directly comparing the strength of these relationships to clinician-rated measures. One potential explanation for the lack of different associations across measurement type in this study stems from the somewhat surprising associations between constructs observed in our sample. Only automated FTD and role functioning were significantly associated, such that better role functioning was linked with greater levels of FTD, the opposite direction of what was expected based on the literature examining these constructs (e.g., Racenstein et al., 1999; Marengo & Harrow, 1997; Holshausen et al., 2014). The inverse nature of this relationship did not appear to be a unique artifact of automated measurement, as clinician-rated FTD was also negatively associated with role functioning (though not statistically significant). The explanation for the weak or counterintuitive associations between FTD and the three outcome variables in contrast to those reported in previous studies is unclear. A recent meta-analysis examining the relationship between social functioning and FTD revealed a high level of heterogeneity across studies (Marggraf et al., 2020). It is possible that associations between FTD, role functioning, and metacognition are also heterogeneous and may be accounted for by a third variable. Future studies comparing associations between FTD measurement types and functional outcomes could potentially provide information about the utility of each measurement type as they relate to everyday outcomes.

Strengths and Limitations

A significant strength of the study is that we examined whether FTD could be reduced using an evidence-based psychosocial intervention, which addresses a significant gap in the literature (see Kircher et al.'s 2018 review). The study also compared potentially more objective, automated FTD ratings to traditional clinician-rated methods, contributing to a growing literature examining differences between these measurement approaches Elvevag et al., 2007; Elvevag et al., 2010; Minor et al., 2018; Willits et al., 2018). A limitation of the current study is that the small sample size reduced our ability to detect small to medium reductions in FTD. Despite this limitation, we were able to demonstrate that automated FTD variables were reduced. Future studies using larger samples could increase the power to detect and better characterize smaller effects of MERIT on FTD. Another study limitation was that this was a secondary analysis of data collected in a parent-study aiming to improve metacognition and thus, recruitment did not specifically focus on participants with FTD. Enrolling participants based on the presence of FTD may not only increase statistical power but also increase our understanding of how more severe levels of FTD change in response to targeted intervention. A third limitation is that a treatment-as-usual cohort was not recruited, limiting our ability to account for natural changes in FTD over time. However, FTD is generally considered to be a relatively stable symptom of psychosis (Docherty, Cohen, Nienow, Dinzeo, & Danglemaier, 2003), with one longitudinal study demonstrating FTD to be relatively stable across 2-, 4.5-, and 7-year follow-up periods (Marengo & Harrow, 1997).

Conclusion

In sum, this is the first known study to specifically examine whether FTD can be reduced with an evidence-based targeted psychosocial intervention (i.e., MERIT). We found significant medium-sized reductions in FTD when assessed with automated instruments but not clinician-rated FTD measures. The effect of MERIT on automated FTD became more robust when analyzing a subsample with at least some detectable level of FTD. In both samples, effect sizes for automated FTD variables were generally larger than clinician-rated variables. Using a larger sample, clinician-rated FTD was not significantly associated with social functioning, role functioning, or metacognition. Automated FTD was not associated with social functioning or

metacognition and, contrary to expectations, was inversely associated with role functioning. The strength of associations between FTD and social functioning, role functioning, and metacognition did not differ based on measurement approach. These findings indicate that although automated FTD was more sensitive to change than clinician-rated FTD, associations between outcome variables did not differ. Future studies should seek to replicate these findings in a larger sample of individuals with FTD and across a longer duration of therapy to determine if more robust changes in FTD can be observed.

REFERENCES

American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders*, Fifth edition. Arlington, VA, American Psychiatric Association.

Andreou, C., Bozikas, V.P., Papouliakos, I., Kosmidis, M.H., Garyfallos, G., Karavatos, A., & Nimatoudis, I. (2008). Factor structure of the Greek translation of the scale for the assessment of thought, language and communication. *Australian & New Zealand Journal of Psychiatry*, *42*(7), 636-642.

Andreasen N.C. (1979a). Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability. *Archives of General Psychiatry, 36,* 1315-1321.

Auther, A.M., Smith, C.W., & Cornblatt, B. (2006). Global Functioning: Social Scale (GF: Social). *Glen Oaks, NY: Zucker-Hillside Hospital*.

Baier, M., DeShay, E., Owens, K., Robinson, M., Lasar, K., Peterson, K., & Bland, R.S. (2000). The relationship between insight and clinical factors for persons with schizophrenia. *Archives of Psychiatric Nursing*, *14*(6), 259-265.

Barrera, A., McKenna, P.J., & Berrios, G.E. (2009). Formal thought disorder, neuropsychology and insight in schizophrenia. *Psychopathology*, *42*(4), 264-269.

Beck, E.M., Cavelti, M., Kvrgic, S., Kleim, B., & Vauth, R. (2011). Are we addressing the 'right stuff' to enhance adherence in schizophrenia? Understanding the role of insight and attitudes towards medication. *Schizophrenia Research*, *132*(1), 42-49.

Beck, A.T., Rector, N.A., Stolar, N., & Grant, P. (2009). *Schizophrenia: Cognitive theory, research, and Therapy*. New York: Guilford Press.

Bedi, G., Carrillo, F., Cecchi, G.A., Slezak, D.F., Sigman, M., Mota, N.B., ... & Corcoran, C.M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia, 1*, 15030.

Bell, M., Milstein, R., Beam-Goulet, J., Lysaker, P., & Cicchetti, D. (1992). The Positive and Negative Syndrome Scale and the Brief Psychiatric Rating Scale: reliability, comparability, and predictive validity. *Journal of Nervous and Mental Disease, 180,* 723-728.

Bleuler, E. (1950). Dementia praecox or the group of schizophrenias. (J. Zinkin,Trans.). New York: International Universities. (Original work published 1911).

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

Breier, A. & Berg, P.H. (1999). The psychosis of schizophrenia: Prevalence, response to atypical antipsychotics, and prediction of outcome. *Biological Psychiatry*, 46, 361–364.

Cavelti, M., Homan, P., & Vauth, R. (2016). The impact of thought disorder on therapeutic alliance and personal recovery in schizophrenia and schizoaffective disorder: an exploratory study. *Psychiatry Research, 239,* 92-98.

Cloutier, M., Aigbogun, M.S., Guerin, A., Nitulescu, R., Ramanakumar, A. V., Kamat, S. A., … Wu, E. (2016). The economic burden of schizophrenia in the United States in 2013. *Journal of Clinical Psychiatry, 77(6),* 764-771.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Cohen, A.S., Alpert, M., Nienow, T.M., Dinzeo, T.J., & Docherty, N.M. (2008). Computerized measurement of negative symptoms in schizophrenia. *Journal of Psychiatric Research, 42*, 827–836.

Cohen AS, Elvevag, B (2014). Automated computerized analysis of speech in psychiatric disorders. *Current Opinions in Psychiatry, 27*, 203-209.

Cornblatt, B.A., Auther, A.M., Niendam, T., Smith, C.W., Zinberg, J., Bearden, C.E., & Cannon, T.D. (2007). Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophrenia Bulletin, 33*(3), 688-702.

De Jong, S., van Donkersgoed, R.J.M., Aleman, A., van der Gaag, M., et al. (2016a). Practical implications of metacognitively oriented psychotherapy in psychosis: findings from a pilot study. *Journal of Nervous and Mental Disease, 204,* 713-716.

De Jong, S., van Donkersgoed, R, Pijnenborg, G.H.M., & Lysaker, P.H. (2016b). Metacognitive Insight and Reflection Therapy (MERIT) with a patient with severe symptoms of disorganization. *Journal of Clinical Psychology, 72,* 164-174.

De Sousa, P., Spray, A., Sellwood, W., & Bentall, R.P. (2015). 'No man is an island'. Testing the specific role of social isolation in formal thought disorder. *Psychiatry Research, 230*(2), 304-313.

Docherty, N.M., Cohen, A.S., Nienow, T.M., Dinzeo, T.J., & Dangelmaier, R.E. (2003). Stability of formal thought disorder and referential communication disturbances in schizophrenia. *Journal of Abnormal Psychology*, *112*(3), 469.

Elvevag, B., Foltz, P.W., Weinberger, D.R., Goldberg, T.E. (2007). Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophrenia Research, 93,* 304–316.

Elvevag, B., Foltz, P.F., Rosenstein, M., DeLisi, L. (2010). An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. *Journal of Neurolinguistics, 23,* 270–284.

Farrelly, S., Brown, G., Szmukler, G., Rose, D., Birchwood, M., Marshall, M., ... & Thornicroft, G. (2014). Can the therapeutic relationship predict 18 month outcomes for individuals with psychosis? *Psychiatry Research, 220*(1-2), 585-591.

Favrod, J., Maire, A., Bardy, S., Pernier, S., & Bonsack, C. (2011). Improving insight into delusions: A pilot study of metacognitive training for patients with schizophrenia. *Journal of Advanced Nursing, 67,* 401-407.

Fulford, D., Niendam, T.A., Floyd, E.G., Carter, C.S., Mathalon, D.H., Vinogradov, S., ... & Loewy, R.L. (2013). Symptom dimensions and functional impairment in early psychosis: more to the story than just negative symptoms. *Schizophrenia Research, 147*(1), 125-131.

Graesser, A.C., McNamara, D.S., & Kulikowich, J.M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*(5), 223-234.

Hamm, J.A., & Firmin, R.L. (2016). Disorganization and individual psychotherapy for schizophrenia: A case report of metacognitive reflection and insight therapy. *Journal of Contemporary Psychotherapy*, *46*(4), 227-234.

Hamm, J.A., Renard, S.B., Fogley, R.L., Leonhardt, B.L., Dimaggio, G., Buck, K.D., & Lysaker, P.H. (2012). Metacognition and social cognition in schizophrenia: stability and relationship to concurrent and prospective symptom assessments. *Journal of Clinical Psychology, 68,* 1303-1312.

Harrow, M. & Marengo, J.T. (1986). Schizophrenic thought disorder at follow-up: Its persistence and prognostic significance. *Schizophrenia Bulletin, 12*(3), 373.

Harvey, P.D., Docherty, N.M., Serper, M.R, Rasmussen, M (1990). Cognitive deficits and thought disorder: II. An 8-month follow-up study. *Schizophrenia Bulletin, 16*, 147–156.

Hazell, C.M., Hayward, M., Cavanagh, K., & Strauss, C. (2016). A systematic review and meta-analysis of low intensity CBT for psychosis. *Clinical Psychology Review, 45*, 183-192.

Holshausen, K., Harvey, P.D., Elvevag, B., Foltz, P.W., & Bowie, C.R., (2014). Latent semantic variables are associated with formal thought disorder and adaptive behavior in older patients with schizophrenia. *Cortex, 55,* 88-96.

Kay SR, Fizbein A, Opler LA. (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13,* 261-276.

Kircher, T., Brohl, H., Meier, F., & Engelen, J. (2018). Formal thought disorders: from phenomenology to neurobiology. *Lancet Psychiatry, 5*, 515-526.

Landauer, T.K. & Dumais, S.T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211-240.

Landauer, T.K., Foltz, P.W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25,* 259-284.

Lenz, G., Katschnig, H., & David, H. (1986). Symptoms, diagnosis and time in hospital. *Psychopathology*, *19*(5), 253-258.

Lysaker, P.H., Buck, K.D., Leonhardt, B., Buck, B.E., Hamm, J.A., Vohs, J et al. (2014). Metacognitively focused psychotherapy for persons with schizophrenia: Eight core elements that define practice. In. P. H. Lysaker, G. Dimaggio, & M. Brune (Eds.). *Metacognition and Social Cognition in Schizophrenia.* New York: Elsevier Press.

Lysaker, P.H., & Dimaggio, G. (2014). Metacognitive capacities for reflection in schizophrenia: implications for developing treatments. *Schizophrenia Bulletin*, *40*(3), 487-491.

Lysaker, P.H., Carcione, A., Dimaggio, G., Johannesen, J.K., Nicolò, G., Procacci, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia: Associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatrica Scandinavica*, *112*(1), 64-71.

Lysaker, P.H., Clements, C.A., Placak-Hallberg, C, Knipschure, S.J., Wright, D.E. (2002). Insight and personal narratives of illness in schizophrenia. *Psychiatry*, 65, 197–206.

Lysaker, P.H., Glynn, S.M., Wilkniss, S.M., & Silverstein, S.M. (2010). Psychotherapy and recovery from schizophrenia: A review of potential applications and need for future study. *Psychological Services*, *7*(2), 75-91.

Lysaker, P.H., Lysaker, J.T. (2006). Psychotherapy and schizophrenia: An analysis of requirements of an individual psychotherapy for persons with profoundly disorganized selves. *Journal of Constructivist Psychology, 19*, 171–189.

Lysaker, P.H., Ringer, J.M., Buck, K.D., et al. (2012) Metacognitive and social cognition deficits in patients with significant psychiatric and medical adversity: a comparison between participants with schizophrenia and a sample of participants who are HIV positive. *Journal of Nervous and Mental Disease, 200*, 130–134.

Lysaker, P.H., Vohs, J., Minor, K.S., Irarr´azaval, L., Leonhardt, B., Hamm, J.A., . . . Dimaggio, G. (2015). Metacognitive deficits in schizophrenia: Presence and associations with psychosocial outcomes. *Journal of Nervous and Mental Disease, 203(7)*, 530–536.

Marengo, J.T., & Harrow, M. (1997). Longitudinal courses of thought disorder in schizophrenia and schizoaffective disorder. *Schizophrenia Bulletin, 23*(2), 273-285.

Marggraf, M.P., Lysaker, P.H., Salyers, M.P., & Minor, K.S. (2020). The link between formal thought disorder and social functioning in schizophrenia: A meta-analysis. *European Psychiatry, 63*.

McNamara, D.S., Graesser, A.C., McCarthy, P.M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

McNamara, D.S., Louwerse, M.M., McCarthy, P.M., & Graesser, A.C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes, 47*(4), 292-330.

Minor, K.S., Lysaker, P.H. (2014). Necessary but not sufficient: Links between neurocognition, social cognition, and metacognition in schizophrenia are moderated by disorganized symptoms. *Schizophrenia Research, 159*, 198-204.

Minor, K.S., Marggraf, M.P., Davis, B.J., Luther, L., Vohs, J.L., Buck, K.D., Lysaker, P.H. (2015). Conceptual disorganization weakens links in cognitive pathways: Disentangling neurocognition, social cognition, and metacognition in schizophrenia. *Schizophrenia Research, 169*, 153-158.

Minor, K.S., Marggraf, M.P., Davis, B.J., Mehdiyoun, N.F., & Breier, A. (2016). Affective systems induce formal thought disorder in early-stage psychosis. *Journal of Abnormal Psychology, 125*(4), 537.

Minor, K.S., Willits, J.A., Marggraf, M.P., Jones, M.N., & Lysaker, P.H. (2018). Measuring disorganized speech in schizophrenia: automated analysis explains variance in cognitive deficits beyond clinician-rated scales. *Psychological Medicine*, 1-9.

Moe, A.M., Breitborde, N.J., Shakeel, M.K., Gallagher, C.J., & Docherty, N.M. (2016). Idea density in the life-stories of people with schizophrenia: Associations with narrative qualities and psychiatric symptoms. *Schizophrenia Research, 172*(1-3), 201-205.

Nicodemus, K.K., Elvevåg, B., Foltz, P.W., Rosenstein, M., Diaz-Asper, C., & Weinberger, D.R. (2014). Category fluency, latent semantic analysis and schizophrenia: a candidate gene approach. *Cortex, 55*, 182-191.

Niendam, T.A., Bearden, C.E., Johnson, J.K., & Cannon, T.D. (2006). *Global Functioning: Role Scale (GF: Role)*. Los Angeles, CA: University of California, Los Angeles.

Park, Y.C., Lee, M.S., Si, T.M., Chiu, H.F., Kanba, S., Chong, M.Y., ... & Rabbani, G. (2019). Psychotropic drug-prescribing correlates of disorganized speech in Asians with schizophrenia: The REAP-AP study. *Saudi Pharmaceutical Journal*, *27*(2), 246-253.

Peralta, V. & Cuesta, M.J. (1994). Psychometric properties of the Positive and Negative Syndrome Scale. *Psychiatry Research,* 31-40.

Perez, V.B., Shafer, K.M., & Cadenhead, K.S. (2012). Visual information processing dysfunction across the developmental course of early psychosis. *Psychological Medicine*, *42*(10), 2167-2179.

Piskulic, D., Addington, J., Auther, A., & Cornblatt, B.A. (2011). Using the global functioning social and role scales in a first-episode sample. *Early Intervention in Psychiatry, 5*(3), 219-223.

Priebe, S., Richardson, M., Cooney, M., Adedeji, O., & McCabe, R. (2011). Does the therapeutic relationship predict outcomes of psychiatric treatment in patients with psychosis? A systematic review. *Psychotherapy and Psychosomatics, 80*(2), 70-77.

Remberk, B., Namyslowska, I., & Rybakowski, F., 2012. Cognition and communication dysfunctions in early-onset schizophrenia: effect of risperidone. *Progress in Neuro-psychopharmacology and Biological Psychiatry, 39*, 348–354.

Roche, E., Creed, L., MacMahon, D., Brennan, D., & Clarke, M. (2015). The epidemiology and associated phenomenology of formal thought disorder: a systematic review. *Schizophrenia Bulletin, 41*, 951-962.

Roche, E., Lyne, J., O'Donoghue, B., Segurado, R., Behan, C., Renwick, L., ... & Clarke, M. (2016). The prognostic value of formal thought disorder following first episode psychosis. *Schizophrenia Research, 178*(1-3), 29-34.

Racenstein, J.M., Penn, D., Harrow, M., & Schleser, R. (1999). Thought disorder and psychosocial functioning in schizophrenia: The concurrent and predictive relationships. *The Journal of Nervous and Mental Disease*, *187*(5), 281-289.

Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., & Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology & Psychotherapy*, *10*(4), 238-261.

Sivec, H.J., & Montesano, V.L. (2012). Cognitive behavioral therapy for psychosis in clinical practice. *Psychotherapy*, *49*, 258-270.

Smith, T.E., Hull, J.W., Huppert, J. D., Silverstein, S.M., Anthony, D.T., & McClough, J.F. (2004). Insight and recovery from psychosis in chronic schizophrenia and schizoaffective disorder patients. *Journal of Psychiatric Research*, *38*(2), 169-176.

Suttajit, S., Arunpongpaisal, S., Srisurapanont, M., Thavichachart, N., Kongsakon, R., Chantakarn, S., ... & Nerapusee, O. (2015). Psychosocial functioning in schizophrenia: are some symptoms or demographic characteristics predictors across the functioning domains? *Neuropsychiatric Disease and Treatment*, *11*, 2471.

Tagamets, M.A., Cortes, C.R., Griego, J.A., & Elvevag, B. (2014). Neural correlates of the relationship between discourse coherence and sensory monitoring in schizophrenia. *Cortex*, 77-87.

Tan, E.J., Thomas, N., & Rossell, S.L. (2014). Speech disturbances and quality of life in schizophrenia: Differential impacts on functioning and life satisfaction. *Comprehensive Psychiatry*, *55*(3), 693-698

Tully, L.M., Lincoln, S.H., Liyanage-Don, N., & Hooker, C.I. (2014). Impaired cognitive control mediates the relationship between cortical thickness of the superior frontal gyrus and role functioning in schizophrenia. *Schizophrenia Research*, *152*(2-3), 358-364.

Van Donkersgoed, R.J.M., DeJong, S., Van der Gaag, M., Aleman, A., Lysaker, P.H., Pijnenborg, G.H.M. (2014). A manual- based individual therapy to improve metacognition: Protocol of a multi-center RCT. *BMC Psychiatry, 14*, 27.

Willits, J.A., Rubin, T., Jones, M.N., Minor, K.S., & Lysaker, P.H. (2017). Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophrenia Research*.

Yalınçetin, B., Ulaş, H., Var, L., Binbay, T., Akdede, B.B., & Alptekin, K. (2016). Relation of formal thought disorder to symptomatic remission and social functioning in schizophrenia. *Comprehensive Psychiatry*, *70*, 98-10

APPENDIX A. MERIT CORE ELEMENTS

1. **Agenda**: involves attending to the patient's immediate wishes and desires
    a. Human behavior is purposeful and all patients are seeking *something* when they come to a session
2. **Dialogue**: involves therapists' sharing their thoughts about patients' mental activities and behaviors without overriding patients' agenda
    a. Patients' offered a chance to reflect about what they think about therapist's presence
3. **Narrative Focus**: involves attending to and reflecting with patients' about their sense of themselves and others within the flow of life
    a. Forge mutual understanding with patients, viewing their experiences as comprehendible
4. **Psychological Problem:** involves attending to patients' sense of the psychological and social challenges they face.
    a. Joint reflection about potential and meaningful problems; does not necessarily mean agreement about the "correct" or "true" problem
5. **Reflection upon Interpersonal Process:** involves attending to patients' sense of how they are relating to the therapist.
    a. Patients encouraged to reflect upon larger interpersonal processes that are taking place in sessions between therapist and client
6. **Perceptions of Change:** involves the therapist attending to the patients' sense of what they are expecting as it is happening within the session
    a. Patients directly invited to reflect upon what is changing and not changing as a result of therapy

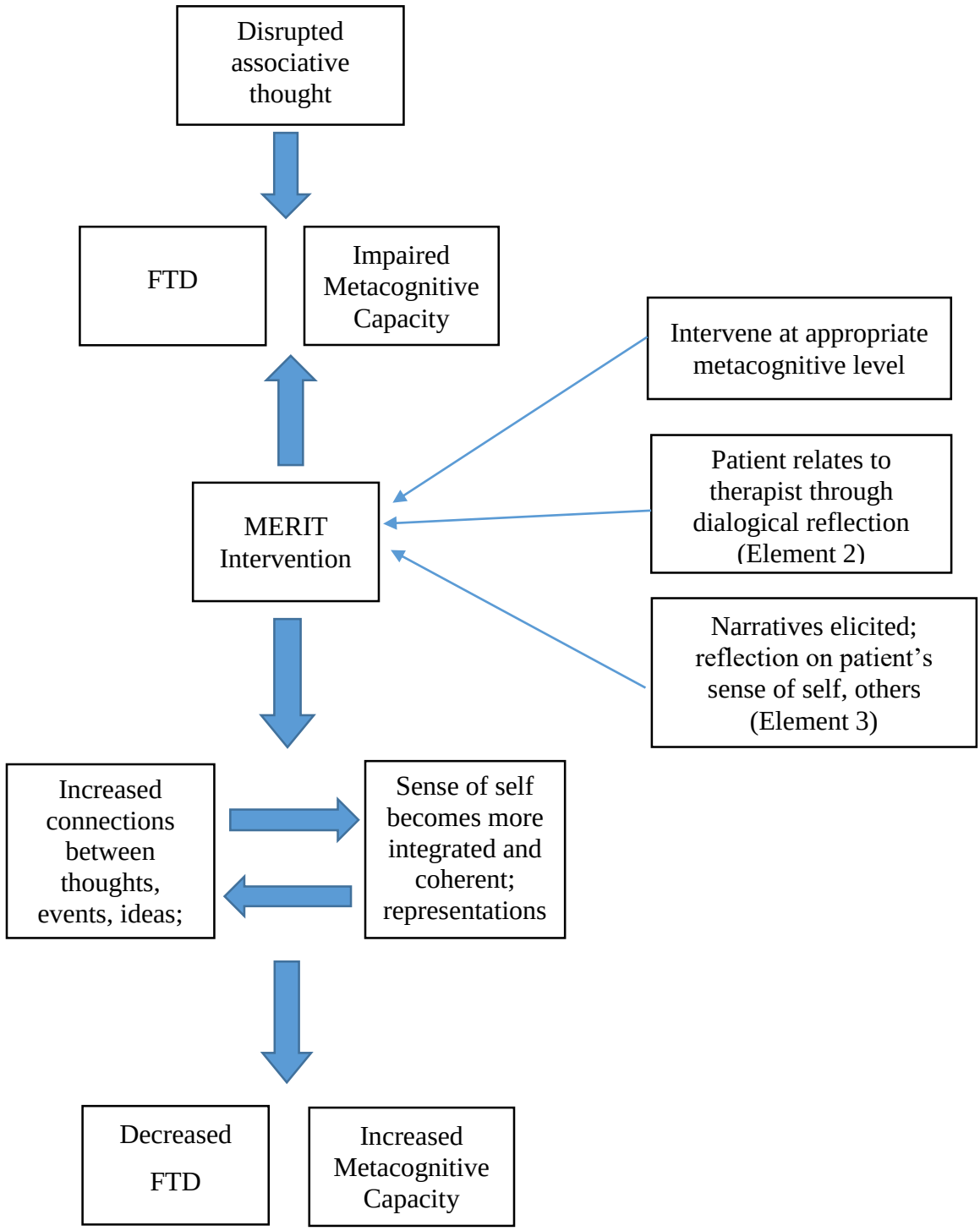
7. **Optimal Stimulation of Reflections About Self and Others:** involves the therapist ensuring that when patients are stimulated to think about themselves and others, the stimulation matches their maximal capacity for metacognitive activity
8. **Optimal Stimulation of Metacognitive Mastery:** involves therapists attending to the patient's use sense of self and others to respond to psychological and social challenges

APPENDIX B. AUTOMATED FTD DESCRIPTIONS

Coh-Metrix Index	Description of indices
Narrativity	Measures whether the text contains characters, events, and places that are familiar to others, and is closely affiliated to word familiarity and with "everyday", oral conversations
Syntactic Simplicity	Reflects the degree to which spoken sentences use simple, familiar syntactic structures that are easy to process
Word Concreteness	Reflects the degree to which content words are concrete, meaningful and easy to understand, rather than abstract and difficult to comprehend
Referential Cohesion	Measures the degree to which words and ideas connect across sentences and throughout conversations
Deep Cohesion	Reflects the degree to which the text contains causal and logical links are present to assist others in forming a deeper and more coherent understanding

APPENDIX C. FIGURES

Figure 1. Conceptual model of MERIT's expected effect on FTD



APPENDIX D. TABLES

Table 1

Participant Data for Aim 1 and Aim 2 Samples

	Aim 1 Sample (n = 20)	Aim 2 Sample (n = 31)
	n, %	*n*, %
Gender (*n*, % Female)	11, 55%	17, 55%
Race		
Black	15, 75%	19, 61%
White	5, 25%	10, 32%
Multiple Races	-	2, 7%
Ethnicity		
Non-Hispanic	18, 90%	27, 87%
Hispanic	-	1, 3%
Unknown	2, 10%	3, 10%
Education		
<HS	6, 30%	10, 32%
HS/GED	3, 15%	4, 13%
>HS	11, 55%	16, 55%
Employment		
Unemployed	15, 75%	25, 81%
Volunteer	1, 5%	1, 3%
<20hrs/week	1, 5%	2, 6%
>20hrs/week	3, 15%	3, 10%
Previous Therapy		
Yes	17, 85%	27, 87%
No	3, 15%	4, 13%
	M (SD)	*M (SD)*
Age	44.25 (10.95)	45.97 (9.80)
Symptoms		
Total Positive	12.30 (3.71)	13.13 (4.29)
Total Negative	15.90 (5.76)	16.71 (6.14)
Total Disorganized	14.45 (4.36)	13.84 (3.98)
Social Functioning	5.60 (0.60)	5.84 (1.16)
Role Functioning	3.70 (2.34)	3.26 (2.19)
Metacognition		
Total MAS	12.5 (2.92)	11.77 (3.28)
Self-Reflectivity	5.18 (1.36)	4.79 (1.42)
Other	3.30 (0.62)	3.18 (0.93)
Decentration	0.58 (0.50)	0.52 (0.51)
Mastery	3.40 (1.31)	3.29 (1.27)

Note. MAS = Metacognitive Assessment Scale – Abbreviated; Other = Awareness of the other's mind

Table 2

Participant data for study completers (n = 20) and non-completers (n = 11)

	Completers (n = 20)	Non-completers (n = 11)	Test of Significance
	n, %	*n*, %	
Gender (*n*, % Female)	11, 55%	6, 55%	X^2 (1) = *.001*
Race			X^2 (1) = *2.57*
Black	15, 75%	4, 36%	
White	5, 25%	5, 46%	
Multiple Races	-	2, 18%	
Ethnicity			X^2 (1) = *0.42*
Non-Hispanic	18, 90%	9, 82%	
Hispanic	-	1, 9%	
Unknown	2, 10%	1, 9%	
Education			X^2 (2) = 0.28
<HS	6, 30%	4, 36%	
HS/GED	3, 15%	1, 9%	
>HS	11, 55%	6, 55%	
Employment			X^2 (1) = *1.15*
Unemployed	15, 75%	10, 91%	
Volunteer	1, 5%	-	
<20hrs/week	1, 5%	1, 9%	
>20hrs/week	3, 15%	-	
Previous Therapy			X^2 (1) = *0.22*
Yes	17, 85%	10, 91%	
No	3, 15%	1, 9%	
	M (SD)	*M (SD)*	
Age	44.25 (10.95)	49.09 (6.60)	*t* (29) = -1.33
Symptoms			
Total Positive	12.30 (3.71)	14.64 (5.01)	*t* (29) = -1.48
Total Negative	15.90 (5.76)	18.18 (6.81)	*t* (29) = -0.99
Total Disorganized	14.45 (4.36)	12.73 (3.04)	*t* (29) = 1.16
Social Functioning	5.60 (0.60)	6.27 (1.74)	*t* (29) = -1.58
Role Functioning	3.70 (2.34)	2.36 (1.63)	*t* (29) = 1.68
Metacognition			
Total MAS	12.5 (2.92)	10.82 (3.57)	*t* (29) = 1.42
Self-reflectivity	5.18 (1.36)	4.27 (1.29)	*t* (29) = 1.81+
Other	3.30 (0.62)	2.95 (1.33)	*t* (29) = 1.00
Decentration	0.58 (0.50)	0.50 (0.50)	*t* (29) = 0.43
Mastery	3.40 (1.31)	3.09 (1.22)	*t* (29) = 0.65

Note. All p's >0.05; MAS=Metacognitive Assessment Scale – Abbreviated; Other = Awareness of the other's mind

+$p < 0.10$

Table 3

Aim 1 analyses: Paired-samples t-tests comparing baseline versus post-intervention FTD ratings (n = 20)

FTD Measure	Baseline *M (SD)*	Post-intervention *M (SD)*	Effect Size Cohen's *d*
PANSS: CD	1.70 (0.98)	1.75 (1.07)	0.06
Automated Coh-Metrix[a,b]			
Composite[c]	0.80 (0.33)	1.01 (0.28)	0.57*
Narrativity	2.32 (0.62)	2.44 (0.45)	0.29
Syntactic Simplicity	-0.08 (0.64)	0.01 (0.85)	0.10
Word Concreteness	-0.52 (0.78)	-0.36 (0.47)	0.22
Referential Cohesion	1.36 (0.97)	1.63 (0.72)	0.21
Deep Cohesion	0.95 (0.63)	1.32 (0.67)	0.64*

Note: PANSS: CD = Conceptual Disorganization item on the Positive and Negative Syndrome Scale

* *t*-test *p*-value < 0.05

[a] For automated variables, higher values indicate less FTD

[b] Coh-Metrix-generated Z-scores are reported and were used in all analyses

[c] Composite variable was calculated by averaging the z-scores of the 5 automated variables

Table 4

Sensitivity Analysis: paired samples t-tests comparing baseline and post-intervention FTD Ratings; (n = 8)

FTD Variable	Baseline *M (SD)*	Post-intervention *M (SD)*	Effect size Cohen's *d*
PANSS:CD	2.75 (0.71)	2.5 (1.20)	0.21
Automated Coh-Metrix[a,b]			
Composite[c]	0.80 (0.25)	1.00 (0.27)	0.90*
Narrativity	2.41 (0.54)	2.52 (0.54)	0.51
Syntactic Simplicity	-0.08 (0.25)	0.15(0.70)	0.33
Word Concreteness	-0.41 (0.48)	-0.43 (0.61)	0.04
Referential Cohesion	1.27 (0.97)	1.31 (0.64)	0.05
Deep Cohesion	0.79 (0.74)	1.45 (0.55)	1.57*

Note: PANSS: CD = Conceptual Disorganization item on the Positive and Negative Syndrome Scale

* p-val < 0.05

[a] For automated variables, higher values indicate less FTD

[b]Z-scores are reported and were used in all analyses

[c]Composite variable was calculated by averaging the z-scores of the 5 automated variables

Table 5

Aim 2: Bivariate correlations and Fisher's r to Z transformations (Fisher's Z-test) comparing associations between FTD and social functioning, role functioning and metacognition (n=31)

FTD Variable	Social Functioning *r*	Fisher's Z	Role Functioning *r*	Fisher's Z	Metacog-nition *r*	Fisher's Z
PANSS: CD[a]	0.18	-	-0.24	-	0.05	-
Automated Composite	0.05	0.523	-0.45*	0.924	-0.09	0.558

Note: PANSS: CD = Conceptual Disorganization item on the Positive and Negative Syndrome Scale

[a]PANSS: CD correlations were reverse-coded such that positive correlations indicate that less FTD is associated with higher levels of functioning

* $p < .05$