

University of Groningen

Social cognition and treatment in psychosis

van Donkersgoed, Rozanne Janna Margaretha

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van Donkersgoed, R. J. M. (2018). Social cognition and treatment in psychosis. [Groningen]: Rijksuniversiteit Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

CHAPTER 7

Metacognitive Reflection and Insight Therapy (MERIT) with a Patient with Persistent Negative Symptoms

De Jong S, Van Donkersgoed RJM*, Aleman A, Wunderink L,
Van Der Gaag M, Lysaker PH, Pijnenborg GHM.
(*contributed equally)*

Accepted for publication in Psychological Medicine

Abstract

Introduction: Impaired metacognition is associated with several difficulties in daily life functioning of people with psychosis. Metacognition can be divided into four domains: Self-Reflection, Understanding the Other's Mind, Decentration and Mastery. Several case studies and two pilot studies demonstrated metacognition to be a feasible target for improvement using therapy. The present study investigates whether Metacognitive Reflection and Insight Therapy (MERIT) can be used to improve metacognition in patients with schizophrenia.

Methods: This study was designed as a randomized controlled trial. Patients in the active condition (n=35) received MERIT, the control group (n=35) received treatment as usual. Multilevel intention-to-treat analysis as well as sensitivity analysis was performed for the four domains of metacognition and secondary outcomes empathy, depression, stigma, social functioning and quality of life.

Results: Intention-to-treat analysis demonstrated that both groups improved between pre- and post-measurements, with no significant differences between the groups: MERIT was not more effective than TAU. However patients who received MERIT continued to improve on metacognitive Self Reflectivity, while performance of the control group dipped back down, leading to significant differences at follow-up. Almost half of the patients in the active condition did not complete the therapy (17/35). Sensitivity analysis including only completers showed improvements on Self Reflectivity and metacognitive Mastery at follow-up. No improvements were found on Understanding the Other's Mind and Decentration.

Conclusion: Significant gains in the MERIT group compared to TAU are only demonstrated at follow-up. Participants in the MERIT group were at follow-up on average more likely to recognize their thoughts as subjective and changeable rather than regarding them as facts. The large drop-out rate in the therapy group might be the result of the duration of therapy (forty sessions). MERIT might be a useful treatment approach for patients whose self-reflection is too limited to benefit from other therapies. Further Limitations and suggestions for future research are discussed.

Introduction

Many persons with schizophrenia have impaired metacognitive capacity; the ability to reflect on thoughts and feelings and to integrate these reflections into detailed representations of oneself and others (Frith, 1992; Lysaker et al., 2011; Hamm et al., 2012; Lysaker et al., 2014). Metacognition can be divided into four semi-independent domains: Self-Reflectivity, Understanding the Other's Mind, Decentration - the ability to understand that one is not at the center of all meaningful activity, and Mastery - the ability to use metacognitive information to deal with stressors (Lysaker, Erickson, et al., 2011; Semerari et al., 2003).

Metacognitive dysfunction is associated with problems in daily life functioning of people with schizophrenia in several ways. Lower levels of metacognition are correlated with lower levels of functional competence (Lysaker, McCormick, et al., 2011), lessened subjective experience recovery (Kukla, Lysaker, & Salyers, 2013), more severe negative symptoms (Hamm et al., 2012; Lysaker, Carcione, et al., 2005; Macbeth et al., 2014; Nicolò et al., 2012) and lower quality of the therapeutic alliance between patient and therapist (L. W. Davis, Eicher, & Lysaker, 2011). Furthermore; social cognition and insight are positively associated with metacognitive mastery (Lysaker, Erickson, et al., 2011). Moreover, metacognition mediates the impact of neurocognitive deficits on social function, even after controlling for symptoms (Lysaker, Shea, et al., 2010).

Metacognition might be a valuable target for treatment of patients with schizophrenia, as improved metacognition may lead to improved daily life functioning. Studies suggest that individual therapy can be a useful form of treatment to improve metacognition. Several forms of individual therapy have successfully improved metacognition in patients with various mental disorders other than psychosis (Choi-Kain & Gunderson, 2008; Dimaggio, Semerari, Carcione, Nicolò, & Procacci, 2007; Fonagy, Gergely, & Jurist, 2002). Several case studies (Brent, 2009; Buck & Lysaker, 2009; de Jong, van Donkersgoed, Pijnenborg, & Lysaker, 2016; Lysaker, Davis, et al., 2005; Lysaker, Buck, & Ringer, 2007; Salvatore et al., 2009; Salvatore, Russo, Russo, Popolo, & Dimaggio, 2012; van Donkersgoed, de Jong, & Pijnenborg, 2016) and two pilot studies (Bargenquast & Schweitzer, 2013; de Jong, van Donkersgoed, Aleman, et al., 2016) have reported improvement of metacognition after individual therapy in people with psychosis.

Lysaker, Buck et al. (2010) proposed a manualized procedure to improve metacognition in people with schizophrenia. The current paper presents a randomized controlled trial investigating the effectiveness of this Metacognitive Reflection and Insight Therapy (MERIT) (Van Donkersgoed et al., 2014), which was conducted based on a pilot study (de Jong, van Donkersgoed, Aleman, et al., 2016).

Methods

The protocol for this study was registered (ISRCTN16659871) and published (Van Donkersgoed et al., 2014). Medical-Ethical approval was obtained via the local ethical board (METc2013.124). All research was conducted in accordance to the principles of the Declaration of Helsinki.

Therapy

The method under investigation, MERIT, aims to stimulate the four elements of metacognition. The treatment protocol is not a step-by-step program, but is driven by the level of metacognition of the patient. The therapist elicits a personal story (narrative) of the patient. In this narrative, the therapist looks for signs of metacognition. Is the patient aware of his thoughts? Can he/she reflect on those thoughts and on the thoughts of others? Does he/she identify and frame psychological distress? The scales of the Metacognitive Assessment Scale (MAS-A, see materials) are used to classify the level of metacognitive functioning. The therapist adjust his or her interventions according to the level of metacognition of the patient and stimulates the patient to perform more complex metacognitive tasks, using eight specific treatment elements (T-MAS, see appendix). The therapy consists of forty individual therapy sessions. The treatment protocol has been translated into Dutch by the research team.

Therapists

Thirteen therapists across seven mental healthcare institutes in the Netherlands were recruited. At minimum all therapists had a master education (MSc) in Clinical Psychology and practical experience in the field, most of them (eleven) had also obtained the post-master health-care-license required for clinical practice in the Netherlands. Therapists received a three-day training program in the Metacognitive Reflection and Insight Therapy (MERIT), delivered by the first author of MERIT, P.H. Lysaker. Once every two weeks a group supervision session by Lysaker was organized for all therapists via internet telephony, in which the therapists received feedback on how they applied the method.

Participants

Patients in the participating treatment facilities were screened on metacognitive difficulties using four screening questions, developed based on the four domains of metacognition (e.g. "To what extent is the patient able to think about his/her own thoughts?"). Answers were given on a Likert scale 0-10, with higher scores reflecting better functioning. These questions were completed by the case manager or staff member most familiar with the patient. Those patients who scored <6 on two or more of the screening questions, were subsequently approached in

person and received basic information and an information letter regarding the study.

Inclusion criteria:

Impaired metacognitive abilities (determined using the MAS-A, see instruments)
 Diagnosis of Schizophrenia or Schizoaffective Disorder according to DSM-IV-TR (MINI-PLUS)
 Being able to give informed consent
 18 years or older
 No change in medication in the thirty days before first assessment

Exclusion criteria:

Acute psychosis at the moment of assessment
 Co-morbid neurological disorder
 Diagnosis of severe substance dependence, but not abuse
 Impaired intellectual functioning ($IQ < 70$)

Interested participants were administered a baseline assessment composed of two meetings with a research assistant. In the first meeting the inclusion and exclusion criteria were verified with the MINI-Plus, IPII, MAS-A and PANSS interview (for materials see below). After inclusion, participants were administered the remainder of the test battery in a second meeting. To ensure blind randomization, an independent third party performed block randomization procedures (Kazdin, 2010) to ensure groups equivalent in size. See figure 1 for a CONSORT diagram detailing participant flow.

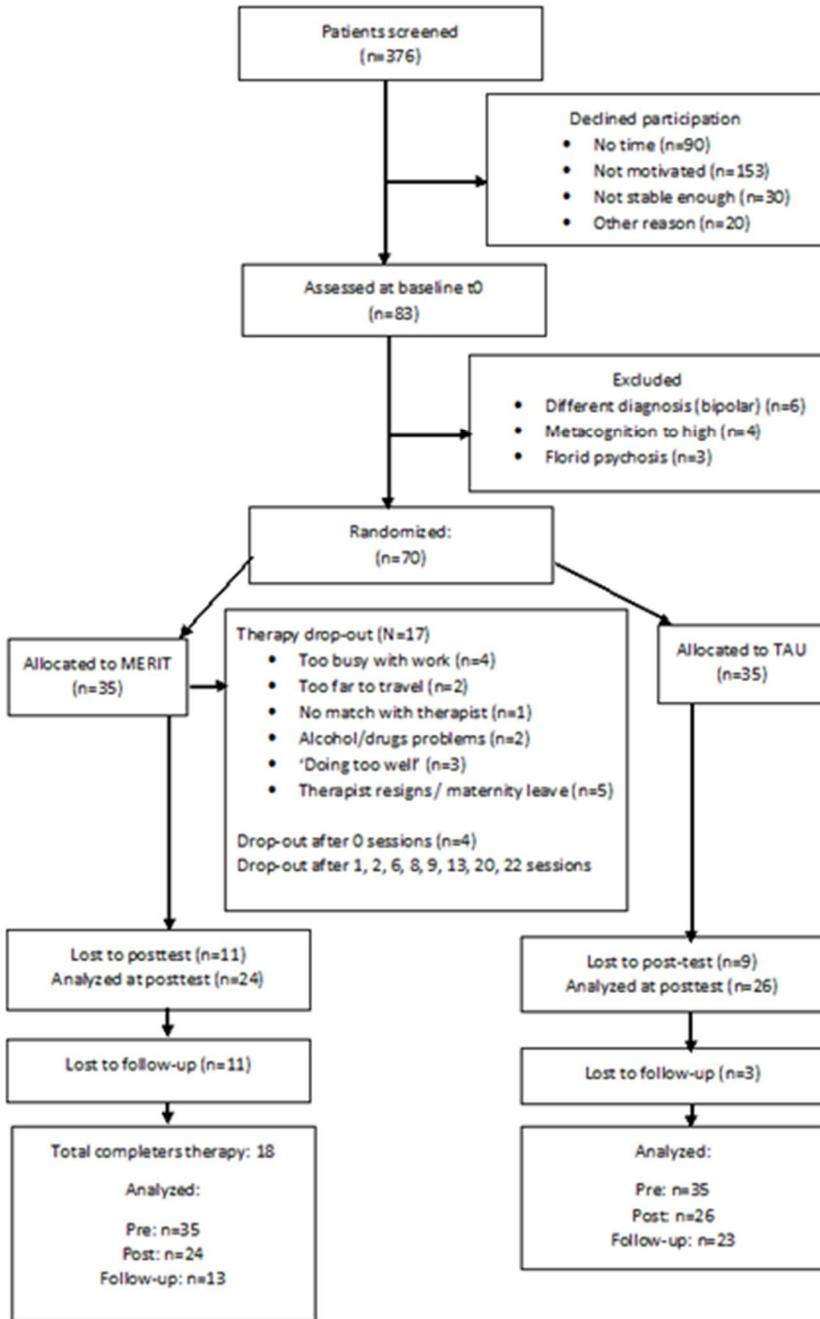


Figure 1. CONSORT diagram of participant flow.

Assessment

All research assistants held at least a bachelor's degree in clinical psychology and were blinded to participant condition. Assessment occurred at three moments: T0 (baseline), T1 (directly following treatment) and T2 (6-month follow-up).

Primary outcome: metacognition

Metacognition Assessment Scale (MAS-A; Lysaker, Carcione, et al., 2005). To assess metacognitive functioning, the Indiana Psychiatric Illness Interview (see below) was conducted and transcribed. Three raters blind to condition and trained in the MAS-A during a 4-hour training, scored this transcript on metacognitive capacity along four axes: Self-Reflectivity (scores 0 (low) - 9 (high)), Understanding the Other's Mind (scores 0 (low) -7 (high)), Decentration (scores 0 (low) -3 (high)) and Mastery (scores 0 (low) -9 (high)). During consensus meetings, final scores on each of the four domains were established. Total scores are analyzed, followed by analyses to determine on which specific domains improvements were found.

Indiana Psychiatric Illness Interview (IPI; Lysaker, Carcione, et al., 2005). The IPI is a semi-structured interview developed to elicit a speech sample during which participants can demonstrate metacognitive capacity. Interviews last between 20 and 60 minutes, and consist of five sections: life narrative, illness narrative, experience of mental illness, the influence of illness on one's life, and the future. The interview is converted into a transcript, which is used to score the level of metacognition of the participant using the MAS-A (see above).

Secondary outcomes

Beck Cognitive Insight Scale (BCIS; Beck, Baruch, Balter, Steer, & Warman, 2004). This 15-item questionnaire measures cognitive insight along the subscales of self-reflectiveness (9 items) and certainty (6 items) using a 4-point Likert scale. A total score is obtained by subtracting the Self Certainty score from the Self-Reflectiveness score, resulting in an Index of cognitive insight (with higher scores indicating better insight), which has demonstrated promising psychometric qualities, including convergent and criterion validity (Riggs, Grant, Perivoliotis, & Beck, 2012).

Clinical Global Impression (CGI; Haro et al., 2003). This rating scale allows for the assessment of the participant's current functioning, along the domains of positive symptoms, negative symptoms and general symptoms using 7 anchor points per scale, ranging from "Not ill" to "Among the most severely ill".

Empathic Accuracy Task (EAT): To measure empathic accuracy we used a Dutch version of the EAT developed by Zaki et al., 2008. A shorter version than the original Dutch task described by (Aan Het Rot & Hogenelst, 2014) was used, this

was necessary to keep the total assessment battery under two hours. The original task was shortened by selecting four out of the ten original videos. Participants were required to continuously rate the valence (positive-negative) of the videos in which a target tells a personal story, using a turning device. Scores of the participants are correlated with the target's own ratings, leading to an index of empathic accuracy. Level of expressivity of the targets is based on their score on the Berkeley Expressivity Questionnaire (BEQ; Gross & John, 1995).

Faux-Pas Test (FPT; Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). During this test of Theory of Mind, ten stories are read aloud to the participant, who can read along using a printed-out version of the story. The participant is asked whether a socially undesirable action was taken by one of the participants, or not, and how the participant in the story must have felt, resulting in 2 scores: the number of faux pas correctly identified (min 0-max. 5) and empathy questions ("How does person X in the story feel") answered correctly (min 0 - max. 5). Interpersonal Reactivity Index (IRI; M. H. Davis, 1983). Using 28 items to be answered on a six-point Likert Scale, this questionnaire measures empathy, with a higher score indicating greater self-reported empathy.

Internalized Stigma of Mental Illness Scale (ISMI; Boyd Ritsher, Otilingam, & Grajales, 2003). The ISMIS measures self-reported internalized stigma of mental illness using 29-items on a 4-point Likert scale. Higher scores are indicative of a greater experience of self stigma. Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). This well-validated structured interview is designed to measure the presence of neuropsychiatric disorders. Sections A through D (mood disorders), K through L (substance abuse) and M (psychotic disorders) were administered to verify in- and exclusion criteria for the study.

Positive and Negative Syndrome Scale (PANSS; Kay, Fiszbein, & Opler, 1987). This semi-structured interview was employed by trained raters to indicate the severity of 30 symptoms using a 7-point Likert Scale, ranging from "Absent" to "Extreme", resulting in a total score between 30 and 210, with higher scores indicating more severe symptomatology.

Personal and Social Performance Scale (PSP; Nasrallah, Morosini, & Gagnon, 2008). Using this rating scale, interviewers rate the impact of the disorder on four domains of social functioning on a 6-point Likert Scale ranging from "absent" to "very severe". Results are converted in a 1 - 100 score of severity, with higher scores indicating more severe impact of the disorder on functioning.

Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers, Corcoran, Drake, Shryane, & Völlm, 2011). Based on factor analysis of several common self-report measures (including the IRI), the QCAE measures self-reported empathy. It consists of 31 items, answered on a 4-point Likert scale, with higher scores indicating greater self-reported empathy.

Quick Inventory of Depressive Symptomatology – Self Report (QIDS-SR; Rush et al., 2003). The QIDS-SR measures depressive symptoms during the last week, using 16-items based on the DSM-IV-TR criteria for Major Depressive Disorder, answered on a 4-point Likert scale. A higher total score indicates greater severity of depressive symptoms.

Self-Rated Manchester Short Assessment of Quality of Life (MANSA; Priebe, Huxley, Knight, & Evans, 1999). Using twelve subjective and four objective questions answered on a 7-point Likert scale, this questionnaire allows the participant to indicate general life satisfaction along several domains, with higher scores indicating greater satisfaction.

Cognition measures

Dutch Adult Reading Test (DART; Schmand et al., 1991). The DART tests the pronunciation of irregularly spelled words and is used to estimate pre-morbid intelligence.

Trailmaking test A&B (TMT; Reitan & Wolfson, 1985). The TMT provides information on visual search, scanning, mental flexibility speed of processing and executive functions. It is part of the Halstead-Reitan Battery. The TMT consists of two parts. Part A requires an individual to draw lines sequentially connecting 25 encircled numbers distributed on a sheet of paper. Task requirements are similar for Part B except the person must alternate between numbers and letters (e.g., 1, A, 2, B, 3, C, etc.). The final score is determined by subtracting the time to complete task A from the time it took to complete task B, with higher scores indicating lower cognition (Tombaugh, 2004).

Digit Symbol Test (part of the Wechsler Adult Intelligence Scale; Wechsler 1995). This test evaluates the recognition and recoding of visual information. The test consists of several rows of paired boxes with a digit in the top box and an empty space in the box below. At the top of the page is shown which symbols are paired to the digits. The participant has to fill in as many symbols in the empty boxes within 90 seconds. The final score consists of the amount of symbols that is filled in correctly within the time, with a higher score indicating better cognition.

Design and statistical analyses

The study was designed as a multicenter randomized controlled trial with a treatment condition in which participants received MERIT, and a control condition in which participants received treatment as usual. Patients in the control group and in the MERIT group met once a month on average with their psychiatrist for medication monitoring and received practical guidance (for example with finances or work related problems) from a social worker. Two out of 35 participants in the control group met with a psychologist during the period between pre and post measurements. Four participants met with a psychologist in

the period between post and follow-up measurements. Patients in the treatment group did not receive any additional psychosocial interventions apart from the MERIT therapy. Participants and their psychiatrists were asked to keep medication changes limited to only crucial adjustments until study end. Data were collected at baseline (T0), post-treatment (T1) and after 6 months at follow-up (T2). Participants received €20 for each completed assessment.

Demographic differences between groups were tested using SPSS Statistics 24 with independent-samples t-tests (age, age at onset of first psychosis, number of psychotic episodes, duration of illness, estimated premorbid IQ, cognition and symptoms) or Pearson's Chi-Square test (gender, diagnosis, education level). These were conducted two-tailed, with significance levels set at $p=0.05$. As reported in Table 1, none of the demographic variables demonstrated statistically significant differences between the groups, and as such none were entered into subsequent analyses.

The effects of the treatment on outcome measures were assessed with multilevel analysis, using MLwiN (Charlton, Rasbash, Browne, Healy, & Cameron, 2017). A separate 3-level model was constructed for each of the outcome variables: Therapists were modelled at level 3, participants at level 2, and time of assessment at level 1. The following predictors were entered as fixed effects: a) dummy variables representing time (T0, T1, T2); and b) the interactions T1*condition and T2*condition. The random effects were the intercepts at levels 2 and 3, and residual at level 1. To assess whether the MERIT group had improved more than the control group at T1 and T2, significance testing was conducted using deviance tests (Snijders & Bosker, 2000) between the models with the interaction between the time of assessment under investigation (T1 or T2) and condition (MERIT/TAU), and a model without the interaction terms. Subtraction of the $-2 \times \text{Loglikelihood}$ value of model B from model A yields a chi-square value which can be compared to the chi-square distribution to test whether the difference between the models is significant, with significance level set at $p<0.05$, one-tailed. An intention to-treat analysis was conducted on the entire sample, followed by a sensitivity analysis in which only the results were modeled of those participants who had completed the therapy.

Results

Demographics

In total, 70 participants were included in the study (figure 1), distributed evenly among the two conditions. None of the demographic variables differed significantly between the groups (Table 1).

Variable	Control	N	MERIT	N	<i>p</i> T-Test / χ^2
Age, mean (SD)	38 (10.61)	35	42 (12.02)	35	.14
Gender		35		35	.43
Male	26		23		
Female	9		12		
Education		35		35	.42
Low	14		11		
Middle	8		13		
High	13		11		
Diagnosis		35		35	.80
Schizophrenia	23		24		
Schizoaffective	12		11		
Age of onset	23.18 (6.26)	34	25.97 (9.31)	33	.15
Episodes, mean (SD)	2.83 (3.04)	30	3.16 (3.07)	31	.68
Years of illness, mean (SD)	12 (9.54)	31	15.53 (11.47)	31	.19
DART*, mean (SD)	77.94 (14.01)	34	78.5 (13.32)	32	.87
Trailmaking, mean (SD)	174.38 (88.88)	34	156.29 (66.19)	35	.34
Digit Symbol, mean (SD)	52.53 (17.51)	34	52.17 (18.28)	35	.93
PANSS* total, mean (SD)	66.29 (17.87)	34	66.17 (15.02)	35	.98

*DART = Dutch Adult Reading Test; PANSS = Positive and Negative Symptom Scale

Table 1. Comparison of demographic variables between the control and MERIT conditions.

Drop-out

Participants were invited for post-measurement and follow-up assessments irrespective of completing all forty sessions of therapy or not. Drop-out in the control condition, as defined by a refusal to take part in the post-measurement and/or follow-up measurement, was 9/35 compared to 11/35 in the MERIT condition for post treatment and 12/35 compared to 22/35 in the MERIT condition for follow-up.

Therapy completion

Eighteen out of 35 participants completed all forty sessions of therapy. Four patients dropped out before the first session of therapy. Others dropped out after 1, 2, 6, 8, 9, 13, 20 and 22 sessions. Reasons for termination of the therapy are noted in Figure 1. No significant differences on the four scales of metacognition were found between the drop-out group and the group that completed the therapy.

Therapist effect

Multilevel analysis did not reveal a significant contribution of the therapist variable as a level to the model, indicating no significant differences between therapists regarding the improvement of metacognition.

Primary outcome

Intention-to-treat analysis (see Table 2) revealed that in both groups metacognition had significantly improved from baseline to post-treatment. Directly after treatment, differences in growth of metacognition were non-significant between the two groups, with the deviance test between a model with and a model without the time (pre-post)* condition (MERIT-TAU) interaction yielding $\chi^2(1) = 0.435, p=.51$. While the total metacognition scores in the control condition dipped back down between post-treatment and 6-month follow-up, the MERIT group continued to improve. At follow-up, differences between the two groups were significant for the MAS-A total score. The addition of the interaction term of time (follow-up)*condition(MERIT / TAU) led to a significant improvement of the model, with deviance tests yielding $\chi^2(1)=3.763, p=.05$. Analyses using the MAS-A subscales as outcome revealed that gains were only significant on the subscale self-reflectivity, with the deviance test yielding $\chi^2(1)=10.295, p=.001$.

Parameter	Self Beta (SE)	Other Beta (SE)	Decentr. Beta (SE)	Mastery Beta (SE)	Total Beta (SE)
Fixed effects					
Time factor					
Baseline	4.23 (0.19)	2.60 (0.10)	1.04 (0.07)	3.18 (0.20)	11.11 (0.50)
Post ^a	0.15 (0.23)	0.02 (0.16)	0.06 (0.11)	0.83 (0.24)	1.08 (0.56)
Post MERIT ^a	0.42 (0.30)	0.30 (0.22)	0.08 (0.15)	-0.04 (0.33)	0.76 (0.75)
Follow-up ^b	-0.06 (0.24)	-0.04 (0.17)	0.16 (0.12)	0.83 (0.26)	0.92 (0.59)
Follow-up ^b	1.22** (0.37)	0.31 (0.26)	-0.05 (0.18)	0.36 (0.37)	1.81* (0.91)
Random effects					
<i>Variances of</i>					
Level 3 - therapist	0.19 (0.14)	0.00 (0.00)	0.00 (0.00)	0.16 (0.16)	1.40 (1.05)
Level 2 - intercept	0.35 (0.15)	0.22 (0.08)	0.10 (0.04)	0.66 (0.22)	3.10 (1.09)
Level 1 - residual	0.87 (0.13)	0.44 (0.07)	0.21 (0.03)	0.94 (0.14)	5.10 (0.77)

^aPost effect: Difference between T0 and T1 (TAU is reference category)

^bFollow-up effect: Difference between T0 and T2 (TAU is reference category)

** = significant at $p < .01$

* = significant at $p < .05$

Table 2. Fixed and random effects on the subscales of the Metacognition Assessment Scale – intention to treat.

Sensitivity analyses (supplement) amplify these findings. When only taking into account those who had completed all 40 sessions of the therapy, differences between the groups in improvements on Self-Reflectivity were significant at post-measurement, with the deviance test between a model with and a model without the time (pre-post)*condition (MERIT-TAU) interaction yielding $\chi^2(1) = 4.219$, $p = .04$. At follow-up, differences between groups were significant for the MAS Total score, the addition of the interaction term of time (follow-up)*condition (MERIT-TAU) led to a significant improvement of the model, with the deviance test yielding $\chi^2(1) = 8.182$, $p = .004$. Analyses using the subscales of the MAS-A indicated that scores on Self-Reflectivity $\chi^2(1) = 12.784$, $p < .01$ and Mastery $\chi^2(1) = 4.793$, $p = .02$ improved.

Secondary outcomes

No sustaining significant differences were found on the secondary outcome measures. In the MERIT condition, at post-measurement, symptoms significantly increased, with deviance tests yielding $\chi^2(1) = 4.278$, $p = .04$, but returned to baseline at follow-up, $\chi^2(1) = .025$, $p = .87$. Tables presenting these results are included as supplemental materials.

Discussion

The current multicenter randomized controlled trial investigated the effectivity of the Metacognitive Reflection and Insight Therapy in improving metacognition. Intention-to-treat analyses showed an improvement in metacognition in both groups, with no significant differences between groups directly post-treatment. A significant improvement was found at follow-up in Self-Reflectivity in patients who received MERIT, when compared to controls. Self-Reflectivity is an important element of metacognition as it is correlated with daily life factors such as subjective sense of recovery (Kukla et al., 2013) and work performance (Lysaker, Dimaggio, et al., 2010). On average, scores on the MAS-A indicated that patients at baseline were able to recognize and distinguish between their different thoughts and emotions, but did not perceive their thoughts as subjective and changeable. In other words: thoughts were accepted as facts. After MERIT, group average scores indicated having moved past being able to recognize that the ideas about oneself and the world are subjective and changeable. It must be noted that this was the average score of the group. There was considerable variance between participants, some people still didn't perceive their thoughts as subjective and changeable after therapy, where others scored far higher after therapy, being able to recognize and connect thoughts and feelings in the moment and over time.

Sensitivity analyses, only including the patients that finished the therapy, showed significant differences on Self-Reflection between groups post-treatment, with better scores in the treatment condition. At the six month follow-up significant

improvements on Self-Reflection and on Mastery were found. At baseline patients' scores indicated that patients in both conditions on average responded to psychological challenges through gross avoidance and passive activities, such as following other's directions. At follow-up, in the MERIT condition, patients' scores indicated that they were able to respond to psychological challenges by actively choosing and engaging in specific activities and behaviors such as medication use, or seeking therapeutic interventions. In the control condition, participants had also improved, but only reached a level where avoidance behaviors were either more specific (i.e. avoiding supermarkets instead of staying indoors completely) or seeking social support. Again it must be noted that this represents the average group score. There was considerable variance between participants, some patients in the MERIT group still weren't able to actively choose solutions for their psychological problems after therapy, whereas others were not only able to change their thoughts to deal with problems but were also able to use knowledge about their own and others cognitions and emotions to come to solutions. General group scores may suggest that MERIT empowered patients to face their challenges in a more active manner, which may make them less dependent on people around them and may allow them to take a more active role in their treatment.

Difference between conditions only became evident at follow-up. One explanation for this effect is that metacognitive gains take some time to develop, even after therapy has been concluded. Such findings are not uncommon; a meta-analysis of cognitive therapy (Gould, Mueser, Bolton, Mays, & Goff, 2001) has shown continued improvements after therapy was concluded. Improvements in the control condition are not likely to be caused by psychological interventions in this group, as at post assessment two out of 35 patients in the control condition indicated having had any contact with a psychologist. Between post- and follow-up, this number increased to four. It is therefore unlikely that interventions in the control group had significant effects on our findings. Possibly, the finding of the control group's raised performance at post measurement reflect natural fluctuations in metacognitive capacity.

Understanding the Other's Mind and Decentration appear less sensitive to change as no significant effects on these scales were found. This is consistent with results from a pilot study of the intervention (de Jong, van Donkersgoed, Aleman, et al., 2016) and long-term case studies (Lysaker et al., 2007). It is possible that it is necessary to be able to think about your own thoughts and feelings before you can understand and think about what is occurring in the others mind (Dimaggio, Lysaker, Carcione, Nicolò, & Semerari, 2008). Meta-analyses of fMRI findings support this, having found that perception of the self and others share higher-order neural pathways in which these processes are combined (van der Meer et al. 2010; van Veluw & Chance, 2014). From a clinical perspective, it seems logical that Self-Reflection has to improve before someone can start to reflect on the mind of others. When someone is not aware of his own thoughts and feelings, how can he understand those of others? One long-term case study has

found the first improvements to Understanding the Other's Mind to occur after about 16 months (Lysaker et al., 2007). More research on this topic is necessary.

A significant increase of symptoms between baseline and post measurement was found in the MERIT group. This difference was no longer present at follow-up. This disappearance could not be contributed to the drop-out at follow-up, as no significant differences were found on post treatment symptoms between follow-up drop-outs and follow-up completers. No other significant group differences on secondary outcomes were found. It is possible that an improvement in metacognition has no effect on the other variables. However, as multiple studies have shown relationships between metacognition and our secondary outcomes (e.g. Hamm et al., 2012; Lysaker, Shea, et al., 2010; Macbeth et al., 2014), including self-reflectivity specifically (Nicolò et al., 2012), another explanation may entail that more time needs to pass for improved metacognition to positively impact secondary outcomes. For example, it may take a while for someone with improved self-reflection to slowly adjust stigmatic views of oneself to a less stigmatic one. It also may take a while for someone with improved Mastery to find a better job or get to know more friends. Another explanation for the lack of results in our study where other studies find significant relationships might be that the secondary variables influence metacognition, instead of the other way around.

Eighteen out of 35 participants completed all forty sessions of therapy. No significant differences on the four scales of metacognition were found between the drop-out group and the group that completed the therapy. The large amount of drop-outs might be the result of the long duration of the therapy. One of the patients stated "no connection with the therapist" as the reason for dropping out. Other reasons for drop-out were: too busy with work, too far to travel, alcohol/drugs problems and 'doing too well'.

In post-treatment interviews conducted with the patients that completed the therapy, all respondents indicated that they had found the therapy useful ("My wife also noticed I was doing better", "More good things about yourself come to the surface. It isn't just your bad sides. I learned to see myself more positively"), and would recommend it to others. The only negative effect that was mentioned was the intensive nature of the therapy ("After sessions, I often needed rest"), by two out of fifteen participants (13%). As no significant contribution of the therapist variable as a level in the multilevel model was found, improvement of metacognition does not seem to depend on specific therapist characteristics.

The current study has several limitations. We investigated precisely 40 sessions of psychotherapy. A psychosocial intervention such as the one used in this study may not lend itself well for studies with a fixed amount of sessions. Several of our participants indicated their motivation for drop-out as 'doing too well', an observation supported by their therapist. In a clinical setting, the ending of therapy in this way is obviously a good thing. Future studies could account for this issue by setting a minimum and maximum amount of sessions. Furthermore, while an effect was found, power analysis for our study indicated a required 120

participants. However, only 70 could be included. Future studies with sufficient power have to be done to support or reject our findings. Finally, the control condition in the current study received treatment as usual. Future studies should compare metacognitive therapy with other active treatments.

In the present study, scores of participants on the Self-Reflection scale of the MAS-A improved from scores indicating they were generally unable to consider one's own thoughts as changeable and fallible, to a perception of thoughts as subjective. The ability to see thoughts as changeable is necessary to be able to engage in Cognitive Behavioral Therapy (CBT), a widely used treatment in people with a psychotic disorder, which focuses on the modification of maladaptive cognitions (Wykes, Steel, Everitt, & Tarrier, 2008). Some patients may not have the necessary level of self-reflection to be able to engage in CBT techniques. Challenging or changing your thoughts is difficult when you are not aware of them or when you are not aware that they can change over time. MERIT may be useful for patients that do not respond (well) to CBT. It can serve as a way to improve self-reflection after which the patient might be able to benefit from CBT methods. Future studies are needed to verify this hypothesis.

Conclusion

Metacognitive Reflection and Insight Therapy did not improve metacognition post treatment. At follow-up however, self-reflection of participants was improved. That is, there was an average change from seeing their thoughts as facts to recognizing their thoughts as subjective and changeable. MERIT might be a useful treatment approach for patients whose self-reflection is too limited to benefit from other therapies such as CBT. Sensitivity analyses also showed improvement of Mastery at follow-up, suggesting that MERIT may potentially empower patients to face their challenges in a more active manner, which will ultimately give them more control over problems in daily life.

Acknowledgments

The authors would like to express their gratitude to all clients who have participated in the study, and their therapists Rahja König, Els Luijten, Dimitri van Wonderen, Maarten Vos, Lia Elenbaas, Henriëtte Horlings, Sanne Swart, Suzanne van den Bosch, Ilona van den Berg and Inez Oosterholt – Ogink. Similarly, we would like to thank management and staff of all participating MHI: GGZ Drenthe, GGZ Friesland, Lentis, Parnassia, Dimence and Yulius. Furthermore, we would like to thank Ymie Bakker, Daniëlle Bandsma, Kim Doodeman, Ann-Katrin Dresemann, Rianne Hiemstra, Jelle Koehoorn, Christin Koopmann, Saskia Rehrmann, Anne Rupert, Bernice Smit, Hilde Span, Michelle van Dam, Rozemarijn van Kleef and Merel Wattel for their limitless efforts as research assistants. We gratefully acknowledge Selwyn B. Renard for his assistance in the randomization procedures. A final word of thanks to Fonds NutsOhra, who subsidized the study.