A randomized controlled trial of Training of Affect Recognition (TAR) in schizophrenia shows lasting effects for theory of mind

Anja Vaskinn [1,2], André Løvgren [1], Maj K. Egeland [1,2], Frida K. Feyer [1], Tiril Østefjells [1,3],

Ole A. Andreassen [1,4], Ingrid Melle [1,4], Kjetil Sundet [1,2]

[1] Norwegian Centre for Mental Disorders Research, Oslo University Hospital, Oslo, Norway
[2] Department of Psychology, University of Oslo, Oslo, Norway
[3] Department for Specialized Inpatient Treatment, Division of Mental Health and Addiction, Akershus University Hospital, Lørenskog, Norway
[4] Institute of Clinical Medicine, University of Oslo, Oslo, Norway

Corresponding author:
Anja Vaskinn
ORCID: 0000-0003-0339-6170
email: anja.vaskinn@medisin.uio.no telephone: + 47 23 02 73 31
Oslo University Hospital
Division of Mental Health and Addiction
Psychosis Research Unit/TOP
PO Box 4956 Nydalen
0424 Oslo, Norway

Abstract

Background: Schizophrenia is characterized by social cognitive impairments that predict functioning. Social cognitive training aims to target these impairments. Although it can improve the targeted social cognitive domain, it is unclear if the training generalizes to non-targeted domains and to functioning, with lasting effects. This randomized controlled trial examined the effect of a targeted facial affect recognition training program, Training of Affect Recognition (TAR), in persons with schizophrenia.

Method: Individuals with schizophrenia were randomized to receive treatment as usual and TAR (n = 24) or treatment as usual (n = 24) after assessments with a comprehensive protocol at baseline (T1). Participants were reassessed immediately after the intervention period (T2: after 8 weeks) and at 3-month follow-up (T3). The protocol included tests of social cognition (facial or body affect recognition, theory of mind), nonsocial cognition (Matrics Consensus Cognitive Battery), clinical symptoms (Positive and Negative Syndrome Scale, Calgary Depression Scale for Schizophrenia), functioning (self-reported, social or nonsocial functional capacity), self-esteem, self-efficacy and insight.

Results: Linear mixed models yielded a significant group x time interaction effect for a non-targeted social cognitive domain (theory of mind) and a trend-level effect for social functional capacity with the intervention group performing better over time. No beneficial effects on nonsocial cognition, other measures of functioning, clinical symptoms, or self-esteem/self-efficacy appeared for the TAR program.

Conclusions: This study provides evidence for transfer and durability effects of facial affect recognition training to theory of mind, but also highlights the need for additional treatments in order to achieve functional benefits.

Key words
Social cognition, emotion perception, social functioning, social cognitive remediation, functional capacity

Introduction

Individuals with schizophrenia experience difficulties in independent living, social relations and occupational outcome that have been traced back to impairments in social and nonsocial cognition [1]. Research has consistently shown that schizophrenia is characterized by social cognitive impairment [2] although a subgroup may have intact social cognition [3, 4]. Social cognition is among the strong predictors of functional outcome [5] and mediates the effects of nonsocial cognition on outcome [6]. As an effort to improve the outcome of schizophrenia, several social cognitive training programs have been developed.

Social cognition is usually divided into the domains of emotion processing, theory of mind (ToM), social perception/knowledge, and attributional style [7]. Social cognitive training programs target one or several of these domains. Their effectiveness in schizophrenia is summarized in several reviews and meta-analyses [8-12]. The most consistent effect is improvement in the domain targeted by the training. For example, a meta-analysis of randomized controlled trials (RCTs) of targeted facial affect training programs concluded with positive effects on facial affect recognition, no effect on positive, negative or general symptoms, and a possible effect on social functioning [11]. The Training of Affect Recognition (TAR) program [13] has been examined in three RCTs [14-16]. All found larger post-training improvement in facial affect recognition in the treatment group. The only study that investigated transfer effects to other social cognitive domains (prosodic affect recognition, ToM) found it to be significant [15]. The same study found TAR to generalize to social competence in role-play, although this effect disappeared in according-to-protocol analyses of completers [15]. Trend-level improvement in social quality of life was reported in the only other study that examined functioning [16]. This study also found that TAR training was associated with improvement in negative symptoms [16]. Although TAR is among the facial affect training programs that have been investigated the most, its transfer and durability effects remain unclear. Further, the meta-analysis of targeted facial affect training programs considered effect estimates to be of low quality (only 8 RCTs were included of which only 3 assessed social functioning) [11]. Thus, further studies of targeted facial affect training programs are warranted.

In this field, we still lack conclusive evidence that social cognitive interventions have meaningful benefits beyond the targeted domain(s) [17]. It is unclear if the training generalizes to other social cognitive domains and to functional outcome. It is also uncertain if effects endure beyond immediate post-training assessments. In the current

study, we aim to address the issues of generalizability and durability by examining if training of a specific social cognitive domain benefits other social cognitive domains, symptom level and functioning, over time. The RCT (Clinicaltrials.gov Identifier NCT01206842), approved by the regional ethics committee, studies the effect of TAR on a range of outcome measures in a longitudinal design with a 3-month post-training follow-up assessment. A group receiving treatment as usual (TAU) was used as the control condition. Larger improvements in social cognition were expected for the group receiving TAR.

Methods

Participants and procedure

Sixty individuals with a diagnosis of schizophrenia or schizoaffective disorder (DSM-IV) [18], recruited from the Thematically Organized Psychosis (TOP) study at Oslo University Hospital, Norway, provided informed consent. Of these, 12 individuals withdrew before baseline assessment. Study participants (n = 48: schizophrenia = 39, schizoaffective = 9) were assessed at three time points: at baseline (T1), immediately after completion of the intervention period, i.e. approximately 8 weeks later (T2), and 3 months after T2 (follow-up: T3). Participants underwent assessment with a clinical and a cognitive protocol at all assessment points. For details, see *Measures* below. The T1 clinical protocol included diagnostic assessments based on the Structured Clinical Interview for DSM-IV (SCID-I) [19], whereas the T1 cognitive protocol included the 2-test version of the Wechsler Abbreviated Scale of Intelligence (WASI) [20]. After completion of T1 assessment, participants were randomized to receive social cognitive training and treatment as usual (TAR group), or treatment as usual (TAU group). Treatment as usual in Norway is provided by the public health care system and consists of individualized antipsychotic medication, psychoeducation and psychotherapy, often cognitive behavioral therapy [21]. Randomization was based on a plan for 50 participants randomized into 5 blocks from www.randomization.com created in September 2012. Assessors were blind to group allocation. There were no significant group differences at T1 for demographic characteristics (see Table 1) and most study variables (Tables 2 and 3), except for facial affect recognition (TAR > TAU) and self-esteem (TAR < TAU) (see Tables 2 and 3, respectively).

Figure 1 presents the CONSORT flow diagram. In the TAR group, 4 participants withdrew after T1 (baseline assessment and randomization). Another 3 participants withdrew between T2 and T3, one of which did not undergo the cognitive assessment at T2. In addition, 2 participants did not wish to undergo cognitive assessment at T3. These

4

2 participants were not considered as dropouts when calculating attrition rates. For yet 2 participants, some data is missing for T2. The attrition rate in the intervention group is 29.2% (7 out of 24 participants that underwent baseline assessment and randomization). In the TAU group, 4 participants withdrew after T1. Between T2 and T3, 5 participants withdrew from the study, yielding an attrition rate of 37.5% (9 of 24 participants that underwent baseline assessment and randomization). There is some missing data at T2 for 2 of the participants that withdrew between T2 and T3. The overall attrition rate for the current study is 33.3%. Details are provided in Table 2.

Treatment intervention

TAR [13, 15] is a targeted 12-session manualized training for facial emotion perception deficits, delivered over 6 weeks and divided into 3 blocks of 4 sessions. The first block involves discrimination of prototypical facial expressions of basic emotions. In the second block processing is faster, more holistic and nonverbal, while the third block focuses on emotional faces in social and situational contexts. TAR principles include restitution, compensation, establishing of alternative strategies, errorless learning and over-learning. It is computer-aided, but also contains desk top exercises. The manual recommends a 2:1 (participants-trainer) administration format. Due to chances of losing participants that had to wait for a second participant, the training was conducted 1:1. The TAR program was translated into Norwegian by the first author for the purposes of this study.

Measures

Primary outcome measures

Social cognitive tests

The social cognitive test battery consisted of four well-validated tests covering different domains. The Pictures of Facial Affect (PFA) is one of several measures based on the Ekman pictures of facial affect [22]. It has been used in previous TAR studies [14, 15] and was therefore chosen as the measure of *facial affect recognition*. It consists of 28 black and white photographs of Caucasian men or women expressing one of the six basic emotions (happiness, sadness, surprise, disgust, anger, fear) or neutral. The total score (percentage correct) is used. The Emotion in Biological Motion test (EmoBio) [23] measures the ability to identify emotions in moving bodies, i.e. human point-light walkers, or *body affect recognition*. The walkers move in a manner indicative of one of four basic emotions (happiness, sadness, anger, fear) or neutral. The total proportional score (0-1), based on Norwegian normative data

[24], is used. *ToM* was measured with the Norwegian version [25] of the Movie for the Assessment of Social Cognition (MASC) [26]. This is an ecologically valid ToM test based on a video-taped movie, depicting social interactions. Questions regarding the movie characters' thoughts, emotions and intentions are asked. The total overall ToM score was used. *Social perception* was measured with a short, validated version [27] of the Relationship Across Domains (RAD) test [28]. RAD consists of short vignettes describing dyadic social exchanges. The social exchange is characteristic of one of four relational models for how people interpret and organize relationships (communal sharing, authority ranking, equality matching, market pricing). The test taker indicates whether three statements about the relationship of a dyad are correct/incorrect. This assesses their implicit knowledge of relational models.

Secondary outcome measures

Neuropsychological tests

Nonsocial cognition was assessed with the MATRICS Consensus Cognitive Battery (MCCB) [29] excluding the social cognition subtest. A composite T-score was used in the analyses. For one TAR participant data was missing for the Brief Visual Memory Test-R, Continuous Performance Test and Category Fluency at T1. The T1 mean of the TAR group was entered. For another TAR participant, the Mazes subtest was not administered at T3. The TAR group T3 mean was entered for this test. We chose this approach to achieve a composite MCCB score for these two individuals at these two assessment points.

Functional assessment

Functioning in schizophrenia can be differentiated based on level (competency under ideal conditions, e.g. functional capacity, versus actual real-life behavior), domain (social and nonsocial), or source (self-reported or other-evaluated). Global functioning was rated using the GAF, split version function subscale (GAF-f) [30]. Self-reported social functioning was measured with the validated Norwegian version [31] of the Social Functioning Scale (SFS) [32]. Social functional capacity was assessed with a short version [33] of the Assessment of Social Problem-Solving Skills (AIPSS) [34]. The AIPSS consists of three subscales. AIPSSrec (receiving skills) is the ability to identify and describe interpersonal problems from short, videotaped social interactions. AIPSSproc (processing skills) is the verbal response to the question of how this problem may be solved, and AIPSSsend (sending skills) is

the actual attempt at solving the problem in a role-play. Nonsocial functional capacity was measured with the brief version of the UCSD Performance-based Skills Assessment (UPSA-B) [35]. It includes two domains of everyday living activities: Finance (counting change, understanding how a bill is payed), and Communication (making emergency phone calls, reschedule an appointment).

Other outcome measures

Clinical assessment

Participants underwent clinical interviews evaluating ongoing symptomatology. Psychotic and other symptoms were measured with the Positive and Negative Syndrome Scale (PANSS) [36]. We present PANSS data using the five-component solution provided by Wallwork et al. [37] where symptoms are divided into positive (PANSSpos), negative (PANSSneg), disorganized (PANSSdis), excited (PANSSexc) and depressive (PANSSdep) components. Depression was also indexed by the Calgary Depression Scale for Schizophrenia (CDSS) [38]. Global symptom level was rated using the Global Assessment of Functioning, split version symptom subscale (GAF-s) [30].

Subjective outcomes

Participants' self-perception was measured with the Revised Self-Efficacy Scale (RSES) [39] and the Rosenberg Self-Esteem Scale [39]. The 19 social RSES items were used. They describe different social exchanges (call a friend, introduce oneself to new people). The participant is asked to indicate how confident (s)he is that (s)he can perform the social behavior, using a 0-100 scale. The Rosenberg Self-Esteem scale [40] consists of 10 appraisals of self-worth to be rated on a 4-point Likert scale (strongly agree – strongly disagree). In addition, participants completed the Birchwood Clinical Insight Scale (BCIS) [41], a measure of awareness of illness and the ability to relabel abnormal subjective experiences as pathological. It is an 8-item self-report where items are rated on a 5-point Likert scale (strongly agree – strongly disagree).

Statistical analyses

All analyses were undertaken in SPSS. The effect of TAR on primary, secondary and other outcome measures was investigated using linear mixed models. This method was chosen in order to account for missing data. Time (measured in months from baseline; T1 = 0, T2 = 2, T3 = 5), group (TAR, TAU) and time x group were fixed in all

models. All models included random intercept. The inclusion of the 3 assessment points ensures testing of durability effects. Mean values for each of the groups (TAR, TAU) at the three assessment points (T1, T2, T3) were estimated from the linear mixed models for primary outcome measures that yielded significant effects and presented in figures.

Specifically, primary outcome variables, entered as dependent variables in four separate analyses using linear mixed models, were PFA, EmoBio, RAD and MASC. Secondary outcome measures, entered in a series of linear mixed models, were MCCB, GAF-f, SFS, UPSA-B and AIPSS (x3). In addition, the effect of TAR on symptoms (GAF-s, CDSS, PANSS x5) and self-reports (BCIS, RSES and Rosenberg), were examined in similar linear mixed models. All significant effects are reported, but main effects are not interpreted if the interaction effect was significant.

In addition, the actual values for participants that completed the assessments at the three assessment points (for n, see Tables 2 and 3) were subjected to statistical analyses of group differences (independent samples t-tests). Further, between-group effect sizes for T1, T2 and T3 were calculated (Cohen's d). Finally, within-group effect sizes (Cohen's d) for T1 and T3 scores were computed in order to increase the understanding of the clinical significance of changes in either group. The use of effect sizes enables comparison with previous studies. Importantly, in the interpretation of results, we focus on effect sizes, not on the p-values for the independent samples t-tests for T2 and T3. In line with this, we have refrained from performing correction for multiple comparisons. We consider effect sizes ≥ 0.45 as medium, ≥ 0.75 as large.

Results

The actual values at T1, T2 and T3 for those assessed are provided in Tables 2 and 3.

Primary outcome measures

The linear mixed models yielded significant time x group interaction effects for MASC ($b = 0.79$, $t(71.48) = 2.52$, $p = 0.014$). See Figure 2. For PFA, the main effects of time ($b = 1.76$, $t(75.93) = 2.91$, $p = 0.005$) and group ($b = 10.12$, $t(72.11) = 3.03$, $p = 0.003$) were significant. See Figure 3. Effect sizes for group differences at T3 were medium-sized or large for all social cognitive tests (see Table 2). In particular, the effect size for PFA was

substantial (Cohen's d = 1.32) and larger than at T1 (Cohen's d = 0.61). Also, the improvement in PFA performance from T1 to T3 in the TAR group was considerable (Cohen's d = 1.32) and much larger than the improvement in the TAU group (Cohen's d = 0.45). For EmoBio, the between-group medium effect sizes at T2 (Cohen's d = 0.48) and T3 (Cohen's d = 0.47) were larger than at T1 (Cohen's d = 0.30). Further, the within-group effect size for the TAR and TAU groups were medium-sized (Cohen's d = 0.64) and small (Cohen's d = 0.22), respectively. For RAD, the between-group effect size was medium-sized at T3 (Cohen's d = 0.58) and small at T1 (Cohen's d = 0.28)

Secondary outcome measures

The linear mixed models yielded a significant time x group interaction effect for SFS (b = -0.91, $t(71.45)$ = -2.59, p = 0.012). Main effects of time were significant for MCCB (b = 0.60, $t(69.43)$ = 3.90, $p < 0.001$) and SFS (b = 0.85, $t(71.78)$ = 3.32, p = 0.001). For GAF-f and UPSA-B all effect sizes were small. For AIPSSrec, the between-group effect size at T3 was medium-sized (Cohen's d = 0.58), in contrast to the small between-group effect sizes at T1 (Cohen's d = 0.14) and T2 (Cohen's d = 0.28). Also, the within-group effect size for the TAR group (Cohen's d = 0.52) was larger than the very small effect size for the TAU group (Cohen's d = 0.05). The TAR within-group effect sizes for AIPSSproc and AIPSSsend were also medium-sized (Cohen's d = 0.46/0.55).

Other outcome measures

The linear mixed models yielded a significant time x group interaction effect for PANSSneg (b = 0.81, $t(78.73)$ = 2.51, p = 0.014). Main effects of time were significant for GAF-s (b = 1.35, $t(77.12)$ = 2.73, p = 0.008) , PANSSpos (b = -0.32, $t(76.50)$ = -2.18, p = 0.032), PANSSneg (b = -0.49, $t(79.07)$ = -2.09, p = 0.040), and RSES (b = 3.45, $t(75.27)$ = -2.00, p = 0.049). Main effects of group were significant for Rosenberg Self-Esteem (b = -3.25, $t(57.11)$ = -2.11, p = 0.039). None of the between-group effect sizes for clinical symptoms at T3 were medium-sized or larger, but the within-group effect size was medium-sized for PANSSdep in the TAR group (Cohen's d = 0.57), and for CDSS in the TAU group (Cohen's d = 0.68). For the self-report measures the between-group effect size was medium-sized at all assessment points for the BCIS (Cohen's ds = 0.54 or 0.55). RSES differed between the groups at T3 at a medium level (Cohen's d = 0.60), and Rosenberg Self-Esteem presented with a medium-sized between-

group difference at T1 (Cohen's d = 0.65), and a medium-sized within-group improvement for the TAR group (Cohen's d = 0.60).

Discussion

This study examined the effect of a targeted facial affect recognition training program on social cognition, nonsocial cognition, functioning, symptoms and subjective experiences of individuals with schizophrenia. The extent of improvements in the intervention group compared to the control group was investigated with linear mixed models where significant time x group interaction effects are consistent with treatment effects.

Among the primary outcome measures, there was a significant time x group interaction effect for the ToM measure (MASC). Over time, the TAR group performed better, the TAU group worse. This is in line with the previous TAR study that investigated transfer effects to other social cognitive domains [15] and provides evidence that this training extends beyond the targeted domain and generalizes to improved ToM.

Although the linear mixed models analysis did not yield a significant time x group interaction effect for facial affect recognition (PFA performance), effect sizes still suggest treatment-related improvements and align with previous TAR studies. The paper that introduced the TAR program in a one-group pre-post design [13] found a within-group Cohen's d effect size of 1.02, which is smaller than the current (Cohen's d = 1.32). Sachs et al. [16] failed to find a significant time x group interaction effect for facial affect recognition, but main effects of time and group were significant. Their between-group post-training effect size was larger than their between-group baseline effect size (Cohen's d = 1.01 vs. 0.55). The corresponding effect sizes in our study were 1.32 and 0.61, respectively. Sachs et al. [16] further found a larger within-group effect size for the TAR group than for the TAU group (Cohen's d = 0.53 vs. 0.17). Our corresponding effect sizes were 1.32 and 0.45, respectively. The lack of significant time x group interaction effect in the current study is probably due to a combination of low statistical power and improvement in both groups. Further, the TAR group performed significantly better than the TAU group at T1, indicating non-successful randomization for PFA baseline performance level. Our results must therefore be interpreted with caution. At the same time it should be noted that the PFA test performance of our TAR group at T3 (mean = 92.2, SD = 6.8) was much better than the performance of Norwegian healthy controls (mean = 82.8, SD = 7.5) [42].

Taken together, it is therefore likely that TAR improved facial affect recognition, although the statistical test was non-significant.

We failed to see a significant time x group interaction effect in the linear mixed models for the last two social cognitive tests (EmoBio, RAD) (see Table 2). The same trend in data is, however, seen here as for the PFA test. Although our study provides no direct evidence, this may suggest that TAR could generalize to body affect recognition and social perception.

For nonsocial cognition, the picture is different. Although there was a significant improvement in MCCB performance over time, this did not differ between groups, evidenced in the effect size of Cohen's $d = 0.09$ at T3. We believe that test-retest effects can explain the improvement seen over time for nonsocial cognition in both groups. Our results thus indicate that the TAR program has a specific effect on social cognition.

Among the measures of functioning, the TAR group unexpectedly self-reported worse social functioning over time (significant time x group interaction effect). For social functional capacity (AIPSS), however, a trend in the data suggests that improvement has taken place for the TAR group. Interestingly, this is not seen for nonsocial functional capacity (UPSA-B), paralleling our findings for social versus nonsocial cognition. Taken together, this suggests that TAR may improve social abilities, but not nonsocial abilities. This is further supported by lack of significant effects and small effect sizes for the global measure of functioning. In summary, the training had few effects on functioning, with a possible exception for social functional capacity.

Both groups show an overall improvement in clinical symptom load over time. The clinical improvement was present for global symptom level as well as for positive and depressive symptoms. This is likely an effect of the clinical treatment as usual, which both groups received, and cannot be attributed to the TAR intervention. This adds to the literature that shows that social cognitive training does not ameliorate psychosis [8, 11]. Only for negative symptoms was the time x group interaction effect significant. The TAU group had higher negative symptom load than the TAR group at baseline, lower at T2 and the same at T3. This is the only symptom variable where the TAR group had a slight increase over time (however, the within-group effect size is small: Cohen's $d = 0.32$). In brief, the training did not have an effect on clinical symptoms.

In regard to subjective outcomes, clinical insight did not change over time, and the TAR group had slightly better insight than the TAU group at all assessment points. For self-esteem there was a significant main effect of group. Whereas the TAR group reported lower self-esteem than the TAU group at T1, this was not so at later assessment points. The within-group effect size suggests that the self-esteem of the TAR group grew over time, possibly related to the social cognitive training they received. This is in contrast to their social self-efficacy, which did not change over time. Also, the between-group effect size for social self-efficacy at T3 was medium-sized, but the TAU group had the highest score. This stands in contrast to what one would expect, namely that a training program focusing on understanding the emotional expressions of others would have positive effects on someone's confidence in their own social abilities. This may align with the findings for self-reported social functioning reported above. For both measures, the TAR group rated themselves lower than the TAU group and with quite consistent scores at the three assessment points. One explanation is that they may have a more realistic view of their current social functioning and their social skills in hypothesized situations (social self-efficacy). This is further supported by their insight score, which is numerically higher compared to the TAU group. Another explanation is that RSES assesses self-efficacy for interacting with others, not for understanding others, and therefore may not be suited to detect specific changes in social self-efficacy.

Limitations of the study include the lack of an active control condition and direct measures of real-life functioning, as well as non-successful randomization for baseline levels for facial affect recognition and self-esteem. Other limitations are the relatively small sample size and the inclusion of individuals with schizoaffective disorder in the schizophrenia group. These issues, along with the choice of refraining from correction for multiple comparisons in this exploratory study, necessitate replications in independent samples.

In conclusion, this study found that a targeted social cognitive training program improved social cognition with durable effects for a non-targeted domain. Lasting transfer effects also appeared at trend-level for social functional capacity. The study provides evidence for the generalizability and durability of targeted facial affect training, but also highlights that additional treatment efforts are required to achieve functional improvement.

Ethical standards

The study was approved by Norway's Regional Ethics Committee for Medical and Health Research (South East) and was performed in accordance with the ethical standards laid down in the Helsinki Declaration. All participants gave their informed consent before their inclusion in the study.

Conflict of interest

None of the authors report any conflict of interests.

References

1. Green MF, Llerena K, Kern RS (2015). The "right stuff" revisited: what have we learned about the determinants of daily functioning in schizophrenia? Schizophr Bull 41:781-785.

2. Savla GN, Vella L, Armstrong CC, Penn DL, Twamley EW (2013). Deficits in domains of social cognition in schizophrenia: a meta-analysis of the empirical evidence. Schizophr Bull 39:979-992.

3. Rocca P, Galderisi S, Rossi A, Bertolini A et al (2016). Social cognition in people with schizophrenia: a cluster-analytic approach. Psychol Med 46:2717-2729.

4. Hajdúk M, Harvey PD, Penn DL, Pinkham AE (2018). Social cognitive impairments in individuals with schizophrenia vary in severity. J Psychiatr Res 104:65-71.

5. Fett AK, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 35:573-588.

6. Schmidt SJ, Mueller DR, Roder V (2011). Social cognition as a mediator variable between neurocognition and functional outcome in schizophrenia: empirical review and new results by structural equation modeling. Schizoph Bull 37(Suppl.2):S41-S54.

7. Pinkham AE (2014).Social cognition in schizophrenia. J Cli Psychiatry 75(Suppl. 2):14-19.

8. Kurtz MM, Richardson CL (2012). Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. Schizophr Bull 38:1092-1104.

9. Fiszdon JM, Reddy LF (2012). Review of social cognitive treatments for psychosis. Clin Psychology Rev 32:724-740.

10. Kurtz MM, Gagen E, Rocha NBF, Machado S, Penn DL (2016). Comprehensive treatments for social cognitive deficits in schizophrenia: a critical review and effect-size analysis of controlled studies. Clin Psychology Rev 43:80-89.

11. Bordon N, O'Rourke S, Hutton P (2017). The feasibility and clinical benefits of improving facial affect recognition impairments in schizophrenia: systematic review and meta-analysis. Schizophr Res 188:3-12.

12. Vass E, Fekete Z, Simon V, Simon L (2018). Interventions for the treatment of theory of mind deficits in schizophrenia: systematic literature review. Psychiatry Res 267:37-47.

13. Frommann N, Streit M, Wölwer w (2003). Remediation of facial affect recognition impairments in patients with schizophrenia: a new training program. Psychiatry Res 117:281-284.

14. Wölwer W, Frommann N, Halfmann S, Piaszek, Streit M, Gaebel W (2005). Remediation of impairments in facial affect recognition in schizophrenia: efficacy and specificity of a new training program. Schizophr Res 80:295-303.

15. Wölwer W, Frommann N (2011). Social-cognitive remediation in schizophrenia: generalization of effects of the Training of Affect Recognition (TAR). Schizophr Bull 37(suppl.2):S63-S70.

16. Sachs G, Winklbauer B, Jagsch R, Lasser I, Kryspin-Exner I, Frommann N, Wölwer W (2012). Training of Affect Recognition (TAR) in schizophrenia-impact on functional outcome. Schizophr Res 138:262-267.

17. Horan WP, Green MF (2017). Treatments of social cognition in schizophrenia: current status and future directions. Schizophr Res doi: 10.1016/j.schres.2017.07.013.

18. American Psychiatric Association (2000). The diagnostic and statistical manual of mental disorders. 4th edition. American Psychiatric Publishing, Washington DC.

19. First MB, Spitzer RL, Gibbon M, Williams JBW (1996). Structured clinical interview for DSM-IV TR axis I disorders, research version, patient edition (SCID-P). Biometrics Research, New York State Psychiatric Institute, New York.

20. Wechsler D (2007). Wechsler Abbreviated Scale of Intelligence (WASI). Norwegian manual supplement. Pearson Assessment, Stockholm, Sweden.

21. Norwegian Directory of Health (2013). National guidelines for the assessment, treatment and continued support for individuals with psychotic disorders. Norwegian Directory of Health, Oslo, Norway.

22. Ekman P, Friesen W (1976). Pictures of Facial Affect. Consulting Psychologists Press, Palo Alto, CA.

23. Heberlein AS, Adolphs R, Tranel D, Damasio H (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. J Cogn Neurosc 16:1143-1158.

24. Vaskinn A, Sundet K, Østefjells T, Nymo K, Melle I, Ueland T (2016). Reading emotions from body movement: a generalized impairment in schizophrenia. Front Psyciatry: 6:2058.

25. Fretland RA, Andersson S, Sundet K, Andreassen OA, Melle I, Vaskinn A (2015). Theory of mind in schizophrenia: error types and associations with symptoms. Schizophr Res 162:42-46.

26. Dziobek I, Fleck S, Kalbe E, Rogers K, Hassenstab J, Brand M (2006). Introducing MASC: a movie for the assessment of social cognition. J Autism Develop Disord 36:623-636.

27. Vaskinn A, Fiske AP, Green MF (2017). Enhancing tolerability of a measure of social perception in schizophrenia: comparison of short and long Norwegian versions of the Relationships Across Domains test. Cogn Neuropsychiatry 22:254-262.

28. Sergi MJ, Fiske AP, Horan WP, Kern RS, Kee KS, Subotnik KL, Nuechterlein KH, Green MF (2009). Development of a measure of relationship perception in schizophrenia. Psychiatry Res 166:54-62.

29. Nuechterlein KH, Green MF (2009). In: Rund BR, Sundet K editors. MATRICS Consensus Cognitive Battery. Norwegian version. MATRICS Assessment, Los Angeles, CA.

30. Pedersen G, Hagtvedt KA, Karterud S (2007). Generalizability studies of the global assessment of functioning-split version. Compr Psychiatry 48: 88-94.

31. Hellvin T, Sundet K, Vaskinn A, Simonsen C, Ueland T, Andreassen OA, Melle I (2010). Validation of the Norwegian version of the Social Functioning Scale (SFS) for schizophrenia and bipolar disorder. Scand J Psychol 51: 525-533.

32. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S (1994). The Social Functioning Scale: the development and validation of a new measure of social adjustment for use in family intervention programs with schizophrenic patients. Br J Psychiatry 157:853-859.

33. Vaskinn A, Ventura J, Andreassen OA, Melle I, Sundet K (2015). A social path to functioning in schizophrenia: from social self-efficacy through negative symptoms to social functional capacity. Psychiatry Res 228:803-807.

34. Donohoe CP, Carter MJ, Bloem WD, Hirsch GL, Laasi N, Wallace CJ (1990). Assessment of interpersonal problem-solving skills. Psychiatry 53:329-339.

35. Mausbach BT, Harvey PD, Goldman SR, Jeste DV, Patterson TL (2007). Development of a brief scale of everyday functioning in persons with severe mental illness. Schizophr Bull 33:1364-1372.

36. Kay SR, Fiszbein A, Opler LA (1987). The positive and negative syndrome scale PANSS) for schizophrenia. Schizophr Bull 13:261-276.

37. Wallwork RS, Fortgang R, Hashimoto R, Weinberger DR, Dickinson D (2012). Searching for a consensus model of the Positive and Negative Syndrome Scale for schizophrenia. Schizophr Res 137:246-250.

38. Addington D, Addington J, Maticka-Tyndale E (1993). Assessing depression in schizophrenia: The Calgary Depression Scale. Br J Psychiatry 163(suppl22):39-44.

39. McDermott B (1995). Development of an instrument for assessing self-efficacy in schizophrenic spectrum disorders. J Clin Psychol 51:320-331.

40. Rosenberg M (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

41. Birchwood M, Smith J, Drury V, Healy J, Macmillan F, Slade M (1994). A self-report insight scale for psychosis: reliability, validity and sensitivity to change. Acta Psych Scand 89:62-67.

42. Frøyhaug M (2018). Theory of mind in severe mental illness. Exploring the psychometric properties of the Norwegian version of the Hinting Task. Master thesis. Department of Psychology, University of Oslo, Oslo, Norway.

Table 1. **Demographic and clinical characteristics in individuals with schizophrenia receiving training of affect recognition (TAR) or treatment as usual (TAU).**

	TAR n = 24 Mean (SD)	TAU n = 24 Mean (SD)	*statistics*	p-value
Age	29.9 (8.9)	30.8 (8.7)	t = 0.36	0.720
Sex (m/f)	18/6	14/10	$x^2 = 1.50$	0.221
Education (years)	12.3 (2.1)	12.0 (1.9)	t = 0.50	0.618
WASI IQ	104.0 (12.5)	99.9 (14.8)	t = 1.02	0.312
Illness duration (years)	7.7 (7.1)	8.0 (8.1)	t = 0.15	0.881
Antipsychotic medication*	1.89 (1.31)	1.34 (1.06)*	t = 1.53	0.134

*Amount of defined daily dose of antipsychotic treatment; n = 21 (three TAU participants did not use antipsychotic medication).

Table 2. **Comparison of social and non-social cognition, and functioning between the intervention and control groups at baseline, post-treatment and 3 months post-treatment**

	Baseline (T1)				Post-treatment (T2)				3 month follow-up (T3)				T1-T3
	TAR n = 24 Mean (SD)	TAU n = 24 Mean (SD)	t (p-value)	Between-group Cohen's d	TAR n = 20 Mean (SD)	TAU n = 20 Mean (SD)	t (p-value)	Between-group Cohen's d	TAR n = 17 Mean (SD)	TAU n = 15 Mean (SD)	t (p-value)	Between-group Cohen's d	Within-group Cohen's d
Social cognition													
PFA	79.7 (12.1)	71.2 (15.6)	2.11 (0.040)	**0.61**	89.7 (7.0)	75.2 (13.4)	4.19 (< 0.001)	**1.42**	92.2 (6.8)	78.0 (14.8)	3.38 (0.002)	**1.32**	TAR: **1.32** TAU: **0.45**
EmoBio	0.80 (0.11)	0.76 (0.16)	1.04 (0.304)	0.30	0.87 (0.08)	0.82 (0.13)	1.32 (0.196)	**0.48**	0.87 (0.11)	0.80 (0.20)	1.13 (0.269)	**0.47**	TAR: **0.64** TAU: 0.22
RAD	24.9 (5.3)	23.5 (4.6)	0.99 (0.326)	0.28	25.6 (5.9)	24.3 (5.7)	0.65 (0.518)	0.22	26.7 (5.4)	23.3 (6.4)	1.59 (0.123)	**0.58**	TAR: 0.34 TAU: 0.03
MASC	28.8 (6.7)	29.3 (8.2)	0.27 (0.787)	0.07	31.3 (7.3)	30.3 (7.9)	0.39 (0.702)	0.13	32.0 (7.6)	26.9 (6.9)	1.94 (0.063)	**0.71**	TAR: **0.45** TAU: 0.32
Non-social cognition													
MCCB	40.3 (8.2)	41.0 (7.8)	0.30 (0.763)	0.09	41.1 (7.7)	43.7 (8.7)	0.99 (0.328)	0.32	42.6 (7.0)	41.9 (9.3)	0.26 (0.800)	0.09	TAR: 0.30 TAU: 0.11
Self-reported functioning													
SFS	103.3 (8.6)	104.3 (8.3)	0.44 (0.662)	0.12	102.7 (6.7)	105.5 (8.5)	1.15 (0.258)	0.37	102.6 (5.4)	106.9 (9.6)	1.59 (0.123)	**0.57**	TAR: 0.10 TAU: 0.29
Clinician-rated functioning													
GAF-f	44.4 (9.6)	44.3 (11.9)	0.01 (0.989)	0.01	46.2 (10.4)	48.8 (15.9)	0.60 (0.554)	0.20	47.0 (10.1)	44.9 (12.4)	0.54 (0.595)	0.19	TAR: 0.26 TAU: 0.05
Functional capacity													
UPSA	75.2 (11.8)	75.6 (14.5)	0.12 (0.905)	0.03	74.4 (13.9)	76.0 (12.9)	0.37 (0.714)	0.12	75.3 (11.7)	77.2 (12.7)	0.66 (0.511)	0.16	TAR: 0.01 TAU: 0.12
AIPSSrec	68.9 (20.7)	66.0 (19.2)	0.50 (0.621)	0.14	74.4 (16.0)	69.5 (18.6)	0.83 (0.410)	0.28	78.9 (17.4)	67.1 (23.4)	1.57 (0.128)	**0.58**	TAR: **0.52** TAU: 0.05
AIPSSproc	47.3 (21.6)	50.1 (25.0)	0.43 (0.667)	0.12	57.8 (14.9)	53.7 (24.2)	0.60 (0.555)	0.21	57.0 (20.2)	56.0 (26.9)	0.12 (0.909)	0.04	TAR: **0.46** TAU: 0.23
AIPSSsend	46.8 (14.6)	45.9 (23.0)	0.16 (0.872)	0.05	49.7 (15.8)	52.1 (23.4)	0.34 (0.735)	0.12	55.9 (18.4)	48.9 (25.4)	0.86 (0.396)	0.32	TAR: **0.55** TAU: 0.12

TAR = Training of Affect recognition (intervention group). TAU = Treatment as usual (control group). PFA = Pictures of Facial Affect. T2 TAR = 19, TAU = 19. T3 TAR = 15. EmoBio = Emotion in Biological Motion. T2 TAR = 19, TAU = 19. T3 TAR = 15. RAD = Relationships Across Domains. T2 TAR = 19, TAU = 18. MASC = Movie for the Assessment of Social Cognition. T2 TAR = 18, TAU = 19. T3 TAR = 15. MCCB = Matrics Consensus Cognitive Battery. SFS = Social Functioning Scale. T2 TAU = 19. GAF-f = Global Assessment of Functioning, function subscale. T2 TAR = 19. UPSA = University of California San Diego Performance Skills Assessment. T2 TAR = 17, TAU = 19. T3 TAR = 15. AIPSS = Assessment of Interpersonal Problem-Solving Skills. rec = receiving skills. proc = processing skills. send = sending skills. T2 TAR = 17, TAU = 19. T3 TAR = 15.
Medium-sized effect sizes (≥0.45) in **bold**.

Table 3. **Comparison of clinician-rated symptoms and self-reported clinical insight, social self-efficacy and self-esteem between the intervention and control groups at baseline, post-treatment and 3 months post-treatment**

	Baseline (T1)				Post-treatment (T2)				3 month follow-up (T3)				T1-T3
	TAR n = 24 Mean (SD)	TAU n = 24 Mean (SD)	t (p-value)	Between-group Cohen's d	TAR n = 20 Mean (SD)	TAU n = 20 Mean (SD)	t (p-value)	Between-group Cohen's d	TAR n = 17 Mean (SD)	TAU n = 15 Mean (SD)	t (p-value)	Between-group Cohen's d	Within-group Cohen's d
Clinical symptoms													
PANSSpos	10.5 (4.4)	9.8 (3.9)	0.59 (0.560)	0.17	9.8 (4.1)	8.8 (5.1)	0.69 (0.498)	0.22	8.9 (3.5)	8.0 (4.9)	0.63 (0.533)	0.21	TAR: 0.41 TAU: 0.41
PANSSneg	12.9 (4.8)	16.0 (5.9)	1.98 (0.054)	**0.58**	14.8 (6.1)	12.9 (6.2)	0.99 (0.331)	0.31	14.8 (6.9)	14.7 (5.7)	0.07 (0.945)	0.02	TAR: 0.32 TAU: 0.22
PANSSdis	5.6 (2.4)	6.3 (3.0)	0.91 (0.369)	0.26	5.6 (3.1)	5.5 (2.5)	0.17 (0.866)	0.04	5.9 (3.5)	6.1 (2.7)	0.17 (0.864)	0.06	TAR: 0.10 TAU: 0.07
PANSSexc	5.3 (1.7)	5.7 (2.0)	0.78 (0.437)	0.22	5.2 (1.4)	5.8 (3.2)	0.69 (0.492)	0.26	4.9 (1.5)	4.9 (2.3)	0.02 (0.982)	0.00	TAR: 0.25 TAU: 0.37
PANSSdep	8.4 (2.7)	6.9 (3.0)	1.83 (0.074)	**0.53**	6.9 (3.3)	5.9 (2.5)	1.07 (0.291)	0.35	6.7 (3.3)	5.7 (2.6)	0.97 (0.340)	0.34	TAR: **0.57** TAU: 0.43
CDSS	4.3 (3.5)	3.6 (3.1)	0.74 (0.463)	0.21	3.3 (4.2)	3.0 (3.8)	0.28 (0.782)	0.08	3.0 (3.5)	1.8 (2.2)	0.13 (0.266)	0.42	TAR: 0.37 TAU: **0.68**
GAF-s	45.6 (10.5)	45.2 (11.8)	0.13 (0.897)	0.04	45.6 (12.5)	51.8 (18.2)	1.22 (0.228)	0.40	48.0 (15.1)	50.1 (13.3)	0.42 (0.676)	0.15	TAR: 0.19 TAU: 0.39
Self-report													
BCIS (range 0-12)	8.3 (2.1)	7.2 (1.9)	1.74 (0.089)	**0.55**	8.2 (2.1)	7.2 (1.6)	1.80 (0.081)	**0.54**	8.6 (2.6)	7.5 (1.4)	1.50 (0.144)	**0.55**	TAR: 0.13 TAU: 0.18
RSES	58.6 (16.4)	62.4 (16.3)	0.80 (0.426)	0.23	60.6 (15.7)	63.7 (20.3)	0.53 (0.598)	0.17	59.7 (15.3)	69.7 (18.3)	1.70 (0.100)	**0.60**	TAR: 0.07 TAU: 0.42
Rosenberg Self-esteem	23.5 (5.2)	26.9 (5.3)	-2.23 (0.031)	**0.65**	26.1 (6.3)	27.3 (3.8)	0.73 (0.473)	0.24	27.1 (6.9)	28.8 (4.6)	0.80 (0.431)	0.30	TAR: **0.60** TAU: 0.38

TAR = Training of Affect recognition (intervention group). TAU = Treatment as usual (control group). PANSS = Positive and Negative Syndrome Scale. pos = positive symptoms. neg = negative symptoms. dis = disorganized symptoms. exc = excited symptoms. dep = depressive symptoms. CDSS = Calgary Depression Scale for Schizophrenia. GAF-s = Global Assessment of Functioning, symptoms subscale. T2 TAR = 19. BCIS = Birchwood Clinical Insight Scale. T2 TAU = 19. RSES = Revised Self-efficacy Scale. Rosenberg Self-esteem. T2 TAU = 19.
Medium-sized effect sizes (≥0.45) in **bold**.

Figure 1. CONSORT Flow Diagram

```
                          ┌─────────────────────┐
                          │      Consented      │
                          │      (n = 60)       │
                          └─────────────────────┘
                                     │
                                     │              ┌─────────────────────┐
                                     ├─────────────▶│      Withdrew       │
                                     │              │     (n = 12)        │
                                     │              └─────────────────────┘
                          ┌─────────────────────┐
                          │ Baseline Assessment │
                          │ and Randomization   │
                          │      (n= 48)        │
                          └─────────────────────┘
                                     │
              ┌──────────────────────┴──────────────────────┐
              ▼                                              ▼
   ┌──────────────────────┐                      ┌──────────────────────┐
   │ Treatment (TAR) group│                      │ Control (TAU) group  │
   │       n = 24         │                      │       n = 24         │
   └──────────────────────┘                      └──────────────────────┘
              │                                              │
┌──────────┐  │                                              │  ┌──────────┐
│4 drop-out│◀─┤                                              ├─▶│4 drop-outs│
└──────────┘  ▼                                              ▼  └──────────┘
   ┌──────────────────────┐                      ┌──────────────────────┐
   │ Follow-up post-      │                      │ Follow-up (8 weeks)  │
   │ treatment (8 weeks)  │                      │       n = 20         │
   │       n = 20         │                      │                      │
   └──────────────────────┘                      └──────────────────────┘
              │                                              │
┌──────────┐  │                                              │  ┌──────────┐
│3 drop-outs│◀─┤                                              ├─▶│5 drop-outs│
└──────────┘  ▼                                              ▼  └──────────┘
   ┌──────────────────────┐                      ┌──────────────────────┐
   │ Follow-up (3 months) │                      │ Follow-up (3 months) │
   │       n = 17         │                      │       n = 15         │
   └──────────────────────┘                      └──────────────────────┘
```
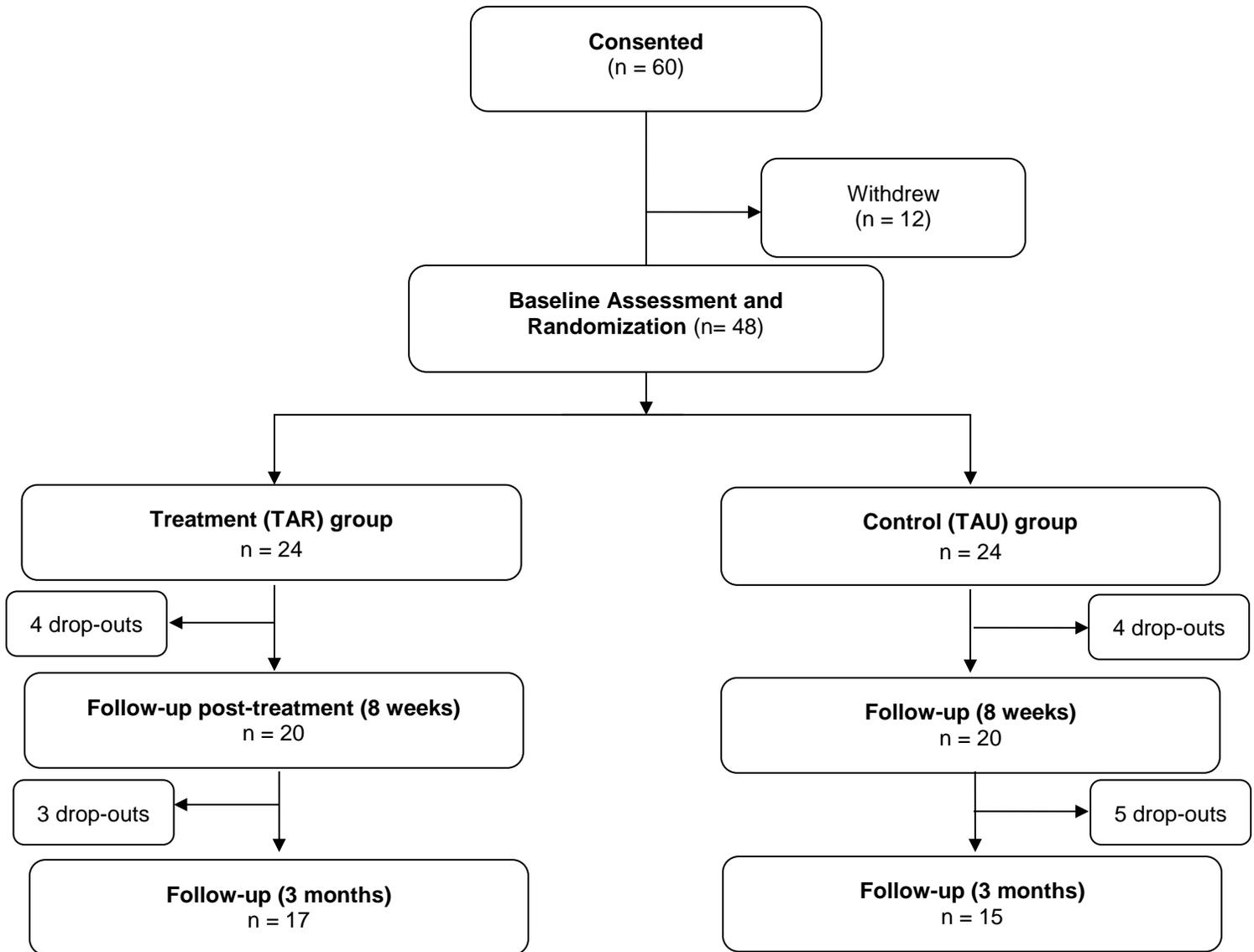
Figure 2. Theory of mind (Movie for the Assessment of Social Cognition) in individuals receiving social cognitive training (TAR) or treatment as usual (TAR). Estimated marginal means from linear mixed models
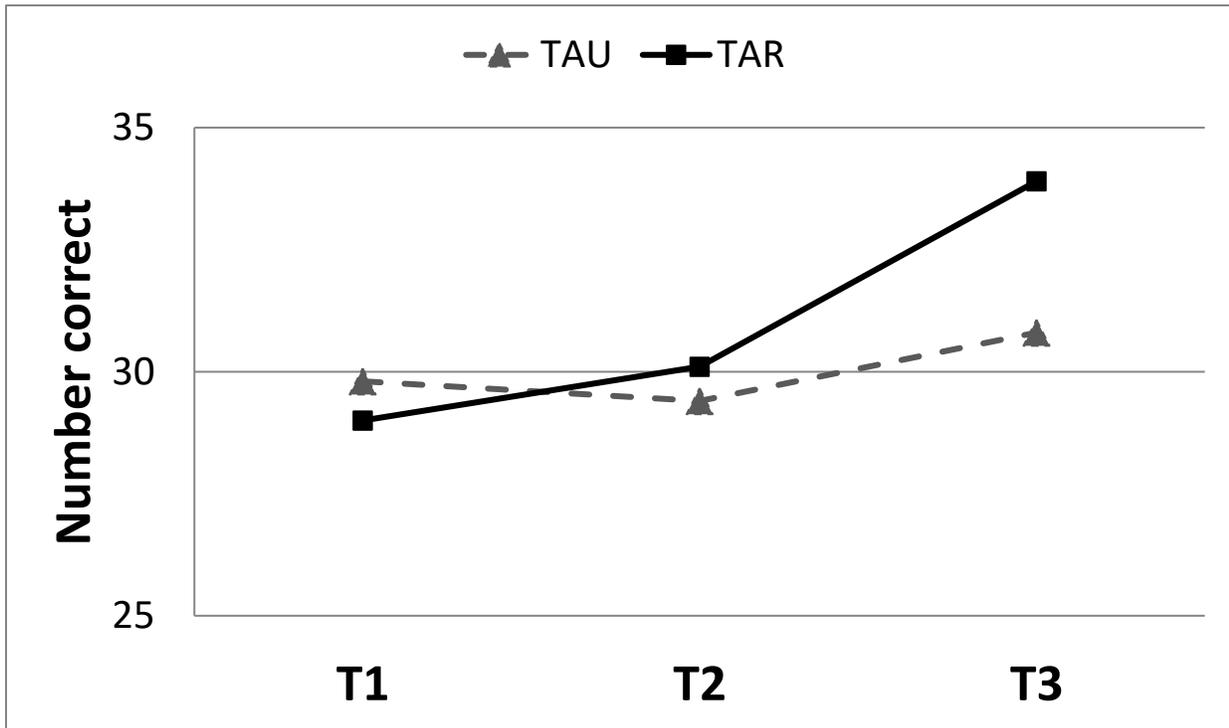

Figure 3. Facial affect recognition (Pictures of Facial Affect) in individuals receiving social cognitive training (TAR) or treatment as usual (TAR). Estimated marginal means from linear mixed models