The Relational Trip Task, a novel ecological measure of relational memory: data from a schizophrenia sample

Ana Elisa Sousa¹², Gabrielle Pochiet¹², Jennifer D. Ryan³, Martin Lepage¹⁴

¹Douglas Mental Health University Institute, 6875 LaSalle Boulevard, Montreal, QC, H4H 1R3, Canada

²Integrated Program in Neuroscience, McGill University, 2001 McGill College Ave, Montréal, QC, H3A 1G1, Canada

³Rotman Research Institute, Baycrest; Department of Psychology, University of Toronto, 27 King’s College Cir, Toronto, ON M5S, Canada.

⁴Department of Psychiatry. McGill University, 1033 Pine Avenue, Montreal, QC, H3A 1A1.

Corresponding Author:

Martin Lepage
Douglas Mental Health University Institute
6875 LaSalle Boulevard
Verdun, Quebec H4H 1R3
Canada
Tel: 514-761-6131 ext. 4393
Fax: 514-888-4064
E-mail: martin.lepage@mcgill.ca


This is an Accepted Manuscript of an article published by Taylor & Francis in Cognitive Neuropsychiatry on 11 Oct 2021, available online: https://doi.org/10.1080/13546805.2021.1987870
The Relational Trip Task, a novel ecological measure of relational memory: data from a schizophrenia sample

Abstract

Introduction: Relational memory (RM) is severely impaired in schizophrenia. Unitization can circumvent RM impairments in clinical populations as measured by the transverse-patterning (TP) task, a well-established measure of RM capacity. We compared memory performance on a new ecological RM measure, the Relational Trip Task (RTT), to that of TP at baseline and examined the effects of a unitization intervention in RTT performance. RTT involves learning relational information of real-life stimuli, such as the relationship between people and places or objects.

Methods: TP and RTT performances were examined in 45 individuals with schizophrenia. TP-impaired participants (n = 22) were randomized to either the intervention or an active control group. TP and RTT were administered again after unitization training. Task validity and reliability were assessed. Intervention group’s pre- and post-RTT accuracies were compared and contrasted to that in the control group.

Results: RTT and TP were moderately correlated. TP non-learners had inferior performance in RTT at baseline. Improvement in RTT performance after unitization training was observed in the intervention group; no pre-post improvement was observed in the control group.

Conclusion: RTT has an acceptable criterion validity and excellent alternate-form reliability. Unitization seemed to be successfully generalized to support associations of real-life stimuli.

Keywords: schizophrenia; psychosis; relational memory; unitization; generalization.
Introduction

Relational memory (RM), a subcomponent of episodic memory that accounts for the associations between individual items, or between items and their context (Öngür et al., 2006), is severely affected in schizophrenia (Armstrong et al., 2012; Avery et al., 2019; Avery et al., 2018; Hannula et al., 2010; Lepage et al., 2015; Öngür et al., 2006; Ranganath et al., 2008; Suh et al., 2020; Titone et al., 2004). RM is critical for complex cognitive processes, including imagination, language, future planning and flexible social cognition and behaviour (Addis et al., 2011; N. J. Cohen, 2015; Duff & Brown-Schmidt, 2012; Moscovitch, 2008; Rubin et al., 2014).

Unitization represents a cognitive strategy that can compensate for RM impairments (D'Angelo et al., 2017; D’Angelo et al., 2015; D’Angelo et al., 2016; Ryan et al., 2013). In unitization, distinct items are fused through an action that retains the nature of their relationship, which is later recovered without the need for storage of the relations themselves (D’Angelo et al., 2016). Some amnesic individuals with lesions relatively circumscribed to the extended hippocampal system (D’Angelo et al., 2015; Ryan et al., 2013) and healthy older adults have been able to benefit from unitization performing a Transverse Patterning (TP) task (D’Angelo et al., 2016). It is hypothesized that unitization supports TP learning by reducing reliance on hippocampal-dependent relational binding and increasing reliance on fused representations supported by semantic memory processes (D'Angelo et al., 2017; D’Angelo et al., 2015; D’Angelo et al., 2016; Ryan et al., 2013).

TP requires learning of previously unknown relations between items. Individuals with schizophrenia present deficits in this task (L. Rowland et al., 2007; L. M. Rowland et al., 2010;
Spieker et al., 2013) even when relational learning is facilitated with extensive, stepwise TP training (Spieker et al., 2013). We recently reported evidence that unitization is also effective in individuals with schizophrenia that previously failed TP (Sousa et al., 2021). It remains unclear, however, whether unitization can support the learning of relationships among closer to real-life stimuli, such as people, familiar places and objects.

To that end, we developed the Relational Trip Task (RTT), an ecological task partly inspired by the Transverse Patterning paradigm (D'Angelo et al., 2016; Hanlon et al., 2005) but using real-life stimuli (pictures of people’s faces, and real and familiar places and objects) instead of the geometrical shapes, symbols or abstract objects used in TP. The “winner/loser” relationship used in TP was replaced by those commonly found between people and places (going to/coming from) and between people and objects (like/dislike). The task was conceived so that participants could still apply the unitization strategy to learn the relationships.

We adapted the unitization training protocol used by D'Angelo et al. (2016) and conducted a pilot study using an active-controlled experimental design with the overall objective of examining the feasibility, acceptability and efficacy of a unitization intervention in schizophrenia. The results of this study were reported in Sousa et al. (2021). In the present study, we examined the criterion validity of RTT as a measure of relational memory in relation to an existing similar measure of relational memory, the TP task. We also examined the reliability of two versions of RTT. Secondly, we compared the use of the unitization strategy in the RTT task to that of the TP task. We hypothesized that 1) The TP and RTT tasks will be highly, positively correlated, and indicative of good criterion validity; furthermore, individuals that fail to learn all TP rules (an indicative of RM-impairment) will perform significantly worse in the RTT compared to those who succeed in learning TP, which is evidence of good construct
validity; thirdly, we expect the two versions of RTT to be strongly, positively correlated, as an indication of high task reliability. 2) RM-impaired individuals will show performance improvement in RTT after unitization training, beyond that found in the control group, in similarity to the improvement observed in TP after unitization training.

Materials and methods

Participants
Recruitment took place at the Center for Personalized Psychological Intervention for Psychosis (Ci3P) of the Douglas Institute in Montreal between August 2018 and September 2019. Inclusion criteria for screening comprised of a) having a diagnosis of schizophrenia spectrum or related psychotic disorder, b) age 18 and over, c) clinical stability, as determined by the research protocol physician, d) ability to provide informed consent and e) French or English proficiency. Exclusion criteria were Verbal IQ score <70, diagnosis of a medical or neurological condition known to affect cognition, family history of hereditary neurological disorders (e.g. Huntington’s disease), and a diagnosis of substance dependence (current or within the past three months). Informed written consent was collected, and participants received compensation for their time. Research protocols were approved by the Douglas Institute’s Research Ethics Committee. A screening procedure was designed to identify participants with a selective relational binding impairment (in contrast to a more pervasive/general cognitive impairment). Screening procedures were detailed in Sousa et al. (2021) and summarized below.

Assessment
Participants’ diagnosis, clinical data and medication dosages were obtained from a semi-structured interview. When possible, that information was confirmed by medical chart review. Forty-five participants recruited for the unitization study completed TP and RTT at baseline. A detailed description of the TP task can be found in Sousa et al. (2021). Participants were classified at screening according to their TP performance. A score equal to or below 67% in this task (indicating that only 2/3 rules are being employed at testing) suggests a lack of relational learning, and the use of a “winner-takes-all” strategy instead (for more details, see Sousa et al. (2020). Individuals with performance ≤ 67% in the TP task at screening were classified as TP non-learners, while those scoring > 67% as TP learners. Subsequently, a subset of non-learners identified for their selective RM impairment (defined by failing relational-binding TP, but scoring above 80% in a semantically-rich version of the TP task) was pseudo-randomly assigned to the unitization intervention group or the control group via sequential assignment. One participant was later excluded due to a change in diagnosis that no longer met the psychosis criterion.

The intervention group and the control group received three similar sessions that included transverse-patterning conditions and different versions of RTT before and after unitization training, and separate symptoms and cognitive functioning assessment session as part of the unitization intervention. The control group was not exposed to unitization training and unitized task conditions prior to session 3. All sessions were performed in either English or French according to participants’ preferences. More details about the unitization intervention can be found in Sousa et al. (2021).

Structure and content of the Relational Trip Task
The Relational Trip Task (RTT) was programmed on E-prime 2.0 and presented on a laptop computer. Twenty stimuli (5 faces, 5 objects and 5 places) plus 5 practice stimuli were used in each of the three versions of RTT. Each version comprised its own set of stimuli (see Figure 1.A for an example of task stimuli). All versions were composed of two parts, each comprising two training blocks, an immediate test block and a delayed test block after a one-hour break. The two task parts differed in whether the presented pairs were composed of people’s faces and pictures of places or people’s faces and pictures of objects.

Initial task instructions required participants to imagine they were going on a trip around the world and that they would meet some people on their way. For each person, they would learn where they were coming from and going to. They were asked to do the best they could to remember the relationships (i.e. “coming from” or “going to”) between people and places, followed by a practice example to favour familiarization with the task.

The first block proceeded by presenting each person, followed by the picture of the place the person was “coming from” and the place that the person was “going to”. Participants were asked to answer about both relationships three times each for each person, in random order, before proceeding to the next person's stimulus. Answers were given by pressing the keys “Q” or “P” on the laptop keyboard, to indicate the relationship on the left (“coming from”) or right (“going to”), which were placed accordingly on the screen. Feedback was provided, so if participants answered correctly, they saw a cartoon face and the words “Good job!” on the screen, and if not, the words “Not quite! Let’s try again…” and an upset cartoon face were shown instead (see Figure 1.B). An accuracy of 83% (5 out of 6 correct) was required to proceed to the next person’s pairs.
Each place stimulus was paired with two different people, each of the pairings having one directionality (so that one person was “coming from”, and a different person was “going to” each of the five places). The repetition of place stimuli was used to encourage participants to encode the directionality of the relationship rather than solely the pairing itself, simulating the transverse-patterning paradigm in which the nature of the relationship between two items depends on the pair, and not on the value of the isolated items.

After learning the relationships between people and places (10 pairs in total), participants proceeded to the second training block, in which they were asked the directionality (“coming from”/”going to”) of the ten previously learned relationships in random order, with feedback, for a total of 30 trials (3 presentations of each pair). If the accuracy of 67% (20 out of 30 correct) was not achieved, the block is repeated, up to five times.

The immediate testing block followed the structure of the second training block (30 trials, 3 presentations of each pair, random order), this time without feedback, which was replaced by a fixation cross. After a one-hour break, participants completed the delayed test, which was identical to the immediate testing block.

The second part of the task followed training and immediate testing of part one. Participants were told they would learn about the preferences of the people they just met (i.e. liked and disliked objects). All blocks in the second part of the task had the same structure as those in the first part, including an immediate test and a one-hour delayed test.

**Unitization training procedures**

Unitization training included learning the core concepts of relational memory and unitization, followed by a step-by-step explanation of how the strategy works (Supplemental file A).
Participants practiced unitization with two transverse-patterning conditions (one in the control group) before using it on RTT (see Sousa et al. (2021) for details). All participants completed a standard version of RTT at screening (RTT1). The intervention group completed a second, unitized version (RTT2-Unitized) after learning unitization. The control group completed a non-unitized version (RTT2) and a unitized version of RTT (RTT3-Unitized) before and after being exposed to unitization training. The unitization strategy steps were revised before participants were asked to perform RTT using unitization, which as per design, happened in session 2 for the intervention group and session 3 for the control group.

RTT-Unitized conditions differed from standard RTT to the extent to which instructions encouraged participants to use unitization to remember the relationships between people and places, and between people and objects during the first training block. For that, the encoding instruction (“do your best to remember”), presented below each new pair of person-place and person-object, was replaced by “Let’s use unitization to remember this”. All remaining task features were identical to that of standard RTT.

During the first part of RTT-Unitized, participants were given an example of unitization during the practice trial (see supplemental file B) and assisted in the creation of their unitizations for each pair. Assistance was offered as needed, to ensure participants were following the four steps of unitization learned during training. During the second part of RTT-Unitized, participants were asked to self-generate their unitization without the experimenters’ assistance. Unitization content was collected by the experiments using the unitization track form after each training and testing block, by asking participants to describe their unitization narrative for each pair. Procedures for RTT2-Unitized and RTT3-Unitized were identical.
The pictures of places were selected from a database of images collected from Google and normalized for visual complexity and familiarity, based on data obtained from 40 healthy individuals. Pictures of familiar objects were obtained from the BOSS database (Brodeur et al., 2010). All the pictures of places and objects were balanced for familiarity and visual complexity across the three versions of RTT. Pictures of people's faces were obtained from the Chicago Faces Database (Ma et al., 2015) and controlled for age, attractiveness, unusually and race across the three versions of RTT. A high level of familiarity with places and objects aimed to facilitate the generation of unitizations.

**Task validity measures**

Testing accuracies at immediate and one-hour delayed testing were the outcomes of TP performance. RTT immediate and one-hour delayed testing accuracies were used to compare task performances. Delayed testing accuracies were considered as the main outcome of TP and RTT, as participants could be relying on working memory (rather than solely on relational memory) to support immediate accuracies. This approach is in line with previous work in TP and unitization (D’Angelo et al., 2016).

Selective RM-impaired participants also completed the Positive and Negative Syndrome Scale 6 (PANSS-6, Østergaard et al., 2016)). The Wechsler Abbreviated Scale of Intelligence (WASI; (Wechsler, 1999)) and CogState Schizophrenia Battery (Pietrzak et al., 2009) in a separate session following the screening session, as part of the unitization intervention protocol. Statistics for these measures can be found in supplemental file C.

**Statistical analysis**
Data analysis was performed using IBM SPSS 21 for macOS and 26 for Windows. Effect sizes were calculated with G*Power version 3.1.9.4 (Faul et al., 2007) and interpreted according to J. Cohen (1988). Group profiles were examined by quantitative and qualitative analysis. The appropriate statistical tests were used to compare means or medians of the target variables across groups, with a 95% interval.

**RTT criterion validity** was assessed by the correlation between RTT and TP tasks accuracies (the last a previously validated measure of relational memory in schizophrenia). As the distribution of TP and RTT task accuracies were non-parametric (Shapiro-Wilk tests $p < .05$), Spearman’s correlation tests were used to compare TP and RTT task performances (one-tailed, as we hypothesized a positive relationship between tasks). As learners and non-learners’ TP and RTT task accuracies were not normally distributed (Shapiro-Wilk tests $p < .05$), Mann-Whitney tests were used to compare between group’s TP and RTT performances at baseline (two-tailed). Effect sizes $r$ for Mann-Whitney tests were calculated using the formula $r = Z/\sqrt{N}$ (as proposed by Rosenthal et al. (1994). Effect sizes $r$ between .1 and .3 are considered small; between .3 to .5 intermediate, and .5 or larger are considered strong (J. Cohen, 1988).

**RTT reliability** between the two different versions of standard RTT (RTT1 and RTT2) was assessed using Spearman’s correlations between task accuracies for those completing both versions of the task (RM-specifically impaired participants randomized to the control group). This test was one-tailed, as we assumed tasks would be positively correlated.

**RTT performance pre- and post-unitization training:** As the distribution of the RTT task scores for both intervention and control group participants approached normality, a 2-way Factorial ANOVA was used to compare participants’ one-hour delayed overall performance, with a within-subjects factor of task (RTT1, RTT2) and a between-subject factor of group
(intervention, control). The same test was repeated with the immediate accuracies (see supplemental file D for immediate test results).

Similar analyses (i.e. same statistical tests used for comparison between group’s TP and RTT at baseline, and equivalent non-parametric tests for RTT performance pre-and post-unitization training) were performed posthoc using reaction time (RT) as the dependent variable.

Results

Forty-five participants completed TP and RTT1 at screening. Table 1 displays baseline sociodemographic and clinical data for TP-learners (N = 21) and non-learners (N = 24), along with group comparisons and its corresponding test statistic when a significant difference was observed. TP-learners had significantly more years of education (MD = 1.4, SED = 0.64) compared to non-learners. No differences in sex, age, education, age of illness onset, illness duration, antipsychotic medication and diagnosis (affective or non-affective psychosis) were identified (ps > .05).

Table 1 near here

Out of the 24 TP non-learners, 22 participants with a selective RM impairment were randomized to the intervention (n = 10) or control group (n = 12) as part of the unitization intervention protocol. The intervention group was significantly more symptomatic, and more cognitively impaired, and had more females compared to control groups (ps < .05, see Supplemental file C for statistics per group). A total of 18 participants completed RTT2/RTT2-Unitized and 10 control participants subsequently completed RTT3-Unitized, after receiving unitization training. One participant in the intervention group that fail to follow procedures
during sessions and explicitly reported not using unitization during training was excluded from further analyses.

**RTT criterion and construct validity**

TP and RTT performances at baseline for Learners and Non-learners are shown in Figure 2.A. Spearman’s correlation analysis showed that TP and RTT accuracies were significantly and positively correlated at immediate ($r_s = .55, p < .001$) and delayed testing ($r_s = .57, p < .001$). Scatter plots of TP and RTT accuracies at immediate and delayed tests are shown in Figure 2.B.

As expected, TP-learners performed significantly better in the TP task than non-learners (immediate testing $U = 2.5, z = -5.73, p < .001, r = .86$; delayed testing $U = 0, z = -5.80, p < .001, r = -.86$). As hypothesized, TP Learners also performed better than non-learners in RTT (immediate testing $U = 102.5, z = -3.40, p < .001, r = -.51$; delayed testing $U = 79, z = -3.942, p < .001, r = -.59$).

**TP and RTT reaction times at baseline**

TP and RTT overall reaction times did not significantly differ at baseline, neither at immediate ($n = 21, \text{Mdn}_{TP} = 2353.8, \text{Mdn}_{RTT} = 2080.3, T = -1.0, p > .05, r = -0.2$) or at delayed testing ($n = 44, \text{Mdn}_{TP} = 2171.1, \text{Mdn}_{RTT} = 2421.4, T = -1.0, p > .05, r = -0.2$). TP and RTT reaction times for Learners and Non-learners, at immediate and delayed test, are shown in Figure 3.A (mean percentages) and B (scatter plots).
TP immediate reaction time data was not available for most non-Learners that completed this condition (23/24) due to the task interruption rule. Therefore, TP within-groups and between-groups comparisons were only performed at delayed testing for this group.

Wilcoxon Signed-Rank tests for the within-group comparison (Learners vs non-Learners) performances across tasks (TP and RTT) suggested that Learners’ reaction times did not significantly differ when comparing at immediate testing ($T = -1.344, p > .05$) or at delayed testing ($T = -1.095, p > .05$). Non-Learners reaction time did not differ when comparing TP and RTT at delayed testing ($T = -.456, p > .05$).

Mann-Whitney tests comparing reaction time of Learners and Non-Learners in each task yield no significant differences ($ps > .05$), suggesting that reaction times in either TP or RTT tasks were not associated with RM-impairment. See supplemental file E for test statistics.

**RTT reliability**

Ten participants with schizophrenia completed the two versions of standard RTT used in the study. Spearman’s correlation tests suggested that task accuracies were strongly correlated at immediate testing, $r_s = .78, p < .01$, and very strongly correlated at delayed testing, $r_s = .91, p < .001$, indicating high reliability between the different versions of the task. Scatter plots of RTT1 and RTT2 accuracies at immediate and delayed tests are available in Figure 4.A.

**RTT performance pre- and post-unitization training**

Selective RM-impaired participants’ (TP Non-Learners) overall performance in each RTT version at immediate and delayed test are shown in Figure 4.B. The interaction between group and task condition at 1-hour delayed testing was statistically significant, $F(1, 16) = 4.48, p =$
.05, ηp² = .219, suggesting that group performances were differentially affected by condition. There was no statistically significant difference in RTT1 performance between intervention (M = 0.77, SE = 0.05) and control group (M = 0.72, SE = 0.06) (F(1,16) = .43, p > .05, ηp² = .026), a task in which both groups were not using unitization. RTT2 performance differences between groups approached significance, F(1,16) = 3.91, p = .065, ηp² = .197, with accuracy in the intervention group (M = 0.90, SE = 0.03) marginally superior to that in the control group (M = 0.74, SE = 0.07), when intervention group used unitization. Simple main effects of task were non-significant in the control group, (MRTT1 = 0.72, SE = 0.06; MRTT2 = 0.74, SE = 0.07, F(1,9) = .74, p > .05, ηp² = .076); in the intervention group, performance was significantly superior in RTT2 (M = 0.90, SE = 0.03) compared to RTT1 (M = 0.77, SE = .05, F(1, 7) = 7.50, p < .05, ηp² = .517). Results yield a similar pattern at immediate testing, except that RTT1 and RTT2 immediate accuracies in the intervention group were both significantly superior to that of the control group (ηp² = .053 and .253, respectively).

RTT2 and RTT3’s 1-hour delayed accuracies comparison in the control group showed that this group also achieved superior performance when they received unitization training, compared to when they performed the task without using unitization (MD = 0.17, SED = 0.05, t(9) = -.328, p = .01, one-tailed, d = 1.05).

Post-hoc comparison of RTT-Unitized accuracies (intervention and control group combined) (Mdn = 0.89) with that of the TP-learners’ RTT accuracies at baseline (Mdn = 0.95) was performed using a Mann-Whitney test (two-tailed). No significant differences were found (U = 165.5, z = -.667, p > .05), suggesting that when unitizing in RTT, RM-impaired participants’ accuracy was not different from that of TP-learners at screening.
**RTT Reaction Times pre- and post-unitization training**

Selective RM-impaired participants’ (TP Non-Learners) reaction time by group in each RTT version at immediate and delayed test are available in the supplemental files F (bar charts) and G (scatter plots). No significant RT differences were observed when comparing pre-and post-unitization training as a function of group, suggesting no overall differences in reaction time in the different versions of RTT (RTT1 vs RTT2) and no significant reaction time differences between intervention and control groups (please see supplemental file H for test statistics). Post-hoc comparison of RTT-Unitized reaction times (intervention and control group combined) ($Mdn = 2895.4$) with that of the TP-learners’ RTT reaction times at baseline ($Mdn = 2402.9$) was performed using a Mann-Whitney U test (two-tailed). A significant difference was found ($U = 111.0, z = -1.98, p = .05, r = 0.32$), suggesting that when unitizing in RTT, RM-impaired participants’ reaction time was significantly superior to that of TP-learners at screening.

**Use of the unitization strategy in RTT**

Overall, participants used unitization during RTT-Unitized conditions, which was verified by the unitization track form administered during these conditions. An example of participant’s unitizations is available in Supplemental File I. Participants’ unitization strategies greatly varied in complexity and to what extent they included additional elements beyond the unitization instructions to enrich their narratives. Overall, it was observed that participants more often followed the steps of unitization when unitizations were assisted by experimenters (first half of the task), compared to when participants unitized by themselves. When self-unitizing with people-objects pairs, unitizations often included a rationalization to explain the person’s ‘liking’ or ‘disliking’ of the object, adding to, or replacing the action-consequence
required by step 2 of unitization instructions (see Supplemental file A for unitization steps). These rationalizations often included elements of prior unitizations with that same person, developed during the first part of the task.

**Discussion**

This study aimed to assess the criterion validity of the Relational Trip Task, a relational-binding task in which both stimuli and task instructions were framed into a real-life-like context, by comparing it to an often-used measure of relational memory in clinical populations, the transverse-patterning paradigm. We also investigated whether unitization, a strategy capable of improving transverse-patterning performance in individuals with relational memory impairment (D’Angelo et al., 2015; D’Angelo et al., 2016), could be translated to the Relational Trip Task. Our results demonstrated that RTT performance was moderately and positively associated with TP, suggesting an acceptable convergent criterion validity for the task (Glen, 2015). Additionally, participants who failed to learn TP at screening had significantly inferior performance in RTT which provides further evidence of construct validity. The comparison between the two versions of RTT suggested a high positive correlation between performance in each version, and indicative of excellent alternate form reliability. We observed evidence of the successful use of unitization to improve performance in RTT for a subset of individuals with schizophrenia who failed the TP task and had inferior RTT performance at baseline, in similarity to what has been observed in TP. Given the importance of reaction time as an indicator of several factors that could affect performance (e.g. task complexity, fatigue, attention, acute symptoms at the time of testing and practice effect), analyses were replicated using this variable as a dependent measure. No significant differences were detected between TP and RTT overall reactions times, between groups (Learners and non-learners) at baseline
for either task, or between non-learners when performing the task a second time using unitization, but reaction times were significantly superior for those using unitization compared to those succeeding at the RTT task at baseline (learners), with a moderate effect size. Superior reaction times using unitization could be associated with the active recall of unitization narratives at testing, rather than using familiarity or an alternative strategy. Due to the small sample size of groups and group unbalance after randomization, important differences in RT may have remained undetected in between-group comparisons.

To our knowledge, the Relational Trip Task is the first relational memory task developed with verisimilitude (the degree of agreement between the task cognitive demands and everyday demands, as defined by Franzen and Wilhelm (1996)) in mind. Previous evidence suggests that overall, memory tests of verisimilitude are better predictors of everyday memory function when compared to standard neuropsychological tests of memory (Chaytor & Schmitter-Edgecombe, 2003). Further evidence suggests that performance-based naturalistic tasks simulating real-life circumstances (in similarity to RTT) are at least moderately associated with real-world skills in people with serious mental illness (e.g. Test of Grocery Shopping Skills, Faith and Rempfer (2018)).

Other tasks used to detected deficits in relational encoding in schizophrenia are often detached from a real-life context (e.g. The Relational and Item-Specific Encoding task (RISE) (Ragland et al., 2011), list learning (Iddon et al., 1998), transitive inference (Titone et al., 2004). For instance, though RISE uses real-life pictures of objects, these items were paired at random, and do not necessarily reflect associations observed in real life. When episodic memory improvement is measured by such tasks as the present one, it remains unknown whether the results can be generalized to reflect functional changes in individuals’ daily lives. Although
showing pictures of individuals and places, and presenting information regarding these (i.e. peoples’ preferences, peoples’ whereabouts) are not equivalent to social interactions, those are closer approximations to a real-life context than that found in previous tasks, representing a closer step towards ecological representation. Other studies investigating short-and long-term memory encoding have used dynamic naturalistic material (e.g. movie scenes) as an approximation of the complex, contextually rich content present in real-life situations (e.g. (Kwok & Macaluso, 2015a, 2015b; Santangelo & Bordier, 2019)

Performance differences between groups when the intervention group was unitizing were significant at immediate testing, but only marginally significant at 1-hour delayed testing. The imbalance of cognitive performance and symptomatology between groups that disfavoured the intervention group may have masked differences in performance at delayed testing. However, the presence of pre-post training improvement in the intervention group (combined with its absence in the control group in corresponded conditions) points to the benefits of unitization even for individuals whose cognitive function is severely compromised. Further studies designed to investigate the efficacy of unitization in association with a wide range of cognitive profiles could provide further information on the matter.

We have reported previous evidence suggesting that unitization may facilitate learning arbitrary TP relationships in individuals with schizophrenia who have relational-binding deficits (Sousa et al., 2021). The suggested mechanism involves reduced reliance on hippocampal-dependent relational binding and increasing reliance on fused representations (Ranganath et al., 2008). In our study, unitization training improved participants’ performance, with a large effect size (J. Cohen, 1988) when the strategy was externally provided (e.g. using
animations depicting the objects interacting so the “winner object” in each pair was made clear, then using stills representing each unitized-pair during training).

We have also investigated the potential of self-unitization in TP, as a first step to the ecological application of the unitization paradigm in daily-life situations requiring relational learning. TP performance significantly decreased when participants had to generate their own strategies, suggesting that both learning task rules and generating strategies may have resulted in a high cognitive load preventing some individuals from successfully developing the unitization strategy and optimal learning. This observation is in line with conclusions of previous works, which pointed to online maintenance of information as one of the processes supporting unitization, along with imagery and access to semantic memory (D’Angelo et al., 2015; D’Angelo et al., 2016; Ryan et al., 2013). In the present study, we separated the self-generation component from the task learning phase, by presenting relationships beforehand and allowing participants to develop a strategy before the task learning phase.

As for the use of the unitization strategy in RTT, unitizations assisted by the experimenter seemed more complete than those generated by participants without assistance. Participants tended to follow unitization "steps" more often when unitizing it with people-places pairs, compared to people-objects pairs. One limitation of this study is that, due to our task design, the condition of the people-places pair was always the one for which unitization was assisted, and the condition of the people-objects pairs was always the one generated by participants without assistance. Consequently, it is hard to parcel out whether following the unitization instructions for people-objects was more challenging than with people-places pairs, or whether this was simply because the experimenter was not requiring all steps to be followed in this part of the task, allowing participants to deviate from or adapt the instructions freely. We suggest
the latter, as self-generated unitizations of people-objects pairs often extrapolate unitization instructions by providing explanations to people’s liked and disliked objects (e.g. Eve dislikes candy cane. She don’t like sugar [because] she’s diabetic) or attributed a personality characteristic to a person based on the object and relationship directionality (e.g. Kate dislikes BBQ [because] she’s vegetarian). It is worth noticing that those unitizations, although often deviating from the unitization instructions, were still unitizing with semantic information, as they provided a frame of reference for a given person from which participants could draw on to answer about the nature of the relationships (e.g. “Kate is going to Cairo [on a] investigation”/ Kate likes the computer. She’s back from the investigation, writing the report [on the computer]).

The possibility that participants’ unitizations in RTT was supported, at least in part, by autobiographical memory should also be considered, as task instructions encouraged one’s to place themselves in a hypothetical social context (i.e. “You’re going on a trip around the world and on your way you will meet some people that are also travelling around.”). Consequently, participants’ unitization often reflected personal experiences or denoted self-insertion. The autobiographical memory literature has indicated that although individuals with schizophrenia seem to not benefit from the self-referential memory boost (an increase in encoding when information has personal implication), they do benefit from imagining a situation from a personal perspective (for a recent review of autobiographical and episodic memory deficits in schizophrenia, please see Kwok et al. (2020)).

Unitization’s effects in other levels of recollection than relational memory have been investigated in recent years. According to Liu and Guo (2019, 2020), unitization of picture pairs could not only facilitate pairs-recognition by relying on familiarity but also supported
item recognition by reducing the contribution of recollection, which is thought to require more neural processing than familiarity recollection (Liu & Guo, 2019). A subsequent study (Liu & Guo, 2020) confirmed unitization’s potential to facilitate item recognition and suggested that this strategy can, in turn, suppress verbatim recognition (the ability to differentiate old stimuli from lure stimuli) in non-clinical individuals. Healthy individuals unitizing with word pairs, exhibited mixed results on the association of unitization and item-recollection, with an overall conclusion that unitization did not interfere with item-recognition performance (Liu et al., 2020).

It remains unclear how unitization affects item recognition in schizophrenia. Additionally, whether unitization differentially benefits individuals with schizophrenia when using verbal, rather than visual stimuli is worth consideration, taking into account that verbal memory is one of the most severely impaired cognitive domains in schizophrenia (Green et al., 2004).

As for the neural processes supporting unitization strategy and memory performance in healthy individuals (Tu & Diana, 2020) and mild-cognitively impaired (MCI) older adults (Delhaye et al., 2019), there has been evidence for the involvement of the perirhinal cortex (PrC) during retrieval of unitized information. PrC function is associated with familiarity-based recollection (Ranganath et al., 2008), and an increase in bilateral perirhinal cortex and decrease in parahippocampal cortex function has been observed during retrieval of unitized colour information in healthy individuals (Tu & Diana, 2020). Meanwhile, MCI older adults, for which perirhinal cortex function is affected (Delhaye et al., 2019), have been reported not to benefit from unitization (D’Angelo et al., 2016; Delhaye et al., 2019). Delhaye et al. (2019) has recently reported on an association between PrC integrity and performance for unitized
associations in this population. The underlying neural mechanisms supporting unitization in schizophrenia, including its reliance on PrC integrity, have yet to be examined.

Other limitations of the present study were the small sample size of participants that completed RTT-unitized conditions, which limits the power of our statistical analyses and robustness of conclusions, and the small sample size of participants completing both versions of the RTT in the standard condition (i.e. not using unitization), which limits the power of our conclusion regarding the reliability between task versions.

We suggest that replications of this study include increased task difficulty by using two different sets of people for each part of the task, and by using less familiar objects for the people-place pairs. Another future step is the validation of the RTT as a measure of relational memory performance, including its comparison to similar relational-memory measures validated in the schizophrenia population (e.g. RISE). In the revised RTT, we propose incorporating comparative measures of episodic memory, such as item recall and transitive inference. Finally, we suggest a larger randomized controlled trial assessing unitization effectiveness in the RTT with a larger sample with true randomization to replicate results, and to investigate the associations between cognitive performance and unitization use in schizophrenia.
Acknowledgments

We thank the members of our lab, Karyne Anselmo, Marie-Christine Boulianne, Charlie Henri-Bellemare and Alina Sabirova, for their help with recruitment and conducting the evaluations, and to all the people who participated in the study.

Disclosure of Interest

M.L. reports grants from Otsuka Lundbeck Alliance, personal fees from Otsuka Canada, personal fees from Lundbeck Canada, grants and personal fees from Janssen, and personal fees from MedAvante-Prophase, outside the submitted work. A.E.S. and G.P. declare no conflict of interest.

Funding details

This study was supported by the Fonds de la recherche en santé du Québec (# 32855) and by the Natural Sciences and Engineering Research Council of Canada (#2015-04913). M. L. holds a James McGill professorship from McGill University and a Research Chair from the Fonds de Recherche Québec –Santé. J. D. R. holds funding from the Canadian Institutes of Health Research (PJT 162274). A. E. S. was partially supported by a Globalink Graduate Fellowship.
References


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2879675/pdf/sbq037.pdf


Table 1. Means (SD), range and percentages for baseline sociodemographic and clinical data for TP learners and non-learners with schizophrenia.

<table>
<thead>
<tr>
<th></th>
<th>TP learners (n = 21)</th>
<th>TP non-learners (n = 24)</th>
<th>Group comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females % (N)</td>
<td>28.6% (6)</td>
<td>54.2% (13)</td>
<td>ns ^3</td>
</tr>
<tr>
<td>Age (years)</td>
<td>36.6 (12.4)</td>
<td>42.2 (11.5)</td>
<td>ns ^2</td>
</tr>
<tr>
<td>Education (years of)</td>
<td>13.2 (2.5)</td>
<td>11.8 (1.8)</td>
<td></td>
</tr>
<tr>
<td>Age of onset (years)</td>
<td>25.9 (8.8)</td>
<td>26.5 (9.8)</td>
<td></td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>11.1 (9.3)</td>
<td>15.8 (11.0)</td>
<td>ns ^2</td>
</tr>
<tr>
<td>Antipsychotic medication (Chlorpromazine equivalents)</td>
<td>455.2 (509.5)</td>
<td>484.4 (406.3)</td>
<td>ns ^2</td>
</tr>
<tr>
<td>Diagnosis % (N)</td>
<td>19% (4)</td>
<td>12.5% (3)</td>
<td>ns ^3</td>
</tr>
<tr>
<td>Affective Psychosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Affective Psychosis</td>
<td>81% (17)</td>
<td>87.5% (21)</td>
<td></td>
</tr>
</tbody>
</table>

Groups’ statistics were compared using independent t-tests ^1, Mann-Whitney tests ^2 or chi-square tests ^3 (two-tailed) when appropriate. Significant differences are reported at *p* < .05.

ns = groups do not significantly differ at the .05 level.
Figure 1.

A. Stimuli Example
B. Experimental Procedure

a) Part 1

i. Stimuli Presentation and Training block 1 (2 pairs at a time, 6 trials before proceeding to next person-pairs, total 10 pairs, minimum of 60 trials)

ii. Training block 2 (10 pairs presented in random order, 3 times each (minimum of 30 trials)

iii. Testing block ((10 pairs presented in random order, 3 times each (total of 30 trials)
Figure 2.

A.

![Immediate Test](image1)

![Delayed Test](image2)

B.

![Immediate Test](image3)

![Delayed Test](image4)
Figure 3.

A.

B.
Figure 4.

A.

IMMEDIATE TEST

DELAYED TEST

B.

INTERVENTION

CONTROL

Task accuracy

RTT1

RTT2 (Unitized)

RTT1

RTT2

RTT3 (Unitized)
Figure Captions:

Figure 1. (A) Example of stimuli used in the RTT task versions 1, 2 and 3 (not including practice stimuli). (B) Experimental Procedure of the Relational Trip Task. The RTT task was split into two parts, differing in whether it paired (a) people with places or (b) people with objects. Each task part was composed of three blocks. (i) Person stimuli were initially present alone on screen, then paired with the other two stimuli (two places or two objects depending on task part), one pair at a time. Each pairing represented a particular relationship between person and item (“going to”/” coming from” for places, “likes”/” dislikes” for objects). Each person’s pairs presentation followed training, in which participants were asked about the relationships just learned before proceeding to the next person. (ii) Training block 2 repeated all previously learned pairings in random order, asking for a pair’s specific relationship. During training blocks, responses were self-paced, and feedback was provided. Unitized training was identical to standard training except that the instruction “Let’s use unitization to remember this” appeared below the pair during its first presentation. ii) Testing blocks were identical to training block 2, except that feedback was replaced by a fixation cross. All test blocks, regardless of the condition, and regardless of whether the test was immediate, after an hour delay, or at one-month follow-up followed the same procedure and the same stimulus arrangement.

Figure 2. TP learners and non-learners immediate and 1-hour delayed accuracy percentage (A) and scatterplots (B) in the Transverse-patterning and Relational Trip Task at baseline. Error bars represent the standard error of the mean. Learners’ and non-learners’ accuracies significantly differ in both tasks at immediate and delayed testing, ***ps < .001. TP learners n = 21; TP non-learners n = 24. Effect sizes r for Mann-Whitney tests ≥ .5 are considered strong (J. Cohen, 1988).

Figure 3. TP learners and non-learners immediate and 1-hour delayed reaction times percentage (A) and scatterplots (B) in the Transverse-patterning and Relational Trip Task at baseline. Error bars represent the standard error of the mean. Learners’ and non-learners’ reaction times did not significantly differ when comparing tasks in either immediate or delayed testing, ps > .05. Immediate test Ns: Learners TP = 20, RTT = 21; non-learners TP = 1, RTT = 22. Delayed test Ns: Learners TP = 21, RTT = 21; non-learners TP = 24, rtt = 23.

Figure 4. RTT scatter plots of accuracy (A) for all Non-learners in RTT1 and RTT2 and accuracy percentages (B) for the intervention (n = 9) and control (n =10) groups at immediate and delayed testing. Error bars represent the standard error of the mean. Horizontal pattern in the bars represents a unitized task condition. Intervention group performances significantly increased after unitization training (ps < .05). The same effect was present for the control group at delayed testing after learning unitization (p < .05).
Supplemental Files

The Relational Trip Task, a novel ecological measure of relational memory: data from a schizophrenia sample

Ana Elisa Sousa, Gabrielle Pochiet, Jennifer D. Ryan, Martin Lepage

A. A step-by-step explanation of the unitization training. This document was provided as a handout to participants when they were asked to generate unitization strategies.

Unitization has four steps:

1. Recognize what is familiar: relate the objects to what you know or to what they remind you of.

2. Picture an action-consequence: imagine the two objects interacting together so that their relationship is made clear.

3. Imagine motion: represent the nature of the interaction in the unit with movement.

4. Fusion: fuse the objects together: combine them in a final single unit that you can easily remember.
B. Example of Unitization provided at practice during RTT2-U condition

1. Practice pairs

Part 1 – Unitization with people and places:

a) This person is coming from this place. To make my unitization, I’ll first make my two elements familiar: I’ll make the person familiar by naming him Tom (because he looks like my friend Tom) and I’ll call this place “Toronto” because it looks a lot like Toronto to me. Next, I imagine the action-consequence of Tom ‘coming from’ Toronto: I imagine him jumping from the top of the CN tower, landing on the water and swimming towards me. “Jumping from the CN tower and swimming towards me” is also my ‘movement’ (the third step of the unitization). My final picture (unit) is Tom in the water, swimming towards me. Note that if I did not know (or was not sure of) the name of the city, I could have named it using some distinctive element from the picture as a reference (for example, I could have named it “the Purple Tower city”). The important is making it familiar enough so I’ll distinguish it from the other places that I see. Finally, every time I see this pair, I know that “Tom” is ‘coming from’ Toronto because I’ll remember my final picture, which is him swimming towards me after jumping from the CN Tower to the lake.

b) Now, Tom is going to that place. I believe it is Montreal, so I’ll imagine Tom swimming down the Saint Laurent River and climbing the Biosphere. My final picture is Tom on the top of the Biosphere waiving at me. Now every time I see this pair, I know Tom is “going to” Montreal because I remember he will swim down the river, climb the Biosphere and waive to me from its top.
c) This person likes this. To make my unitization, I’ll use the same name I have used before, “Tom”, because he looks like my friend Tom. The object is already familiar to me, I know it is an onion. Next, I imagine the action-consequence of Tom ‘liking’ the onion: I’ll imagine him making a big onion sandwich. Making an onion sandwich is also my action. My final picture is a very satisfied Tom biting his onion sandwich. Now when I see the pair “Tom” and “onion” I’ll think of him biting his onion sandwich, and I’ll know he likes it.

d) Now, Tom dislikes this, which I know’s a cupcake. I’ll imagine Tom throwing the cupcake on the floor and stepping on it. My final picture is Tom smashing the cupcake with his foot, icing all spread on the floor. Now every time I see this pair, I’ll think of Tom smashing the cupcake and I know he dislikes it.
### C. Means (SD), range and percentages for baseline sociodemographic and clinical data for eligible participants by group after randomization

<table>
<thead>
<tr>
<th></th>
<th>Control group (n = 10)</th>
<th>Intervention group (n = 12)</th>
<th>Group Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females % (N)</strong></td>
<td>80% (8)</td>
<td>25% (3)</td>
<td>$X^2 (1) = 6.6, p &lt; .05$</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>45.2 (11.6)</td>
<td>41.5 (11.5)</td>
<td>ns $^1$</td>
</tr>
<tr>
<td></td>
<td>24-57</td>
<td>27-59</td>
<td></td>
</tr>
<tr>
<td><strong>Education (years of)</strong></td>
<td>12.0 (1.5)</td>
<td>12.0 (2.1)</td>
<td>ns $^1$</td>
</tr>
<tr>
<td></td>
<td>9-14</td>
<td>8-16</td>
<td></td>
</tr>
<tr>
<td><strong>Age of onset</strong></td>
<td>31.2 (13.2)</td>
<td>22.9 (4.1)</td>
<td>ns $^2$</td>
</tr>
<tr>
<td></td>
<td>16-56</td>
<td>18-32</td>
<td></td>
</tr>
<tr>
<td><strong>Duration of illness (years)</strong></td>
<td>14.0 (10.5)</td>
<td>18.8 (12.9)</td>
<td>ns $^1$</td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>2-42</td>
<td></td>
</tr>
<tr>
<td><strong>Antipsychotic medication</strong></td>
<td>569.0 (478.5)</td>
<td>448.0 (360.7)</td>
<td>ns $^1$</td>
</tr>
<tr>
<td>(Chlorpromazine equivalents)</td>
<td>0-1450</td>
<td>100-1142</td>
<td></td>
</tr>
<tr>
<td><strong>Diagnosis % (N)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Psychosis</td>
<td>30% (3)</td>
<td>0% (0)</td>
<td>ns $^3$</td>
</tr>
<tr>
<td>Non-Affective Psychosis</td>
<td>70% (7)</td>
<td>100% (12)</td>
<td></td>
</tr>
<tr>
<td><strong>WASI T scores$^a$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>42.1 (9.4)</td>
<td>39.0 (8.7)</td>
<td>ns$^1$</td>
</tr>
<tr>
<td></td>
<td>26-52</td>
<td>25-54</td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>42.3 (9.0)</td>
<td>52.2 (6.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31-55</td>
<td>44-60</td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td>47.4 (7.0)</td>
<td>47.8 (9.04)</td>
<td>ns$^2$</td>
</tr>
<tr>
<td></td>
<td>37-53</td>
<td>34-64</td>
<td></td>
</tr>
<tr>
<td>Matrix</td>
<td>48.0 (13.7)</td>
<td>48.9 (7.0)</td>
<td>ns$^2$</td>
</tr>
<tr>
<td></td>
<td>20-63</td>
<td>31-56</td>
<td></td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>92.0 (11.4)</td>
<td>90.3 (12.5)</td>
<td>ns$^1$</td>
</tr>
<tr>
<td></td>
<td>74-104</td>
<td>74-110</td>
<td></td>
</tr>
<tr>
<td>Performance IQ</td>
<td>91.8 (15.3)</td>
<td>100.9 (9.2)</td>
<td>ns$^1$</td>
</tr>
<tr>
<td></td>
<td>66-110</td>
<td>81-111</td>
<td></td>
</tr>
<tr>
<td>Full 2 IQ</td>
<td>97.3 (13.8)</td>
<td>90.6 (9.6)</td>
<td>ns$^1$</td>
</tr>
<tr>
<td></td>
<td>68-112</td>
<td>74-105</td>
<td></td>
</tr>
<tr>
<td>Full 4 IQ</td>
<td>91.1 (13.9)</td>
<td>94.9 (11.4)</td>
<td>ns$^1$</td>
</tr>
<tr>
<td></td>
<td>68-107</td>
<td>76-113</td>
<td></td>
</tr>
<tr>
<td><strong>PANSS 6$^b$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>3.3 (1.6)</td>
<td>4.6 (1.0)</td>
<td>$t(18) = -2.2, p &lt; .05, d = 1.0$</td>
</tr>
<tr>
<td></td>
<td>1-6</td>
<td>3-6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>Intervention</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---</td>
</tr>
<tr>
<td>Conceptual Disorganiz.</td>
<td>2.4 (1.5)</td>
<td>2.0 (1.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td>1-5</td>
<td></td>
</tr>
<tr>
<td>Hallucinations</td>
<td>1.9 (1.7)</td>
<td>3.8 (1.8)</td>
<td>U = 19.5, z = - 2.4, p &lt; .05,</td>
</tr>
<tr>
<td></td>
<td>1-6</td>
<td>1-6</td>
<td>d = 1.1</td>
</tr>
<tr>
<td>Blunted Affect</td>
<td>2.6 (2.2)</td>
<td>2.2 (1.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-6</td>
<td>1-4</td>
<td></td>
</tr>
<tr>
<td>Social Withdrawal</td>
<td>3.2 (1.9)</td>
<td>2.0 (1.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td>1-5</td>
<td></td>
</tr>
<tr>
<td>Lack of Spontaneity</td>
<td>1.8 (1.7)</td>
<td>1 (0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>PANSS Total score</td>
<td>15.2 (5.6)</td>
<td>15.6 (4.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6-22</td>
<td>9-22</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Control and intervention participants statistics were compared using independent t-tests\(^1\), Mann-Whitney tests\(^2\) or chi-square tests\(^3\) (two-tailed) when appropriate. Significant differences (Sig.) are reported at \(p < .05\) (*). Identical to that reported in Sousa, 2020.

WASI, the Wechsler Abbreviated Scale of Intelligence; PANSS-6, Positive and Negative Syndrome Scale 6; CogState, CogState Schizophrenia Battery.

\(^a\)WASI vocabulary and Matrix T scores were missing for 2 participants (dropouts). The remaining WASI subtests scores were missing for 4 participants (2 dropouts, 1 intervention group and 1 control group).

\(^b\)PANNS 6 scores were missing for 2 participants that did not complete the evaluation session, both of which dropped-out before completing the study.

Ns = groups do not significantly differ at the .05 level.
D. Two-way Factorial ANOVA comparing immediate accuracies between groups.

Within-subjects factor of task (RTT1, RTT2) and between-subjects factor of group (intervention, control)

The results of the Two-Way Factorial ANOVA showed that the interaction between group and task condition at immediate testing was statistically significant, $F(1, 16) = 5.85, p < .05, \eta^2_p = .268$, suggesting that group performances were differentially affected by condition. In RTT1, performance was statistically superior in the intervention group ($M = 0.82, SE = 0.05$) compared to the control group ($M = 0.76, SE = 0.05$) ($F(1,16) = 0.90, p = .05, \eta^2_p = .053$). RTT2 performance was also superior in the intervention group when using unitization ($M = .93, SE = 0.04$) compared to that in the control group ($M = .74, SE = 0.07$; $F(1,16) = 5.41, p < .05, \eta^2_p = .253$).

Simple main effects of task were non-significant in the control group, $M_{RTT1} = 0.76, SE = 0.05$; $M_{RTT2} = .74, SE = 0.07$, $F(1,9) = .12, p > .05, \eta^2_p = 013$; in the intervention group, performance was significantly superior in RTT2 ($M = 0.93, SE = 0.04$) compared to RTT1 ($M = 0.82, SE = 0.05$, $F(1, 7) = 10.19, p < .05, \eta^2_p = .593$).
E. Comparison between Learners and Non-Learners reaction times by task
(Mann-Whitney tests)

<table>
<thead>
<tr>
<th></th>
<th>Learners N/Mdn</th>
<th>Non-Learners N/Mdn</th>
<th>Mann-Whitney U</th>
<th>Z</th>
<th>Exact sig (2-tailed)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate</td>
<td>1/(N/A)</td>
<td>1971.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Delayed</td>
<td>21/ 1882.8</td>
<td>24/ 2196.1</td>
<td>212.0</td>
<td>-.910</td>
<td>.372</td>
<td>-.12</td>
</tr>
<tr>
<td><strong>RTT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate</td>
<td>21/ 2219.6</td>
<td>22/ 2097.8</td>
<td>218.0</td>
<td>-.316</td>
<td>.764</td>
<td>-.05</td>
</tr>
<tr>
<td>Delayed</td>
<td>21/ 2402.9</td>
<td>23/ 2613.0</td>
<td>201.0</td>
<td>-.952</td>
<td>.351</td>
<td>-.14</td>
</tr>
</tbody>
</table>
Error bars represent the standard error of the mean. Horizontal pattern in the bars represents a unitized task condition. No significant differences in reaction time as a function of group was observed ($p > 0.5$). However, non-learners using unitization in RTT (intervention and control group combined) had superior reaction time when compared to learners completing the RTT task at baseline ($p = .05$), not shown in the figure.)
G. Scatter plots of RTT1 and RTT2 **reaction times** at immediate and delayed tests, intervention and control groups combined.

A) Immediate testing

![Immediate testing scatter plot]

B) Delayed testing

![Delayed testing scatter plot]
H. Test statistics for RTT Reaction Times pre- and post-unitization training

The interaction between group and task condition at 1-hour delayed testing was not statistically significant, $F(1, 16) = 0.25$, $p = .63$, $\eta^2 = .015$, suggesting that group reaction times were not differentially affected by conditions. Simple main effects of task, $F(1,16) = 0.74$, $p = .41$, $\eta^2 = .044$, and group, $F(1, 16) = .34$, $p = .57$, $\eta^2 = .02$, were also not significant, suggesting respectively no overall differences in reaction time in the different versions of RTT (RTT1 vs RTT2) and no significant reaction time differences between intervention and control groups.

Results yield a similar pattern at immediate testing, with no significant interaction or main effects of task and group when comparing reaction times before and after unitization intervention ($ps > .05$).

Control group’s reaction times also did not significantly differ before and after receiving unitization training (RTT2 vs RTT3-Unitized), either at immediate ($MD = -484.11$, $SED = 296.02$, $t(7) = -1.635$, $p = .146$, two tailed) or delayed testing ($MD = -322.55$, $SED = 481.79$, $t(7) = - .669$, $p = .525$, two tailed).
I. Example of participant’s unitizations

*Note: Information between brackets were added for clarity.*

**People-place pairs**

1. **Joe is coming from Paris.** Joe is standing at the top [of the Eiffel tower] and ready to jump. Very sad.
2. **Joe is going to Beach [Sidney].** But he didn't jump. Visits his father living in the white thing. Joe is sitting with his father, looking at the sunset, happy.
3. **Tracy is coming from New York.** Working at the empire state as maintenance. Mopped the floor, tired, bored.
4. **Tracy is going to Washington.** She's going to Washington to protest or strike or for a better job. Outside [of the white house] on the grass with a sign, yelling.
5. **Eve is coming from Cairo.** She's a musician, hip-hop artist. Straight in the middle [of the street] doing a hip-hop show.
6. **Eve is going to Paris.** She's on a tour. She's going on vacation. She's around, [has] no show, needs a break.
7. **Timothy is coming from Beach [Sidney].** Beach bum, surfing all day. Timothy tanning in the sun.
8. **Timothy is going to New York.** Going to take a tour of the big building. Having lunch in the middle of the building.
9. **Kate is coming from Washington.** She's walking on the ground. Kate walking, stressed, on the steps.
10. **Kate is going to Cairo.** She saw the one-dollar bill and said, "I'm going to the Cairo investigation". She's at the hotel waiting for the investigation.

**People-object pairs**

11. **Joe dislikes pen.** He doesn't like writing or reading. Only watch TV. Don't know what to do, just watches TV.
12. **Joe likes candy cane.** He loves eating sweets. Loves junk food and sugar against depression. Watches TV, eating candies.
13. **Tracy dislikes computers.** [She has] no patience for computers. Walks in her home. Sees kids on the computer. Says to kids "I don't understand".
14. **Tracy likes showerhead.** She to shower. She's coming home, taking shower. Taking shower and feeling relieved.
15. **Eve dislikes candy cane.** She doesn't like sugar. She's diabetic. Throws candy away to the garbage.
16. **Eve likes BBQ.** [She is a] strong woman so needs lots of meat and protein. Eating a big burger on BBQ.
17. **Kate dislikes BBQ.** She's vegetarian. Don't eat meat. She's tiny. With family, looking strange at them, eating her salad.
18. **Kate likes the computer.** She's back from the investigation, writing the report. Sitting on the computer writing the report.
19. **Timothy dislikes showerhead.** [He] doesn't take showers. No shower at his place.
20. **Timothy** likes pens. He is still on the beach doing the crossword, writing in his journal.