Objective, search whether a specific deficit in Theory of Mind (ToM) can be found in fronto-temporal dementia (FTD) but not in Alzheimer Type dementia (AD). The search was further supported with brain neuro-images analysis. Method. Neuropsychological tests, executive function tests, Yesavage depression scale and ToM tests using Happe’s strange stories tests (SST) were administered, to highlight any differences between Alzheimer and Frontotemporal dementia patients in their capacity to build inferences of other subject’s mental states and interpreting a non literal or metaphorical sentence. With this purpose 20 patients with AD and 20 with FTD, age, sex, literacy and IQ matched were selected for this study. Analysis: data were analyzed with ANOVA and correlation tests Results. Patients with FTD performed worst than AD in SST but had identical outcomes in physical events control task. These results didn’t show a meaningful correlation with EF tests. Discussion. Both groups offered low contextual content answers, but the percentage was greater with FTD. Those results are correlated with hypometabolism in ventromedial and dorsolateral brain areas, demonstrated with SPECT. Conclusions. FTD patients show a specific deficit in Theory of Mind which renders them incapable of building inferences on mental states.

Keywords. Executive Function, Theory of Mind, Frontotemporal Dementia, Alzheimer Disease.

Citation. Serrani D. Frontotemporal dementia deficit in Theory of Mind assessed with Happé’s Test.

RESUMO

Objetivo. O objetivo foi avaliar déficit específico na Teoria da Mente comparando-se pacientes com demência frontotemporal e doença de Alzheimer. Método. administraram-se testes neuropsicológicos, de função executiva, escala de depressão de Yesavage e teste de Teoria da Mente utilizando o Teste de Histórias Estranhas de Happé para avaliar capacidade de inferir os estados mentais e interpretar afirmações não literais (metafóricas). Foram avaliados 20 pacientes com demência Alzheimer e 20 pacientes com demência Frontotemporal, pareados por idade, sexo, escolaridade e IQ. Os dados foram analisados com teste ANOVA e correlação. Resultados. os pacientes com demência frontotemporal tiveram pior desempenho que os pacientes com doença de Alzheimer no teste de Histórias Estranhas, mas tiveram um desempenho similar na tarefa de controle de eventos físicos. Estes resultados não mostraram uma correlação significante com os testes de função executiva. Ambos os grupos tiveram poucas respostas corretas de conteúdo contextual, mas a porcentagem foi menor (pior desempenho) para demência frontotemporal. Estes resultados tiveram uma correlação positiva com maior hipometabolismo cerebral nas áreas ventromedial e dorsolateral, demonstrada em imagens de SPECT. Conclusões. os pacientes com Demência têm um déficit específico em Teoria da Mente o que os volta incapazes de construir inferências sobre os estados mentais.

Unitermos. Função Executiva, Teoria da Mente, Demência Frontotemporal, Doença de Alzheimer.

Citação. Serrani D. Frontotemporal dementia Déficit em Teoria da Mente avaliada com Teste de Happé.
INTRODUCTION

The interpretation of non literal, or metaphorical, language such as the subtleties in irony, jokes or sayings remains a difficult issue in dementia. This has been noticed in other types of illnesses as well, like autism, but not in normal child development from the 3-4 years, pointing to a deficit in the development of the Theory of Mind. This deficit in non literal sentence comprehension is highlighted in the Strange Story Test. Theory of Mind is a concept that implies the ability of inferring other people mental states and reading its thoughts and emotions. This ability became deranged in fronto-temporal dementia type. It is already known that sound changes in personality and social functioning reflects impairment in orbito-frontal area, including dis-inhibition, lack of empathy and disruption of manners and behavior rules that are socially accepted. It also involves pragmatic difficulties and errors in guessing non verbal cues. Between the many types of dementia, this pattern of changes is markedly noticeable in frontal lobe degenerative dementia type, previously known as Pick dementia, fronto-temporal dementia, or lobar fronto-temporal dementia; and is different from the parietal and hippocampal impairment which distinguishes Alzheimer type dementia, leading to a lose of semantic comprehension and an aphasic type syndrome. Usually this type of dementia begins at 40 to 60 years, with subtle changes in personality and behavior, including lack of social attunement, empathy with other people, dis-inhibition or inappropriate social behavior, disrespect for personal aspect and lack of introspection there is a tendency toward developing ritualistic and stereotyped behaviors that seem those of autism. In the initial stages it can be a hard duty to distinguish between the different types of dementia. But despite impairment in interpersonal and social behavior, this patients can exhibit a normal performance in central executive functions reflecting impairment in orbito-frontal areas, as executive tasks are sensitive to dorso-lateral functioning. ToM includes a) ability to infer other people mental and emotional states and social cognition, which is independent of general reasoning capacity, b) develops from childhood, c) is dissociated from other cognitive development areas. Regarding neuro-images, ToM comprises damage to frontal lobules and limbic system, particularly amygdala nuclei.

METHOD

Subjects

20 patients with AD and other 20 with FTD age, sex, literacy and IQ matched (Table 1). Diagnosis of probable FTD was achieved following NINCS-ADRDA criteria. Diagnosis of FTD followed criteria settled in Lund and Manchester Consensus. Inclusion criteria were: age >65 years, >7 years of education, lack of pyramidal, extra-pyramidal, neuro-muscular dysfunction, or systemic illness capable of diminishing intellectual performance (hypothyroidism, diabetes), drugs or alcohol abuse history, cranial trauma with loss of consciousness or depression. Depressive illness was ruled out using Ye-savage depression scale. Reference Number of local Ethics Commitee: 000-1379435-1.

Table 1

<table>
<thead>
<tr>
<th>Demographic data</th>
<th>AD</th>
<th>FTD</th>
<th>F (p&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>69.7±2.3</td>
<td>56.3±3.4</td>
<td>2.45 (0.59)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>11±1.5</td>
<td>10±1.8</td>
<td>2.48 (0.56)</td>
</tr>
<tr>
<td>Sex m/f</td>
<td>8/12</td>
<td>9/11</td>
<td></td>
</tr>
</tbody>
</table>

AD=Alzheimer dementia. FTD=Fronto-temporal dementia. Scores are expressed as Median and Standard Deviation. No meaningful differences were found between both groups. Scores are expressed as median and Standard Deviation.

Neuro-psychological Exam

The following general tests were administered: MMSE, Addenbrooke Cognitive Exam (ACE), Intelligence Scale Coefficient and Verbal and Performance Coefficient (WISC) and forward and backward digit span test and logical memory sub-test (story recall) from the Wechsler-Memory Scale Revised, in which two subjects listen to two short stories and then are required to provide the recall of each one, immediately and after a delay of 30 minutes. Each passage contains 25 informative tips and the score is the average of both stories. Complex Rey Figure immediate and 45 minutes delay copies, where subjects are required to copy with no time restrictions the figure provided. After a period of 45 minutes, and without previous warning, subjects are required to draw the
same image, on memory grounds. Copy and memory are scored up to 36 points. Boston Nomination Test short version of 12 cards\textsuperscript{19}, Pyramids and palm trees associative semantic knowledge test\textsuperscript{20} with 52 cards in which three stimuli are presented at the same time, one in the upper part of the card and the other two in the bottom. Of both stimulus, the patient must choose the one which fits with the upper stimulus; verbal fluency test for words beginning with the letters FAS in one minute, and semantic fluency test\textsuperscript{21}, several executive function evaluation tests (EF) including Trail Making Test (TMT) part A (TMTA) and B (TMTB)\textsuperscript{22}, phonological verbal (VF) and semantic (SF) fluency\textsuperscript{23}, Milwaukee Card Sorting Test (MCST)\textsuperscript{24}. EF includes goal selection and action strategies, performance monitoring and flexibility for change according to contextual cues.

Neuro-psychiatric Exam

Neuropsychiatric Inventory (NPI) was administered to both groups\textsuperscript{25}. This exam consists of semi-structured interview administered by a clinician where 12 areas of behavior and neuropsychiatric functioning are evaluated, including delirium, hallucinations, restlessness, anxiety, euphoria, aberrant motor behavior, sleep disturbances, appetite abnormalities, apathy, depression, irritability, dis-inhibition. Each item is scored according to the frequency ($f$) of its presence (scale 0 to 4) and severity ($s$) of the abnormality (scale 0 to 3), giving a total score of $f \times s$, maximal score is 144.

Brain 99mTc-ECD SPECT neuroimages acquisition and assessment

Both groups of patients were administered an intravenous injection of 1110 MBq 99mTc-ECD (ethylcysteinate dimer) while lying in a rest condition in a quiet, dimly room. Subjects were imaged using a dual-head rotating gamma camera (VG MILLENIUM GE) with a low energy, high-resolution collimator, 30 minutes after intravenous injection of 99mTc-ECD. A 128 $\times$ 128 pixel matrix was used for image acquisition with 120 views over a 360° orbit, with a pixel size and slice thickness of 1 mm, in 27 min or more if total counts were lower than 5 $\times$ 106. Image reconstruction was performed by a ramp filtered-back projection and three dimensionally smoothed with a Metz filter (order 3, enhancement 1.24, FWHM 6.7 mm, cut-off 0.61 cycles cm$^{-1}$). The reconstructed images were corrected for gamma ray attenuation using the Chang method (attenuation coefficient: 0.11 cm$^{-1}$).

Image pre-processing and analysis

Statistical Parametric Mapping (SPM2, Welcome Department of Cognitive Neurology, University College, London), and Matlab 6.1 (Mathworks Inc., Sherborn, MA) were used for image pre-processing. Images were spatially normalized to a reference stereotactic template (Montreal Neurological Institute, MNI), and smoothed by a Gaussian kernel of $8 \times 8 \times 8$ mm FWHM SPECT data analysis was performed in blind to clinical diagnosis. Comparisons between groups were made using t-statistics with appropriate linear contrasts in order to evaluate brain hypoperfusion in FTD patients. Slides were assessed anonymously. Intra-raters differences was good ($\kappa = 0.7$).

Theory of Mind Assessment with Strange Story Test (modified from Happe)

In the present study a series of short stories describing everyday situations where the characters tell things that must be interpreted with a metaphorical sense are presented to the two groups of patients. For example, in a scene someone gets a present and says “just what I needed”, this in turn can be interpreted in several ways; because the guy really needed and wanted the present, or may be he said that politely not to dismiss the other person. In real life the purpose of what persons tell can be distinguished using some elements that belong to the context, by the way emotional expression or relationship between the teller and the listener. In the Strange Stories Test the purpose of what characters tell can be inferred in a relatively simple way. These stories are more ecological and represent a refinement in the evaluation of patients with dementia, compared with standardized ToM tests. Original stories were modified to adapt and fit the context and cultural background of the subjects. Also a control task was required to evaluate physical events and comprehension ability independent of any deficit in ToM.
**Materials**

Stimulus consists of short stories, with 6 mental stories and 2 physical content control stories. Mental stories include the following types: 1) joke, 2) white lie, 3) figured language, 4) irony, 5) misunderstanding, 6) forgetfulness; and the physical content control stories included: 7) battle and 8) glasses. Physical content control stories didn't included mental states, nor had a social component, but demanded a global inference to be made beyond what explicitly was told in the text. Mental stories tests included 2 or 3 test answers; the comprehension answer had the form of: “is it true what S said?” and the justification answers took the form of: “why do you suppose that S said that?” Physical content control stories had only one answer, which took the form of “why a particular action has been made?”. Mental test was followed by a drawing which contained main characters, with emotional expressions and the proper context. These were obtained from Francesca Happé and modified to fit in the local cultural context. Mental tests were presented first followed by control task, which served to signal a state of fatigue or lack of motivation, and to test the existence of global and no social inferences as well. These tests were presented in a random order and once in front of the subject it was told to listen carefully to the story because later he will be prompted to answer some issues related to the stories. After reading each mental story the examiner made some questions of comprehension and of mental state, while the story remained in front of the examinee, to avoid mistakes due to memory deficits. Comprehension questions took the form of: “is it true what S said?”; and when the examinee made a mistake the story was repeated until a correct answer was obtained or a justification that demonstrated that the subject had understood the story.(e.g. “well, it is not true if you see it in a literal way, but the expression is correct”). Mental state questions took the form of “why did the S said that?”. Examinees had enough time to answer and were prompted to read again the story by their own to ease the explanations they gave about the expressions in the story after completing the mental stories they accomplished the physical content control tasks, which were read aloud. After that, the patients were asked the inference question about why a certain action had been done in that way.

**Scoring**

For the mental stories each subject was assigned a score according to the correct answers on comprehension and justification. For the former the original answer was taken into account along with the probable corrections. For the latter the score was obtained according to a number of options. Justifications could be adequate or inadequate, in two different forms: from a physical perspective or a mental one. For example in the story in which Julia (one of the characters) is playing and she pretends that a banana stands as a telephone, the explanation is that the banana is a phone can only be correctly justified in two different ways: in a physical sense (“banana has a shape which resembles that of a telephone”), or in a mental sense. (“because Julia is pretending that the banana is a telephone”). In a similar fashion, an incorrect justification could be made on two different ways: in a physical sense “because Julia is about toe at the banana” and in a mental sense: “because Julia is playing”. When both correct and incorrect justifications were made, subjects were classified by the correct answer. In this way the got credit scores for the best answer. In a similar fashion if answers where both mental and physical, justification was classified as if it was a mental one. Answers of physical states included terms such as big, similar to, has the shape of, to sell, to get rid of, for not receiving an X (a physical result, such as get fired, get wet, etc). Answers of mental state included those referred to thoughts, wishes, trends, feelings, dispositions. Justifications of mental states included terms such as: wish, like, think, be, joke, pretend, lie, fear of, hurt, expectation and cheat. For the physical control stories the subjects were assessed according their answers to global inference questions, which asked why something had happened or why a certain action had taken place. Best answer was always accepted. Answers adequacy assessment was made on subjective basis.

**Statistical analysis**

Were performed using the SPSS program. For group comparisons were employed variance analysis (ANOVA) with paired ad-hoc comparison using Turkey Test or t-tests for non paired comparisons. Parametrical statistics were used for normal distributions and variables of the ToM and SST were compared using co-variance
analysis (ANCOVA). For SST and brain atrophy concordance, Spearman’s rank correlation and moment product analysis were used, admitting that impairment takes place when the score is below 1.5 SD of usual performance.

RESULTS

Comparison between group of patients with FTD and AD didn’t show statistically meaningful difference in the following: MMSE, ACE, WAISC, VIC, PIC, forward and backward digit span ($p>0.05$). Scores obtained for WCST showed general performance derangement specially in WCST average score, impairment in set maintenance and first category trials for both AD and FTD, but no statistically meaningful difference between both groups assessed with ANOVA was found, with the exception of completed categories (conceptual abilities) ($p<0.01$) and perseverative errors (cognitive flexibility) ($p<0.001$) with worst scores for FTD (Figure 1).

AD: Alzheimer Disease; FTD: Frontotemporal Dementia; SF: Semantic Fluency; T1stC: trial first category; NPE: Non-perseverative Errors; PE: Perseverative Errors; CC: Completed Categories; GS: Global Score.

Results obtained in TMT A and B didn’t showed meaningful differences between FTD and AD. In the episodic memory test there was a statistical difference between both groups of FTD and AD for immediate ($p<0.01$) and delayed ($p<0.001$) logical story recall and Rey Complex Figure 45’ ($p<0.01$); showing that FTD group has a greater impairment compared to AD. In verbal and semantic fluency there was a worst performance in those with AD compared with FTD ($p<0.01$) and the opposite happened with the Boston Naming Test short version with 12 cards in which case FTD performed better than AD ($p<0.01$). Neuropsychiatric Inventory (NPI) showed scores of $46.3 \pm 4.5$ in the case of FTD and of $5.2 \pm 2.2$ in the case of AD; analysis of variance (ANOVA) showed statistically meaningful difference between both $[t (40) = 8.9<0.001]$ (Table 2).

The answers to the SST allow assessing if there are differences between AD and FTD groups in the ability of giving explanations and contextual justifications to the metaphorical sentences used by the subject in the story in the mental condition. It also assess if the patients could make inferences on the purpose by which a certain action had been made in the physical condition. Scores in the mental and physical conditions were measured as percentages, to match the 6 stimulus in the mental task with the 2 stimulus in the physical task. The scores ranged from 0% to 100% for each one of the conditions (Table 3).

Each group scores were analyzed with ANOVA test, with a group variable between subjects and a condition variable within subjects (mental and physical). ANOVA test showed group $[F (2,49) = 4.76, p<0.05]$ and condition $[F (1,92) = 8.02, p<0.001]$ statistical significance effect, suggesting that both groups differed in the correctness of their answers and that the condition variable (whether mental or physical) had an effect in the performance. Group and condition correlation was meaningful $[F (2,57) = 8.72, p<0.01]$. Different group effects in both conditions was assessed using simple effect analysis comparing different groups in each condition, which showed a meaningful effect for mental condition only $[F \text{ mental} (2,49) =17.30, p< 0.001; F \text{ physical} (2,53) = 1.47, p = 0.63]$. Group and condition correlation was assessed with contrast t-test of cells averages, which showed that the average percentage of patients with AD in mental condition was statistically different from the patients with FTD $[t \text{ DFT} (20) = 6.43, p<0.001; t \text{ DTA} (20) = 4.74, p = 0.001]$ and the average percentage for the group with FTD in mental condition had a marked tendency to be lower than the group with AD $[t (20) = 2.76, p = 0.07]$. The analysis of average percentages show that both groups have poor performance when giving (correct) explanations that had fit to the context in the stories (Figure 2).

Quality of justifications, either mental or physical,
### Table 2
**Neuropsychological exam outcomes**

<table>
<thead>
<tr>
<th></th>
<th>AD (Median SD)</th>
<th>FTD (Median SD)</th>
<th>ANOVA F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>22.2 (3.8)</td>
<td>21.9 (3.7)</td>
<td>2.22</td>
<td>NS</td>
</tr>
<tr>
<td>ACE</td>
<td>81.3±3.1</td>
<td>82.1±2.7</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>WAIS (n=40)</td>
<td>85.5±2.7</td>
<td>86.2±1.5</td>
<td>2.34</td>
<td>(0.97)</td>
</tr>
<tr>
<td>VIC (n=40)</td>
<td>84.4±3.4</td>
<td>84.1±3.1</td>
<td>2.65</td>
<td>(0.60)</td>
</tr>
<tr>
<td>PIC (n=40)</td>
<td>82.3±3.6</td>
<td>83.6±2.8</td>
<td>2.45</td>
<td>(0.66)</td>
</tr>
<tr>
<td>NPI</td>
<td>5.2±2.2</td>
<td>46.3±4.5</td>
<td>5.57</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EDY (n=40)</td>
<td>5.5±1.6</td>
<td>6.0±1.1</td>
<td>2.45</td>
<td>(0.77)</td>
</tr>
<tr>
<td>WAIS Comprehension</td>
<td>14.12±4.30</td>
<td>15.51±5.49</td>
<td>3.47</td>
<td></td>
</tr>
<tr>
<td>WAIS Vocabulary</td>
<td>23.11±9.53</td>
<td>35.48±17.03</td>
<td>0.208</td>
<td></td>
</tr>
<tr>
<td>WAIS Cue Number</td>
<td>9.97±5.16</td>
<td>10.15±6.12</td>
<td>0.339</td>
<td></td>
</tr>
<tr>
<td>WAIS Cubes</td>
<td>10.12±4.45</td>
<td>11.02±5.36</td>
<td>0.613</td>
<td></td>
</tr>
<tr>
<td>WAIS puzzle</td>
<td>10.59±4.49</td>
<td>11.31±5.11</td>
<td>0.449</td>
<td></td>
</tr>
<tr>
<td>Digit Span backward</td>
<td>4.5±1.2</td>
<td>4.6±1.5</td>
<td>NS 2.56</td>
<td></td>
</tr>
<tr>
<td>Logic Recall Immediat</td>
<td>5.9±3.67</td>
<td>8.7±3.8</td>
<td>5.1</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Logic Recall Delayed</td>
<td>2.1±4.1</td>
<td>6.0±3.6</td>
<td>3.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Rey F. copy</td>
<td>26.3±4.8</td>
<td>28.1±5.1</td>
<td>2.44</td>
<td>NS</td>
</tr>
<tr>
<td>Rey F. recall.45’</td>
<td>6.3±1.6</td>
<td>13.1±2.1</td>
<td>1.33</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>12.2±1.2</td>
<td>7.2±1.8</td>
<td>3.55</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Semantic Fluency</td>
<td>14.9±3.2</td>
<td>6.2±1.5</td>
<td>7.45</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Trail making A</td>
<td>9.4±2.8</td>
<td>8.2±1.9</td>
<td>2.44</td>
<td>NS</td>
</tr>
<tr>
<td>Trail making B</td>
<td>5.4±2.4</td>
<td>4.2±2.8</td>
<td>2.55</td>
<td>NS</td>
</tr>
<tr>
<td>MCST GA</td>
<td>4.2±1.9</td>
<td>3.9±2.8</td>
<td>3.45</td>
<td>NS</td>
</tr>
<tr>
<td>Completed Category</td>
<td>5.1±2.3</td>
<td>2.8±2.1</td>
<td>2.78</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Perseverative Errors</td>
<td>4.7±3.8</td>
<td>9.4±4.6</td>
<td>2.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non perseverative Err</td>
<td>8.5±4.1</td>
<td>11.3±2.9</td>
<td>4.67</td>
<td>NS</td>
</tr>
<tr>
<td>Trials 1* category</td>
<td>16.4±3.6</td>
<td>18.2±1.7</td>
<td>6.78</td>
<td>NS</td>
</tr>
<tr>
<td>Set Maintenance Fail</td>
<td>1.8±0.5</td>
<td>3.3±1.6</td>
<td>2.44</td>
<td>NS</td>
</tr>
<tr>
<td>P &amp; P Test</td>
<td>42.1±1.9</td>
<td>41.4±2.6</td>
<td>5.66</td>
<td>NS</td>
</tr>
<tr>
<td>Boston Naming</td>
<td>9.35±3.3</td>
<td>4.21±2.8</td>
<td>7.51</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Both groups matched on age (years), education (years) and sex (male/female). MMSE = mini-mental state examination ACE = Addenbrooke Cognitive Evaluation. WAISIQ = Weschler Adult Intelligence Scale Intelligence Quotient. VIC = Verbal Intelligence Coefficient. PIC = Performance Intelligence Coefficient. YDS = Yesavage Depression Scale. MCST = Milwaukee Card Sorting Test; GS = global score. P&P Test = palms trees and pyramids Test. Scores are expressed as Median and Standard Deviation.

### Table 3
**Types and number of justifications in mental condition**

<table>
<thead>
<tr>
<th>Group</th>
<th>Mental</th>
<th>Physical</th>
<th>Mental Justification</th>
<th>Physical Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD(n=20)</td>
<td>3.74±1.1</td>
<td>2.26±1.2</td>
<td>1.92±1.2</td>
<td>1.06±1.1</td>
</tr>
<tr>
<td>FTD(n=20)</td>
<td>2.91±0.7</td>
<td>3.02±0.9</td>
<td>1.11±1.5</td>
<td>2.02±2.9</td>
</tr>
</tbody>
</table>

AD = Alzheimer dementia. FTD = Fronto-temporal dementia. Scores are expressed as Median and Standard Deviation.
right or wrong, offered by the examinees in mental condition was assessed. The number of answers corresponding to each category was assumed as unit of measure. Scores for each type of justification ranged from 0 to 6. Mental and physical justifications were analyzed with ANOVA test of variance, including correct and incorrect answers as well, with the aim of assessing differences between AD and FTD groups. The analysis revealed differences between groups in the number of mental justifications \[ F(2,47) = 0.37, p<0.05 \], and physical \[ F(2,47) = 0.18, p<0.05 \], in the number of correct mental justifications \[ F(2,47) = 0.61, p<0.05 \] and correct physical justifications \[ F(2,47) = 2.34, p<0.05 \]. No statistical meaningful differences were found between physical and mental incorrect justifications \[ F(2,47) = 5.31, p = 0.71 \]. Patients with FTD showed fewer correct mental and physical justifications than the AD group \[ t_{DFT}(20) = -3.65, p = 0.001; t_{DTA}(20) = -2.18, p<0.005 \] but no differences were found in the number of incorrect mental and physical justifications \[ t(40) = 0.86, p = 0.40 \]. No meaningful differences between groups in the number of comprehension mistakes and omissions in mental stories, evaluated with Chi Square analysis.

**Relationship between SST and ventromedial/dorsolateral brain hypometabolism in FTD**

Correlation between outcomes in SST and SPECT was assessed in patients with a diagnosis of FTD and AD. Of those with FTD, but no with AD, almost 90% showed marked impairment in ToM tasks. A concordance between worse results and greater hypoperfusion in ventromedial frontal cortex \( x, y, z = 5, 18, -14; T = 5.27, \) cluster size = 872), left dorsolateral frontal cortex \(-12, 15,-42; T = 4.78, \) cluster size = 459) and left temporal pole \(-46,-6,-48; T = 5.61, \) cluster size = 992) was found, and product-moment correlation statistical analysis showed correlation values \( r = 0.873 (p<0.01) \) between degree of ventromedial cortical hypometabolism and SST deficits. Cortical hypoperfusion was assessed with a visual scale. In most cases hypoperfusion degree was greater in ventromedial \( r = 0.83; p<0.002 \), than in dorsolateral area \( r = 0.67; p<0.007 \) (Figure 3).

**DISCUSSION**

In mental condition both groups differed meaningfully in the number of mental justifications, as a consequence of being unable of using words describing mental states (such as hurt, laugh, pretend) according to the context depicted in the short stories. Both groups made more context inappropriate mental justifications than appropriate ones, but FTD group had the worst scores. Mistakes in mental condition were remarkably noticeable in certain occasions, for example, when Ana’s mother says: “Ok, that’s fine, it’s what I call to be educated”, a patient’s answer with FTD was that the expression of the mother was correct and he justified that because Ana had a good behavior and hadn’t screamed. In simulation story another patient’s answer was that the
expression of Julia “look, the banana is a telephone!” was correct because she had put the banana on her ear. Both groups had a good performance in physical condition, in which they were required to justify why a certain action had happened. Analysis of the type of justifications in mental condition reveals that both groups differ in the number of mental and physical justifications and in the number of correct justifications as well, but didn’t show any differences in the number of incorrect justifications. The number of correct physical justifications and the absence of differences in number of incorrect justifications, mistakes and omissions as well keep in line with the hypothesis of an impairment in ToM in FTD patients. If that wouldn’t be the case, more physical incorrect justifications should have happened in both groups provided that both had a general comprehension difficulty. One possible explanation may be that in the strange stories two kinds of answers are made. First one refers to the truthfulness of the non literal statement in the story, which demands a lower level of execution. The subject must guess whether the metaphorical, non ambiguous statement, is in accordance with the context of the depicted situation in the story. The second one refers to justification answer, which requires that the patient give a contextually appropriate explanation. An additional load is the requirement to take the place of the characters, both mentally and physically, using the inferential, pragmatic and situational resources, which in time demands a higher level of execution. This situation reflects the lower comprehension abilities and integration of the informative content of the stories in FTD compared to AD. Lesser motivation doesn’t seem to be the cause of the difficulties exhibited by FTD patients, as they have the same number and quality of omissions as AD patients. Subjects of each group generated equal number of answers, but FTD patients had impairment in providing answers with mental, contextual and social content. Lower number of mental condition answers with contextual content in FTD patients suggests they had impaired central coherence ability, as they focus on literal statement ignoring contextual meaning. In this way they keep local coherence but ignoring the global one. For example, in white lie story, Aunt Juana asks Pedro: “do you like my new dress?” This in time answers “it’s gorgeous!”. When asking to a FTD patient why Peter had said that, the patient replied that the dress was new and for that reason it was gorgeous. In that way, by making an inference with correct physical content, shows a difficulty in inferring how the character feels in that moment of the story. These keep in line with actual knowledge on deficits in ToM and not in lack of familiarity with mental states. The better performance in physical content compared with mental ones in FTD deserves an explanation. While in mental condition it is

Maps of significant voxels representing regions of hypoperfusion in FTD superimposed on T1-weighted MRI images. Statistical threshold, P<0.001, T≥3.99, minimum cluster size = 50 voxels. Neurological convention: left is on the left side and vice versa.

Figure 3. SPECT to assess ventromedial and dorsolateral frontal cortical hypometabolism in FTD.
required to infer the meaning of the expression, in physical control tasks it is required wider cause-effect knowledge instead of a social one. Summarizing, differences in mental condition performance between AD and FTD suggests an impairment in ToM in the latter, due to either a damage in inferring the meaning of the speaker expression with the aid of the context in which it is located (ToM) or because they have an impairment when guessing the mental states that are explored with the SST. This in turn can be discarded due to the fact that they make equal number of physical and mental wrong justifications, and less correct justifications.

CONCLUSION

According with the above results, patients with FTD seem to exhibit a specific deficit in Theory of Mind which renders them incapable of building inferences on mental states.

REFERENCES

http://dx.doi.org/10.1162/jocn.2007.19.4.671
http://dx.doi.org/10.1136/jnnp.2006.098228
http://dx.doi.org/10.1176/appi.neuropsych.21.1.100
http://dx.doi.org/10.1093/brain/aws314
http://dx.doi.org/10.1007/BF02172093
http://dx.doi.org/10.1136/jnnp.2006.100669
http://dx.doi.org/10.1016/j.neuropsychologia.2009.08.001
http://dx.doi.org/10.1016/j.neubiorev.2008.09.008
http://dx.doi.org/10.1044/nnsl19.4.117
http://dx.doi.org/10.1136/jnnp.57.4.416
http://dx.doi.org/10.1016/0022-3956(82)90033-4
http://dx.doi.org/10.1016/0022-3956(75)90026-6
http://dx.doi.org/10.1093/brain/awp81
23. Tombaugh T, Kozak J, Rees L. Normative data for the controlled oral word


http://dx.doi.org/10.1093/arclin/11.6.541


http://dx.doi.org/10.1098/rstb.2008.0318