

Early language experience has long-lasting impact on the neural basis of theory of mind

Jorie Koster-Hale^{1*}, Naomi Berlove², Rachel Magid¹, Rachel Benedict³, Jennie Pyers⁴, and Rebecca Saxe¹

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge MA 02139

² Department of Psychology, Tufts University, Medford MA 02155

³ Department of Deaf Education, Boston University, Boston MA 02115

⁴ Department of Psychology, Wellesley College, Wellesley MA 02481

Abstract

The ability to reason about people's internal states, "theory of mind", has been shown to rely on a distinct and reliable group of brain regions. What triggers the specialization of these brain regions: hard-wired maturation or experience during development? Here, we ask this question by examining the neural selectivity of theory of mind brain regions in adults with different developmental experience: late access to language. Delayed access to language results in delayed development of theory of mind. Despite matched language competency and intelligence in adulthood, adults with delayed language experience (Deaf adults born into non-signing households; DoN) show less selectivity in their theory of mind brain regions than adults with early language access (Deaf adults born to signing households; DoS), in both verbal and nonverbal theory of mind tasks. Moreover, individual differences in neural selectivity are predicted by extent of early language experience, suggesting that the emergence of strong selectivity in these brain regions is influenced by experience in childhood. In social brain regions that become highly specialized by adulthood, the degree of specialization is determined by aspects of developmental history.

Keywords: theory of mind; language; deafness; fMRI; development

Background

Social life depends on recognizing that people's actions are consequences of an unobservable, internal causal structure: goals, intentions, and beliefs. In typically developing children and adults, this ability recruits a distinct and reliable group of brain regions, sometimes called the "mentalizing" or "theory of mind" network (e.g. Aichhorn et al. 2009; Saxe & Kanwisher 2003), which includes regions in human temporo-parietal junction (TPJ), medial precuneus (PC), and medial prefrontal cortex (MPFC). One of these regions, right TPJ, shows increased activity when participants think about others' beliefs and desires, but not about others' appearance, social background, stable traits, or even physical sensations (Bruneau & Saxe 2012; Saxe & Powell 2006). Moreover, this selectivity develops late and relates to successful mental-state reasoning on behavioral tasks (Saxe et al. 2009; Gweon et al. 2012). A key question, then, is what triggers the specialization of these brain regions: hard-wired maturation or experience during development?

One possibility is that some number of cognitive functions are implemented by innately specified brain regions (Fodor 1983). These (potentially) modular brain regions would allow rapid and efficient computation of evolutionarily significant cognitive processes, for example, vision. On this view, the association between a selective functional profile, such as increased response to mental state inference, and a specific brain region, such as the RTPJ, would provide evidence that mental state inference is subserved by an innate cognitive module.

An alternative possibility is that specific brain regions acquire highly specialized functions through experience. In particular, extensive early experience may drive the organization of cortex, so that populations of neurons adopt specific functions. For example, behavioral training in monkeys increases neuronal selectivity in cortex recruited for object recognition (Baker et al. 2002; Sigala & Logothetis 2002; Freedman et al. 2006; Cox & DiCarlo 2008). If these populations of neurons are spatially contiguous, functionally coherent patches of cortex may be visible using techniques like fMRI. One such example is the visual word form, an area in human temporal cortex responds selectively to words in familiar (but not unfamiliar) alphabets (Baker et al. 2007).

If specialized brain regions can be organized by intense early experience, then RTPJ may acquire its highly specific response to other minds over the course of childhood experiences. One way to test this hypothesis is to study the neural mechanism of ToM in individuals who have had atypical developmental histories with regard to theory of mind. Here, we leverage prior evidence that early development of ToM depends on exposure to language and to conversation about the mind (e.g. Milligan et al. 2007; Moeller & Schick 2006; de Villiers 2005).

D/deaf¹ children born to non-signing parents (DoN) often have delayed access to language, including language about mental states compared, to D/deaf children born to signing parents (DoS). As a consequence, they exhibit delayed performance on standard measures of ToM reasoning (Figueras-Costa & Harris 2001; Meristo et al. 2007; Peterson & Wellman 2010; Peterson & Siegal 1999; Schick et al. 2007; Woolfe & Want 2002). However, later in childhood, DoN children appear to catch up to DoS children in ToM performance (Peterson & Wellman 2010; Schick et al. 2007). Since most d/Deaf children are

*jorie@mit.edu

¹ Deaf with an uppercase "D" usually refers to adults and children who share the use of American Sign Language and Deaf culture, while a lowercase "d" is usually an audiological description of a person's hearing level.

otherwise typically developing, but vary in their early language acquisition, they provide a unique window into the effects of early language experience on theory of mind. Here, we ask whether this difference in language experience, and subsequent delay in ToM development, leave traces on the neural basis of ToM.

Methods

Participants

Nine Deaf adults born to signing parents (DoS, 4 women, 31±11 years old), and 11 Deaf adults born to non-signing parents (DoN; 4 women, 35±16 years old, average age of ASL acquisition = 1.75 years) participated. All participants were profoundly (90+DB loss) deaf, and fluent, daily-users of American Sign Language (ASL). None of the participants suffered from neurological disorders or had ever sustained head injury. All subjects gave informed consent and were compensated \$30 an hour. An additional 3 DoN participants were scanned, but are excluded from all analyses; two because of high scanner motion and one because of poor (chance) behavioral performance.

Language Screening and Background Survey

All participants were screened by a native ASL signer to determine current ASL fluency. Only fluent signers were included in the study. In the post-scan debriefing, all participants reported that they fully understood the stories presented and were familiar with the language used. Additionally, participants filled out a survey about their language background, answering questions about the language(s) they used growing up, in school, at work, the age that they were first exposed to ASL, etc. Total ASL experience was calculated using a composite score for questions related to language exposure in childhood, on a scale from 0-15. Current ASL use was calculated using a composite score for questions related to language use at home, with friends, and at work.

fMRI Protocol and Tasks

Verbal theory of mind and language localizer: Stimuli in the verbal task were 60 video recordings of a Deaf native signer, who was naive to condition labels and hypotheses. Modeled off of previous tasks design to elicit theory of mind (Perner & Aichhorn 2008; Saxe & Kanwisher 2003; Dufour et al. 2012; Dodell-Feder et al. 2011), 36 stimuli consisted of a Deaf native signer telling short narratives in American Sign Language, in three conditions: 12 stories featuring mental states (e.g. beliefs, desires, emotions), 12 stories featuring socially-relevant information (physical appearance, physical sensation, social relationships), 12 stories featuring physical events (e.g. physical causality, weather), **Figure 1**.

The design allowed us to test for both the existence and selectivity of brain regions recruited for theory of mind. Comparing mental and physical stories allows us to localize regions in the brain that are recruited for mental state reasoning more than for physical reasoning; mental vs physical contrasts are common for localizing the theory of mind network. Second, looking at the response of these brain regions to mental versus social information allows us

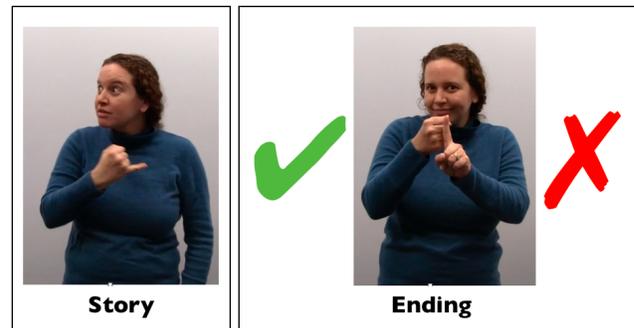


Figure 1: Frames from example stimuli.

to ask whether these regions are selective for mental state content above general social content (Gweon et al. 2012).

All 36 stories were normed by 10 naive, Deaf native signers. Based on behavioral ratings, stories were matched for linguistic features (e.g. syntactic complexity, number of signs, number of verbs), psychological features (e.g. how easy to understand, how interesting), and imageability. Stories were told using simple language, in an enthusiastic, narrative way.

To encourage engagement in the scanner, stories were presented in two consecutive segments: the main story and a final sentence. Half of the stories had mismatching final sentences, taken randomly from another story. After each trial, participants judged whether the final sentence matched the main story. During the final sentence, participants saw a green check mark on the left and a red X on the right side of the screen as response reminders (left button for match, right button for nonmatch). These response images remained on the screen until the participant made a yes-or-no button response. Reaction time was measured from the onset of the final sentence.

The remaining 24 stimuli consisted of the same Deaf native signer signing strings of nonsense signs. The nonsign lists consisted of a series (3 or 8) of phonologically possible, but meaningless, signs, paired with grammatically possible, but meaningless, facial expressions. After each list, participants saw an identical or novel sequence of non-signs and had to judge whether the final sequence matched the signs in the list. Non-sign sequences were normed by 10 Deaf native signers to exclude signs similar to known signs or regional slang, and for task difficulty. Non-signs and stories were matched for low-level visual properties (e.g. motion energy) and duration.

The nonsign condition was included as a control for language processing. Processing pronounceable non-words engages many of the low-level processes required for language processing, such as visual /auditory processing, phonological recognition and composition, and working memory, without recruiting higher-level processes, such as lexical access, word and sentence level composition, syntactic structure building, or semantic computation. Thus, comparing meaningful sentences to lists of pronounceable non-words serves to localize the regions specifically recruited for language processing at the word and sentence level (e.g. Fedorenko et al. 2010; Cutting et al. 2006; Friederici et al. 2000; Hagoort et al. 1999; Heim et al. 2005; Humphries et al. 2006).

Stimuli were presented via Matlab 7.6 running on an Apple MacBook Pro. The complete experiment included

five 12-minute runs. Each run consisted of 12 experimental blocks. Each block was 60s long, consisting of the main story (29-41s), the ending sentence and answer period (7-11s), and a rest period (8-20s). The order of conditions was counterbalanced across runs.

Nonverbal theory of mind: Because the groups differed in early access to language, we also included a non-verbal measure of theory of mind that did not directly rely on language processing. The non-verbal task was free viewing of a 5.5-minute clip of Pixar's short film "Partly Cloudy" in a single run. The movie was independently coded by 5 researchers for 2-10s segments in four conditions: segments that elicited thinking about mental information (e.g. false beliefs, emotional expression), someone else's physical sensations (e.g. pain), social interactions between two characters (e.g. hugging), and the physical environment (e.g. natural scenes). Coding identified 4 mental segments totaling 44s, 5 social interaction segments totaling 28s, 7 social sensation segments totaling 26s, and 3 physical segments totaling 24 s.

Acquisition and Preprocessing

fMRI data were collected in a 3T Siemens scanner at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, using a 32-channel head coil. Using standard echoplanar imaging procedures, we acquired blood oxygen level dependent (BOLD) data in 26 near axial slices using 3x3x4 mm voxels (TR=2 s, TE=40 ms, flip angle=90°). For steady state magnetization, the first 4 seconds of each run were excluded.

Data processing and analysis were performed using SPM8 and custom software. Data were motion corrected, realigned onto a common brain space (MNI template), spatially smoothed using a Gaussian filter (FWHM 5mm) and subjected to a high-pass filter (128 Hz).

fMRI Analysis

All fMRI data were modeled using a boxcar regressor, convolved with a standard hemodynamic response function (HRF). For both tasks, a general linear model was used to analyze the BOLD data from each subject, as a function of condition, using a slow event related design. Model included nuisance covariates for run effects, global mean signal, and an intercept term. Additionally, in the verbal ToM task, the response portion (including the final sentence) of the task were entered as two separate regressors for matching and mismatching endings.

Functional Localizer: Whole brain analysis (Verbal)

A second-level random effects analysis was performed on the contrast images generated for each individual to identify brain regions showing reliable differences between belief and physical stories in the verbal task ('Theory of Mind' regions) and between physical stories and non-signs ('Language' regions) by performing Monte Carlo permutation tests using SnPM5 (Nichols and Holmes, 2002; Hayasaka and Nichols, 2004; <http://www.sph.umich.edu/ni-stat/SnPM/>), with a false positive rate controlled at $p < 0.05$, with 5000 permutations and a variance smoothing of 3mm.

Functional Localizer: Individual ROIs (Verbal)

Based on the results of the whole-brain analysis, individual regions of interest (ROIs) were defined using a whole-brain contrast of mental state stories relative to physical control stories ($p < .001$, uncorrected).

Functional Localizer: ROIs response (Verbal)

To measure the response of these ROIs to the localizer stimuli without the bias of non-independent data, we used a cross-validation to define ROIs and extract percent signal change. Individual subject ROIs were defined using two runs of data, and the response was extracted from the excluded, independent run. This process was iterated over all six runs, allowing us to calculate, in each of the individual regions of interest (ROIs) defined using the localizer, the average percent signal change (PSC) relative to baseline for each time point in each condition, averaging across all voxels in the ROI and across all blocks in the condition, where $PSC(t) = 100 \times (\text{average BOLD magnitude for condition } (t) - \text{average BOLD magnitude for fixation}) / \text{average BOLD magnitude for fixation}$. We averaged the PSC across the entire stimulus presentation, offset 6s from presentation time to account for hemodynamic lag, to get a single PSC for each condition, in each ROI, in each participant (Poldrack, 2006). These values were then averaged across subjects to get a PSC value for each condition for each ROI.

Based on prior evidence for changes in the functional profiles of some of the ROIs as a function of age (Saxe et al., 2009), we examined response selectivity in each ROI for each participant; that is, we determined the degree to

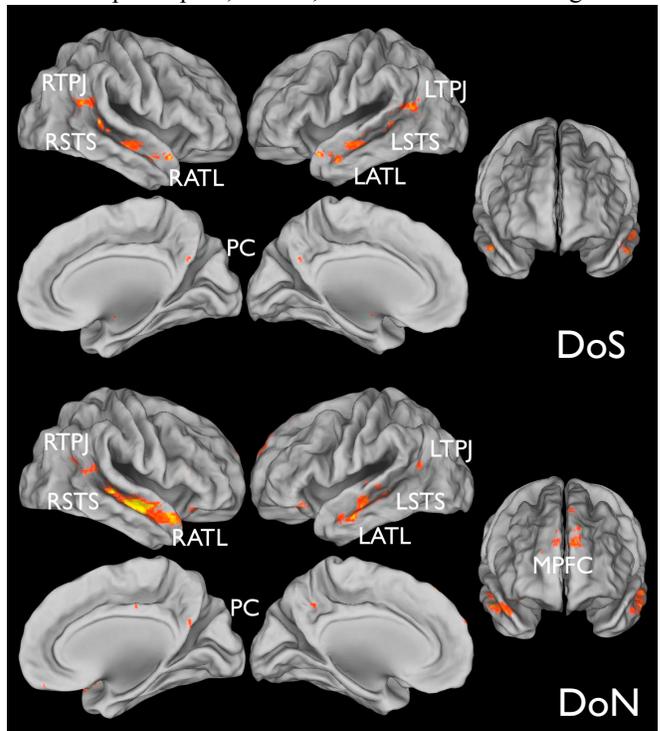


Figure 2: Verbal Theory of Mind task – Whole brain results. DoS (top, N=9) and DoN (bottom, N=11). Brain regions in which the bold signal is higher for stories about mental events compared to stories about physical events (random effects analysis, $p < 0.001$, uncorrected).

which responses were selective for mental state information versus general to any social information. Selectivity was measured by comparing PSC of mental stories to social stories: a larger difference between mental and social stories indicated higher selectivity.

Nonverbal Theory of Mind task: individual ROIs

In each of the individual regions of interest (ROIs) defined using the verbal localizer, we calculated the response to the four segment types.

Results

Survey

From the survey, DoN adults were exposed to ASL significantly later than DoS adults (DoN: mean±sd = 1.75±2.1 years; DoS: 0±0; $t(17)=2.5, p=.02$), and had significantly less ASL exposure during childhood (DoN: 6.5±2.4; DoS: 11.6±1.6; $t(18)=5.4, p<.001$). However, there was no difference between groups in current ASL use (DoN: 1.5±0.7; DoS: 1.8±0.7; $t(18)=1.1, p=.3$).

Verbal Theory of mind task: behavior

Both DoS and DoN performed well on the scanner task, with an average of 87% correct ± 1.5% standard error across conditions, and 6.4±0.1 s reaction time. There were no significant effects of group or condition on either accuracy or RT (Accuracy: group: $F(1,18)=, p=.97, \eta^2=0$; condition: $F(3,54)=1.16, p=.33, \eta^2=.02$; group by condition: $F(3,54)=1.23, p=.31, \eta^2=.03$. RT: group: $F(1,18)=1.44, p=.25, \eta^2=.03$; condition: $F(3,54)=1.42, p=.25, \eta^2=.04$; group by condition: $F(3,54)=1.25, p=.3, \eta^2=.04$).

Verbal Theory of Mind task: theory of mind

Whole brain analysis (Verbal): Both DoS and DoN adults had canonical-looking neural responses to mental stories relative to physical stories, with activation in bilateral temporal-parietal junction (RTPJ), bilateral superior temporal sulcus (STS), bilateral anterior temporal lobe (ATL), and precuneus (PC). Additionally DoN adults had activation in regions in bilateral inferior front gyrus (IFG) and medial prefrontal cortex (MPFC), as well as the left frontal eye field, though those regions showed no

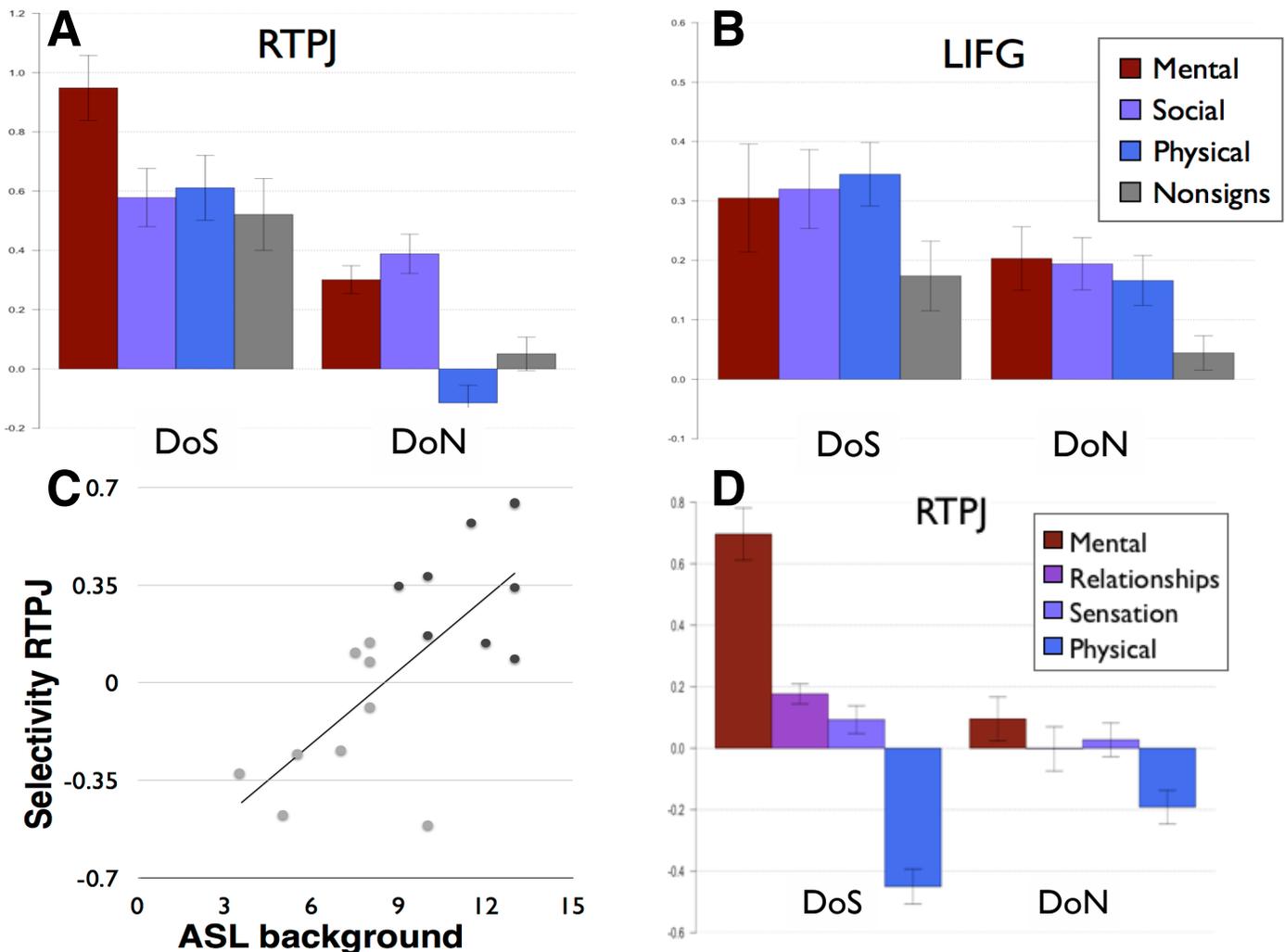


Figure 3: Individual ROI analyses. (A) Average percent signal change during the verbal ToM task (cross-validation) in RTPJ and (B) in LIFG. (C) Selectivity (mental-social) correlates with self-reported ASL background. (D) Average PSC during nonverbal ToM Task in RTPJ.

significant difference between groups. In sum, DoS and DoN adults recruit largely the same brain regions to process mental state stories (**Figure 2**).

ROI Analysis (Verbal): Based on the whole brain analysis and previous work, six functional ROIs were identified in each participant, RTPJ, LTPJ, RSTS, PC, dorsal MPFC, and ventral MPFC. In RTPJ, DoS adults exhibited strong selectivity to mental state information, with a stronger response to mental state stories (0.95 ± 0.22) relative to physical stories (0.61 ± 0.22 ; $t(8)=3$, $p=.02$) and to social stories (0.58 ± 0.2 , $t(8)=5.2$, $p<.0001$), and no difference between physical and social stories ($t(8)=02$, $p=.83$).

DoN adults also had a stronger response to mental state stories (0.3 ± 0.09) relative to physical stories (-0.11 ± 0.12 ; $t(8)=3$, $p=.02$) in the RTPJ. However, DoN showed no selectivity for mental state stories relative to social stories (0.39 ± 0.13 , $t(10)=1$, $p=.35$), with a stronger response to social stories relative to physical stories ($t(10)=5.6$, $p<.001$).

Comparing groups, DoN showed significantly less selectivity than DoS in RTPJ, with a significant main effect of group, such that DoS adults have overall higher responses to all conditions ($F(1,18)=6.01$, $p=.02$, $\eta^2=.22$), a main effect of condition ($F(2,36)=13.14$, $p<.001$, $\eta^2=.1$) and, critically, a significant interaction ($F(2,36)=7.19$, $p<.001$, $\eta^2=.06$; **Figure 3A**).

No other region (LTPJ, RSTS, PC, DMPFC, VMPFC) showed any effect of group, or any group by condition interactions (all $F<2$, all $p>0.1$).

Verbal Theory of Mind task: language history

Across participants, selectivity (the difference in activation between mental and social stories) correlated with self-reported exposure to ASL during childhood ($r(18)=0.72$, $p<.001$), suggesting that the emergence of strong selectivity in these brain regions is related to language exposure in early childhood (**Figure 3C**).

Verbal Theory of Mind task: language

To test whether this effect is related to online language processing, we examined whether regions typically implicated in language processing, such as left inferior frontal gyrus, show differences in selectivity between conditions or groups. We localized the left IFG (Broca's area) in individual subjects, contrasting all stories to nonsigns. In both DoS and DoN adults, this region responded robustly to all three story conditions relative to non-signs, ($F(4,88)=5.44$, $p<.001$, $\eta^2=.06$) with no differences between story conditions ($F(2,32)=0.02$, $p=.98$, $\eta^2=.$), or across groups ($F(1,16)=0.82$, $p=.38$, $\eta^2=.04$), and no interaction ($F(2,32)=0.2$, $p=.82$, $\eta^2=.$, **Figure 3B**).

Nonverbal Theory of Mind task

Finally, we used the RTPJ identified in the verbal task to probe the response to the non-verbal free-viewing task.

We find a main effect of condition ($F(3,30)=15.14$, $p<.001$, $\eta^2=.46$), and a group by condition interaction ($F(3,30)=15.14$, $p<.001$, $\eta^2=.46$), with no main effect of group ($F(1,10)=1.69$, $p=.22$, $\eta^2=.07$).

Follow up t-tests suggest that the RTPJ of DoS adults showed a selective response to mental information

presented non-verbally, with a robust response to mental segments (0.7 ± 0.17) relative to physical (-0.45 ± 0.11 , $t(5)=4.6$, $p<.01$) and both social conditions (relationships: 0.18 ± 0.06 , $t(5)=2.7$, $p<.04$); sensation 0.09 ± 0.09 , $t(5)=4.1$, $p<.01$).

In contrast, the RTPJ of DoN responded only marginally more to non-verbal mental segments (0.1 ± 0.14) than physical segments (-0.19 ± 0.11 ; $t(5)=2.2$, $p<.08$), with no difference between mental segments and social segments (relationships: 0 ± 0.14 , $t(5)=0.8$, $p<.48$; sensation: 0.03 ± 0.11 , $t(5)=0.5$, $p<.62$; **Figure 3D**).

Discussion and Conclusion

In this study, we investigated whether the neural specialization of brain regions associated with theory of mind reasoning is driven by early experience, specifically linguistic experience. We found that, despite matched language use in adulthood, DoN adults show decreased selectivity in the right TPJ, compared to DoS adults.

During the verbal ToM task, the RTPJ in DoS adults showed a highly selective response, with a stronger response to stories about mental states relative to social information. In contrast, the RTPJ in DoN responded equally to mental state information and social information. Across groups, RTPJ selectivity was correlated with individual differences in childhood language-exposure to ASL, suggesting that the emergence of strong selectivity in RTPJ is influenced by amount of language experience in childhood. The selectivity profile we find in DoN adults mirrors the neural pattern found in young, typically developing children (Gweon et al. 2012; Saxe et al. 2009).

We replicated this group difference in selectivity in the non-verbal, free-viewing task. DoS adults also recruited their RTPJ significantly more during segments of the nonverbal task with mental state information than DoN adults. Because this task did not explicitly draw attention to the character's mental states or require mental state inference, this pattern may suggest a difference in spontaneous mentalizing between groups.

Because a key difference between our groups was early language access, it is important to ask whether the differences we observed in the RTPJ between DoS and DoN adults are due to differences in the neural basis of theory of mind, and not to differences in language comprehension. First, the groups were both screened for ASL proficiency, were matched in current ASL use, and reported that they fully understood the stimuli. Second, our groups were matched on the task in both RT and accuracy. Third, we observed no differences between groups and no differences in story type in a region strongly associated with language processing (Left IFG, i.e. BA44/45, "Broca's area"). Finally, we observed differences across groups in the neural response to nonverbal events, as well as verbal stories, suggesting that online language comprehension is not driving these effects.

Despite these neural differences, DoN children catch up in ToM task performance by early adolescence (Schick et al. 2007; Peterson & Wellman 2010). What does it mean if there are profound changes in the selectivity of brain regions that have been causality implicated in ToM reasoning (Young et al. 2010), but no commensurate

difference in behavior? One possibility is that having a fully specialized RTPJ is only one way to get a fully developed ToM, and that DoN adults may have developed full ToM competency using other mechanisms. Another possibility is that DoN adults have subtle deficits, measurable only on “harder” tasks, for example mental state reasoning performed under cognitive load or with speeded judgements, or on ToM measures that do not have a single right answer, such as making moral judgments (Koster-Hale et al. 2013; Cushman 2008; Miller et al. in press; Young & Saxe 2009).

Together, these data suggest that emergence of strong selectivity in these brain regions is influenced by experience in childhood. We take this to be the first evidence that, in brain regions that become highly specialized by adulthood, the degree of specialization is determined by aspects of developmental history.

Acknowledgments

This material is based upon work supported by the the National Science Foundation under Grant 095518 and a National Science Foundation Graduate Research Fellowship, Grant 0645960. Thanks to Amy Wilson and Jenny Lu to their help recruiting and testing, to Saxelab and Kanwisher lab at MIT for helpful feedback, and our participants.

References

Aichhorn, M.M. et al., 2009. Temporo-parietal junction activity in theory-of-mind tasks: falseness, beliefs, or attention. *Journal of Cognitive Neuroscience*, 21(6), pp.1179–1192.

Baker, C.I. et al., 2007. Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *PNAS*, 104(21), pp.9087–9092.

Baker, C.I., Behrmann, M. & Olson, C.R., 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience*, 5(11), pp.1210–1216.

Bruneau, E.G. & Saxe, R., 2012. Distinct roles of the 'shared pain' and 'theory of mind' networks in processing others' emotional suffering. *Neuropsychologia*, 50(2), pp.219–231.

Cox, D.D. & DiCarlo, J.J., 2008. Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? *Journal of Neuroscience*, 28(40), pp.10045–10055.

Cushman, F., 2008. Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), pp.353–380.

Cutting, L.E. et al., 2006. Differential components of sentence comprehension: Beyond single word reading and memory. *NeuroImage*, 29(2), pp.10–10.

de Villiers, P.A., 2005. The Role of Language in Theory-of-Mind Development: What Deaf Children Tell Us. *Why language matters for theory of mind*.

Dodell-Feder, D. et al., 2011. fMRI item analysis in a theory of mind task. *NeuroImage*, 55(2), pp.705–712.

Dufour, N. et al., 2012. What explains variability in brain regions associated with Theory of Mind in a large sample of neurotypical adults and adults with ASD.

Fedorenko, E. et al., 2010. New Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects. *Journal of Neurophysiology*, 104(2), pp.1177–1194.

Figueras-Costa, B. & Harris, P., 2001. Theory of mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education*, 6(2), p.92.

Fodor, J.A., 1983. *The Modularity of Mind*, MIT Press.

Freedman, D.J. et al., 2006. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 16(11), pp.1631–1644.

Friederici, A., Meyer, M. & Cramon, von, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, 75(3), pp.289–300.

Gweon, H. et al., 2012. Theory of Mind Performance in Children Correlates With Functional Specialization of a Brain Region for Thinking About Thoughts. *Child Development*, pp.no–no.

Hagoort, P. et al., 1999. The neural circuitry involved in the reading of German words and pseudowords: A PET study. *Journal of Cognitive Neuroscience*, 11(4), pp.383–398.

Heim, S.S. et al., 2005. The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords. *Brain research Cognitive brain research*, 25(3), pp.12–12.

Humphries, C.C. et al., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, 18(4), pp.665–679.

Koster-Hale, J. et al., 2013. Decoding moral judgments from neural representations of intentions. *PNAS*, pp.1–14.

Meristo, M. et al., 2007. Language access and theory of mind reasoning: Evidence from deaf children in bilingual and oralist environments. *Developmental Psychology*, 43(5), pp.1156–1169.

Miller, Hannikainen & Cushman (in press) Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*. Milligan, K., Astington, J.W. & Dack, L., 2007. Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2), pp.622–646.

Moeller, M. & Schick, B., 2006. Relations between maternal input and theory of mind understanding in deaf children. *Child Development*, 77(3), pp.751–766.

Partly cloudy. Dir. Peter Sohn. Pixar, 2008. DVD.

Perner, J. & Aichhorn, M.M., 2008. Theory of mind, language and the temporoparietal junction mystery. *Trends in Cognitive Sciences*, 12(4), pp.123–126.

Peterson, C.C. & Siegal, M., 1999. Representing Inner Worlds: Theory of Mind in Autistic, Deaf, and Normal Hearing Children. *Psychological Science*, 10(2), pp.126–129.

Peterson, C.C. & Wellman, H.M., 2010. From fancy to reason: scaling deaf and hearing children's understanding of theory of mind and pretence. *British Journal of Developmental Psychology*, 27(Pt 2), pp.297–310.

Saxe, R. & Kanwisher, N., 2003. People thinking about thinking people: The role of the temporo-parietal junction in. *NeuroImage*, 19(4), pp.1835–1842.

Saxe, R. & Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, 17(8), pp.692–699.

Saxe, R. et al., 2009. Brain Regions for Perceiving and Reasoning About Other People in School-Aged Children. *Child Development*, 80(4), pp.1197–1209.

Schick, B. et al., 2007. Language and theory of mind: A study of deaf children. *Child Development*, 78(2), pp.376–396.

Sigala, N. & Logothetis, N.K., 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, 415(6869), pp.318–320.

Woolfe, T. & Want, S., 2002. Signposts to development: Theory of mind in deaf children. *Child Development*, 73(3), pp.768–778.

Young, L.L. & Saxe, R., 2009. Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, 47(10), pp.2065–2072.

Young, L.L. et al., 2010. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), pp.6753–6758.