Can Changes in Inhibitory Control Explain Child-Level Theory of Mind Development?

Elif N. Poyraz (elifnur@psych.rutgers.edu)[1]
Pernille Hemmer (pernille.hemmer@rutgers.edu)[1]
Alan M. Leslie (aleslie@ruccs.rutgers.edu)[1]
Department of Psychology & Center for Cognitive Science, 152 Frelinghuysen Road, Piscataway, NJ 08854 USA[1]

Abstract

A central canon in theory of mind research is that between the ages of three and four a drastic performance difference in children's understanding occurs. However, the reason for the 'three to four shift' has yet to be settled. One account, Theory of Mind Mechanism (ToMM) theory (Leslie, 1994), posits that change in inhibitory power can account for this difference. This is supported by a recent computational implementation of the theory, showing that differences in inhibitory power can account for age differences at an aggregate level (Wang, Hemmer, & Leslie, 2019). However, as Baker et al. (2016) point out, established findings are entirely based on group-aggregated findings, yet computational and developmental processes do not take place in the 'aggregated mind'. What remains largely unexplored is what happens at the level of the individual child. Here we combine the computational implementation of ToMM with data from Baker et al., 2016, who assessed longitudinal developmental change in Theory of Mind performance by repeated testing of individual child over the three-to-four shift period on standard 'Sally and Anne' false belief tasks, to obtain a cumulative record for each child. Specifically, we found that children's age was not directly informative of developmental change in theory of mind reasoning. Instead, the main contributor to theory of mind performance at the individual learner level is inhibitory power.

Keywords: Theory of mind; Computational Model; Developmental; Inhibition; Longitudinal

Introduction

The ability to attribute mental states and goals, or to *infer* other people's mental contents is called theory of mind (ToM) and is considered essential for our social life. Importantly, this ability appears to emerge over early childhood. A robust empirical finding is that performance in the standard false belief task improves dramatically between the ages of three and four (Wellman, Cross, & Watson, 2001). However, the contributing factors underlying performance are still poorly understood, and theories are divergent. One view is that conceptual understanding changes between ages 3 and 4, while another view is that inhibitory (executive) control skills improve over this period (Leslie, German, & Polizzi, 2005). Furthermore, one could easily anticipate that there would be individual differences in the development of ToM performance, but due to the one-trial nature of most empirical studies, individual differences have remained largely unexplored. In this paper, we seek to assess individual differences in inhibitory control both across children and *within the individual child*, using a unique longitudinal dataset in combination with a computational modeling approach.

Over the last few decades, ToM has been a major topic of interest in the cognitive developmental literature. The standard false belief task (i.e., the Sally-Anne task) is the most commonly used measure of ToM (Baron-Cohen, Leslie, & Frith, 1985). In this task, an actor, Sally has a marble and hides it in Location A (e.g., a basket), then leaves the scene. After Sally's departure, another actor, Anne, takes the marble from location A, and hides it in Location B (e.g., a box). Children are then asked: "When Sally comes back, where will she look for her marble?" The robust finding obtained from this paradigm is that the majority of typically developing children older than four are successful at predicting that Sally will look in Location A; correctly attributing the *false belief* to Sally. On the contrary, children younger than four usually fail, and predict that Sally will look in Location B, where the marble currently is, i.e., the *true belief* (Wellman et al., 2001; see Baillargeon, Scott, & He, 2010 for much earlier understanding of false beliefs).

A number of theorists have argued that this difference is due to a *lack* of *conceptual* understanding for false beliefs before the age of four (e.g., Perner, Leekam, & Wimmer, 1987; Gopnik & Astington, 1988). Others have argued for *performance related* explanations. This view proposes that the limited processing resources of younger children mask an underlying intact conceptualization of false beliefs (FB) (Leslie, 1994; Leslie, et al., 2005). Indeed, recent findings suggest that adding *uncertainty* to the child's own belief about the location of the desired object, improves younger children's performance (Setoh, Scott, & Baillargeon, 2016; Grosso et al., 2019; also see Wang & Leslie, 2016 for performance in non-verbal tasks).

The age of a child is a pervasive factor in development but enters into a computational process mediated by some other factor or factors rather than directly. Our model, drawing on Wang, Leslie, & Hemmer (2019) hypothesizes that age is mediated by increasing inhibitory power.

The Role of Inhibition in Belief Desire Reasoning

There has been wide interest in understanding how inhibition interacts with FB performance. One suggestion has been that inhibitory control plays a crucial role in theory of mind development by enabling children to entertain others' mental states (e.g., Carlson, Moses, & Hix, 1998; Carlson & Moses, 2001). Carlson and colleagues reported that performance in inhibitory control tasks was significantly related to ToM performance. One study, by Sabbagh et al. (2006), compared performance of Chinese and US children in inhibitory control and ToM tasks. They found that although Chinese children were ahead of their US counterparts in inhibition, their ToM performance was not significantly different raising the possibility of that some unidentified cultural factor(s) may influence the role of age-mediated inhibitory power. Nevertheless, individual differences in inhibitory control *within* the Chinese sample still predicted later ToM performance. Therefore, we seek to evaluate how children's performance in ToM relates to their inhibitory control skills, both *across* individuals and also *within* an individual *across time*.

Notably, these findings are in agreement with the predictions of a more specified theory, the Theory of Mind Mechanism (ToMM) (Leslie, 1994; Leslie et al., 2005). This theory posits that the ToM mechanism is on-line from infancy onward, and children already have a mature competence that would result in successful theory of mind reasoning. However, the standard task poses certain processing demands that are beyond the executive control resources of the young mind, specifically inhibitory control demands.

ToMM + SP Model In the the Sally-Anne scenario, ToMM will *spontaneously* compute the *true belief* (TB), which is the object's current location, and also the *false belief* (FB), which is based off of a crucial aspect; calculation of the visual/informational access (V) of the person. Then, these two belief candidates will be evaluated. Since beliefs ought to be, and generally are true (Dennett, 1989), the TB will be selected by default. In TB scenarios, this default serves well, and results in correct attributions. However, in a FB scenario, like the Sally-Anne, this prepotent response needs to be *inhibited*, so that the correct belief candidate (i.e., FB) can be selected as a response (i.e., *an action prediction* (A)). The inhibition (I) process is represented by the addition of the Selection Process (SP) to the ToMM (ToMM +SP).

The ToMM + SP model is not only supported by experimental evidence (Friedman & Leslie, 2004; Setoh et al., 2016) but also by a computational approach used to quantitatively estimate the role of inhibition in ToM reasoning (Wang et al., 2019). In their model, Wang and colleagues used behavioral data from both low (LD) and high demand (HD) FB tasks to infer three- and four-year-olds' inhibitory power. Their model captures inhibition's role to account for the 3 to 4 'shift', rather than a conceptual shift in understanding false beliefs. The current study will employ the same model as Wang et al. (2019). See the Modeling section for more details.

Contribution of Longitudinal Data

A characteristic of child data on ToM using the Sally-Anne task is that in a majority of studies each child only completes one trial. Thus, all analysis is at an aggregate level across children, obscuring individual differences and only allowing for a crude age assumption (with up to a year's age difference between children assumed to be of the age group). This has substantially limited theorists in assessing individual performance over time. Since developmental change does not take place in the 'aggregated brain', change should be sought out in individual-level data. Baker et al. (2016) sought to resolve this problem by first collecting longitudinal data by following preschoolers between the ages of 3 to 5, and then implementing a novel Bayesian change-point analysis. Most notably, they found that the generally assumed 'sudden insight' shift in ToM understanding was only a minority among different patterns observed (9.6%), with the majority of children showing unstable records (44%).

Their findings lay bare the fact that age itself is not directly predictive of ToM performance. What remains unanswered is what underlies the developmental change in ToM reasoning. The current study seeks to elucidate this question by combining the model developed by Wang et al. (2019) with the longitudinal data in Baker et al. (2016). We explore how the ToMM+SP model can be helpful not only in understanding the 3 to 4 'shift' (ala Wang et al.), but more specifically how fluctuations in individual performance records over a temporal window could potentially be explained by changes in inhibitory power estimates, rather than by time (i.e., age) alone.

Modeling

The Data Set

The data set for this model comes from Baker et al. (2016)'s longitudinal study, in which an extensive record of each individual child was reported. The main aim in their study was to test preschool children in the transitional period of theory of mind performance, in repeated trials over a course of unrestricted time (i.e., as long as possible).

The final data set included 52 children from the US and UK. The preschoolers' age at first testing session ranged from 34 months to 62 months, with an average of 47 months (*sd*=6.3 months). Each testing session had no more than six weeks in between and occurred approximately monthly (for more details about data collection and analysis see Baker et al. 2016). The number of trials completed ranged from 10 to 36, with an average of 21 testing sessions (*sd*=7.6).

Subject Level Model

Following convention, we adopt graphic models as a way to visualize structure. In the graphic model, circular nodes represent all variables of interest. Nodes for observed variables are shaded while nodes for latent variables are unshaded. Further, stochastic variables are differentiated with a single border and deterministic variables with double borders. Arrows indicate conditional dependencies.

We implemented two models. Figure 1 gives the graphical model for the 'subject level' model, which assumes that inhibition (I) is fixed for each individual child—That is to say that each child's inhibitory power is inferred from that child's aggregate performance. Figure 4 gives the 'temporal subject level' model, which assumes that I for each child is allowed to vary over a temporal window.

In the standard false-belief task, Sally wants her toy back,

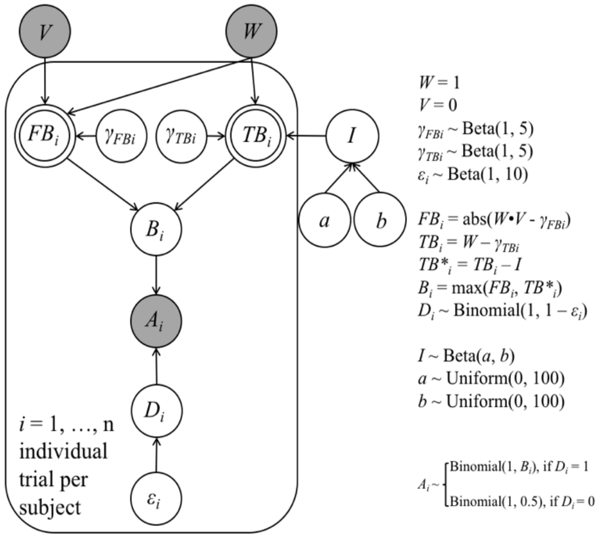

Figure 1 The graphical model for ToMM+SP at the subject level. *W*: world, *FB*: false belief; *TB*: true belief; *D*: desire; *A*: action; *V*: visual access; *I*: inhibition. This 'subject level' model assumes that inhibition (I) is fixed for each individual child but varies between children.

and will go to the location where she believes the toy to be. Thus, Desire (D) is assumed to have a large prior probability, $1 - \varepsilon$. Following Wang et al., we *1)* use an asymmetric beta prior on ε, with $\varepsilon \sim$ Beta (1, 10), indicating that Sally is unlikely to change her desire (this is also the assumption used by Goodman et al., 2006); *2)* assume the TB content one attributes to Sally is determined by one's own knowledge of the world with probability $W - \gamma$TB (W = 1 is change in the world or 0 is no change), and is the variations in how one may incorrectly represent the world; *3)* assume the FB content is determined by Sally's visual access to the world (V=0, or no access, in a false-belief scenario), with the uncertainty about FB given by the absolute value of ($W \bullet V$ - γFB). Following Goodman et al., we used asymmetric beta priors on γTB and γFB, with $\gamma TB \sim$ Beta (1, 5) and $\gamma FB \sim$ Beta (1, 5), indicating that one has an accurate representation of the world and others visual access.

Under the assumption of ToMM + SP, to succeed in the standard false-belief task, one has to deploy sufficient inhibition to overcome the TB prior. Following Wang et al., $I \sim$ Beta (α, β), where $\alpha \sim$ Uniform (1,100) and $\beta \sim$ Uniform (1,100), indicating no prior assumptions about the shape of the beta distribution of I and to allow for a complete estimation of I from the observed data.

The attributed belief (B) is the stronger belief between FB and TB with inhibition, i.e., the probability of B is the larger of the probability of the true belief with an inhibition (W - γTB - I) and the probability of the FB (absolute value of $W \bullet V$ - γFB). After the selection, $B \sim$ Binomial (1,max (FB, TB-I)), one attributes either the belief that there is no change (B = 0) or the belief that there is a change (B = 1) to the agent, which coupled with desire determines action prediction (A). If the attributed desire is to retrieve the object (D = 1), prediction follows the attributed belief (B) and $A \sim$ Binomial (1, B); if the attributed desire is ambiguous (D = 0), action prediction is random and $A \sim$ Binomial (1, 0.5). See Wang et al. for the conditional probabilities amongst all variables.

In sum, A is conditioned on the B, and D. B is determined by selecting between the *TB* and the *FB* through applying I. Having multiple trials for each subject's A, we can now reverse the causal chain to infer to the cognitive parameter I that produced such performance for a given subject. This inference is crucial in terms of understanding, for the first time, individual differences in inhibitory power, and how these differences in I explain ToM performance.

The models were implemented using WinBUGS (Lunn et al., 2000; Ntzoufras, 2009). Our results are based on drawing 1,000 samples from two separate chains with a 100 burn-in period for each of the models. Convergence of the chains was assessed using the $\hat{R}$ statistic (Brooks and Gelman, 1998).

Subject Level Model Results Figure 2 shows the results for the Subject Level Model. In each graph, darker shades indicate older children. Because performance is bounded at 0 and 1 the distributions are heavily skewed, therefore each dot represents the mode of a given measure. In Panel A, subject's overall performance (calculated as proportion correct across trials) is plotted as a function of age in months at the start of testing. There was a moderate correlation between the age at first test, and overall success $r(50) = .36$, $p = .007$. As it can be seen in the figure, performance is quite scattered as a function of age. Similarly, in Panel C the mode of the posterior distribution of the inhibition parameter for each subject is plotted as a function of age at the start of testing. There was a moderate correlation between age and I, $r(50) = .4$, $p = .003$.

Panel B can be thought of as a combination of Panel A and C, showing overall success as a function of inhibition. This shows how subjects' ToM performance relates to estimated inhibitory power, $r(50) = .98$, $p < .001$, suggesting that the effect of age is mediated by inhibitory power. Indeed, differently (age) shaded nodes are scattered across the fitted line, indicating that some young children perform better than many older children, and this relationship can be captured by inhibition.

In Panel D, the standard error (unbiased estimate of variability) is plotted as a function of overall performance by subject. This is to demonstrate that variance in a child's trials alone is not informative, as the child that performed at the ceiling (100% correct), and the child performed at the floor (100% incorrect) produced the exact same variance, namely *0*. This is due to the dichotomous nature of data that the standard FB task yields, and the traditional inferential statistics based on averaged data alone cannot describe the richness of the longitudinal data used here.

In Panel E, the standard error for I is plotted as a function of overall performance by subject. This demonstrates that variation in inhibition follows a non-linear pattern similar to

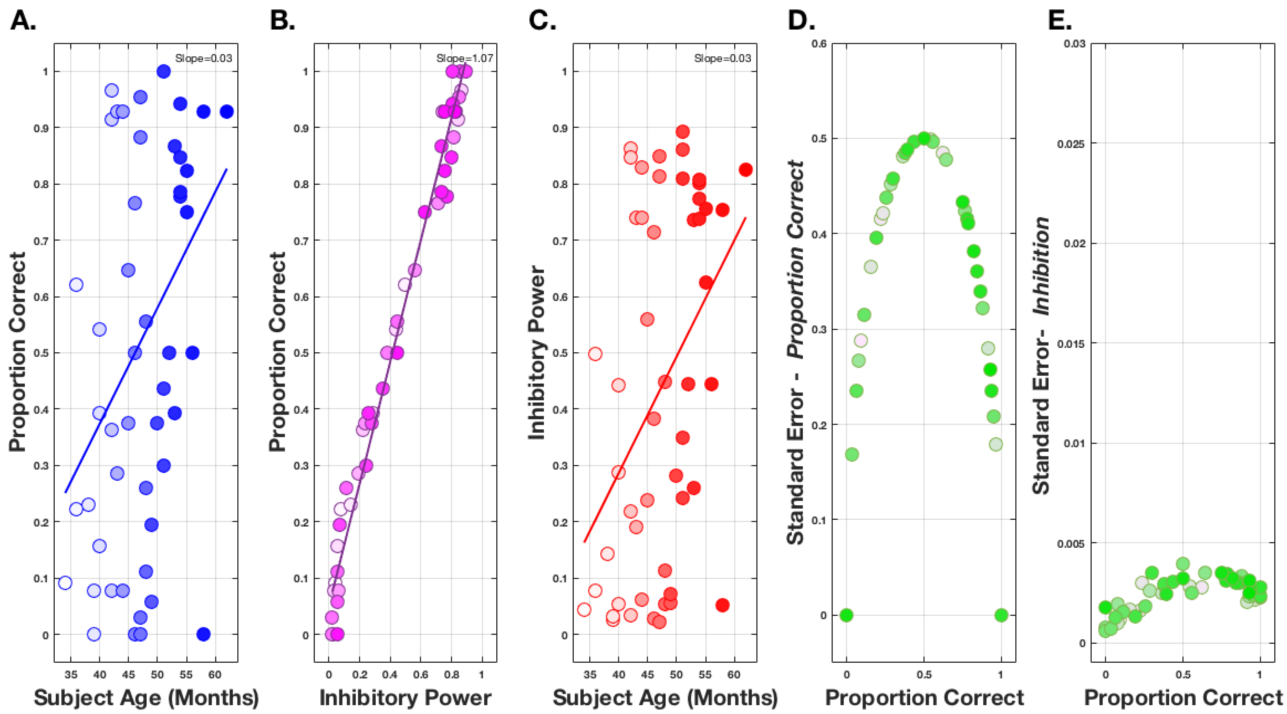

Figure 2 Each dot represents an individual child's value. Darker shades mean older age at the start of testing. Lines in A, B, C indicate the best fitting line, with slope given in top right. **A.** y axis: proportion correct across all observed trials per child, x axis: children's age at the start of testing. **B.** y axis: proportion correct across all observed trials per child, x axis: Mode of the Posterior distribution of inhibition for each child. **C.** y axis: Mode of the Posterior distribution of inhibition for each child, x axis: children's age at the start of testing. **D.** Standard Error values for each subject's performance. Y axis rescaled for visual clarity. **E.** Standard Errors for Inhibition. Y axis rescaled for visual clarity.

that of proportion correct. For very low performance variability is also low – suggesting that when a child fails, they do so consistently. Variation then increases across accuracy, such that as inhibition comes on line performance gets more unstable. Finally, for higher accuracy variation again decreases, such that once a child has inhibitory power performance becomes more stable.

Figure 3 shows the mode posterior predictives for action prediction simulated by the model. This simulation can be thought as: if the same subject were to undergo the same number of trials, using the parameters inferred from the observed data, what would be the expected outcome in terms of their overall success. The mode of the posterior predictive distributions for each subject is indicated by the red shaded dots. The blue shaded dots indicate the observed proportion correct for each subject. It can be easily observed that the model predictions follow a similar trend to that of observed data, with slight underestimates for children with higher inhibitory power, and slight overestimates for children with lower inhibitory power.

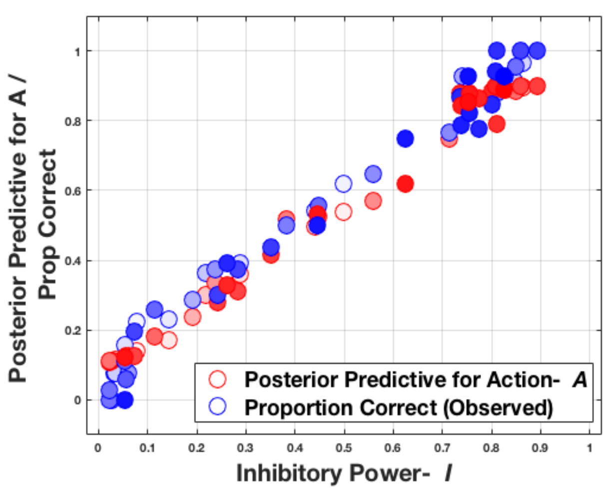

Figure 3 Each dot represents an individual child's value—lighter shades indicate younger age. Red shades represent the simulated Action prediction values by the model based on the inferred *I* estimates. Blue shades represent the proportion correct for the observed data.

Temporal Subject Level Model

Inplementing the subject level model we found only a moderate relationship between age and ToM performance (suggesting that age is not directly informative), but captured how individual differences in inhibitory power can account for ToM performance, regardless of subject age. Next, we sought to evaluate development within individual children—

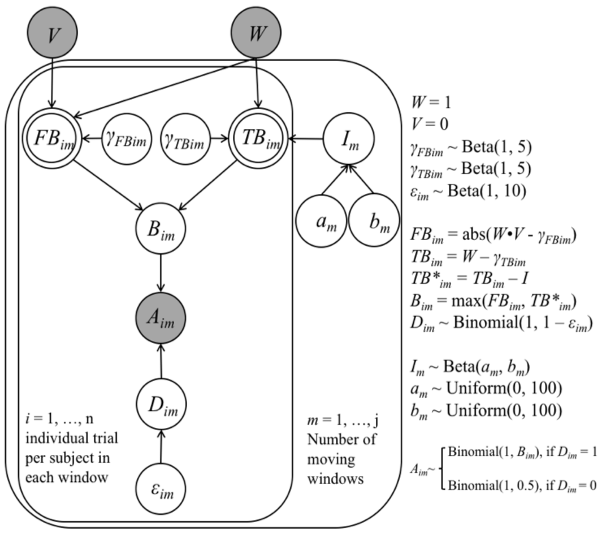

Figure 4 The graphical model for implementing moving windows. This temporal subject level model assumes that I for each child is allowed to vary over a moving temporal window.

namely, how does a child's inhibitory control change *over time*?

It would be expected that, like any other cognitive process mechanism (i.e., executive contol) that mature with age, inhibition should follow a similar trend.We did not find evidence of a 3 to 4 'shift' in the aggregate child, assuming fixed inhibition. However, it is still possible that inhibitory control develops over the testing period of the individual

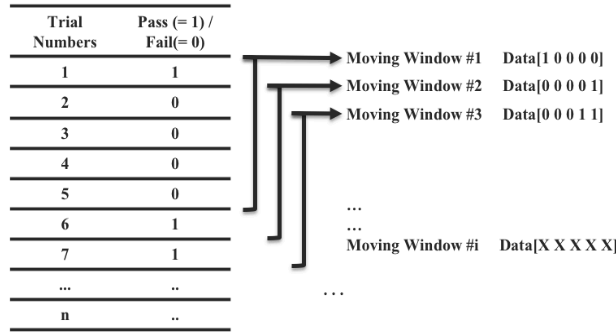

Figure 5 Illustration of how data was selected by the use of a moving window of 5 trials.

child. The next model, the 'temporal subject level' model, aims to capture this change within each individual subject.

Figure 4 shows the graphical model for the 'temporal subject level' model, which assumes that I for each child is allowed to vary over a temporal window. Using the same data set, we implement a 'moving window' that iterates over a set number of trials per subject. A depiction of how this window was implemented on data can be found in Figure 5. This window size was set to five for the current model, indicating that for each inference, the model received five data points for A. Because this is a first attempt to analyze inhibitory control's change over time, the criterion for the least amount of trials that we could expect a change to be manifested was largely unknown. Since each trial was separated, on average, by a month, having the window size set to five seemed like a reasonable number that could capture change.

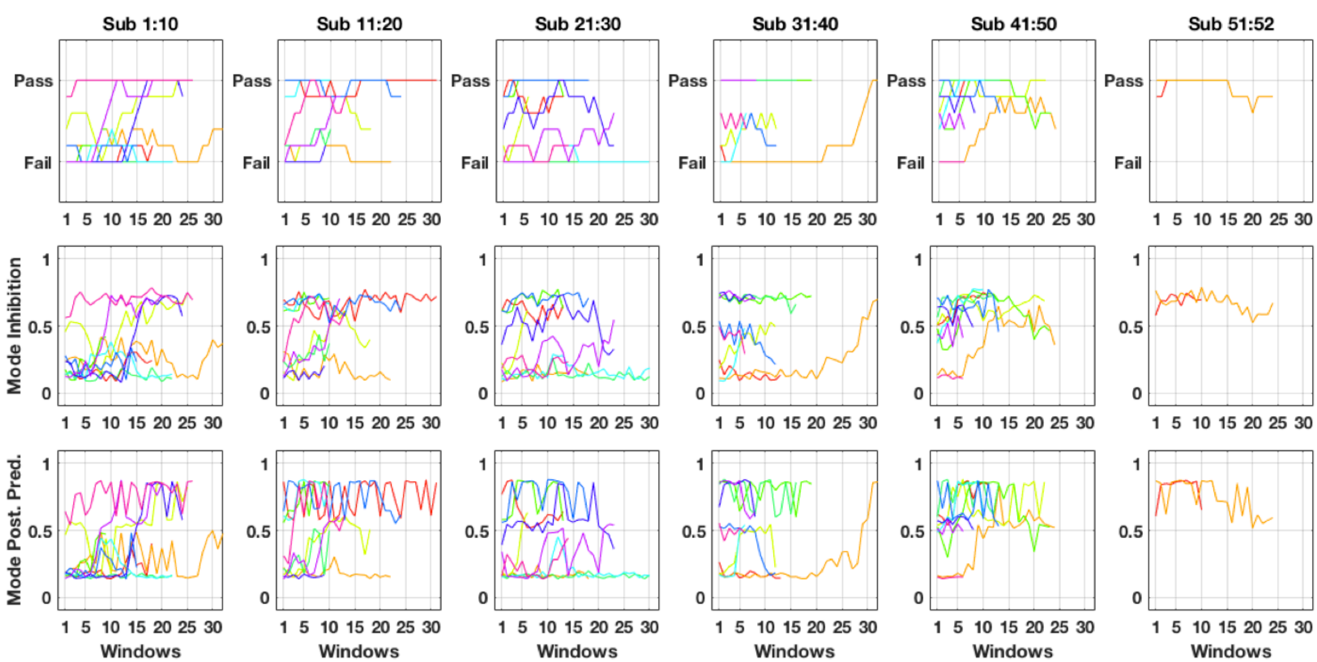

Figure 6 Each subplot has 10 subjects (except the last column). Age increases column-wise across plots. The same color across plots in each column is for the same subject. **Top Panel** – Shows the raw performance across windows. **Middle Panel** – Shows the mode inhibitory estimates across windows. **Bottom Panel** – Shows the mode action prediction.

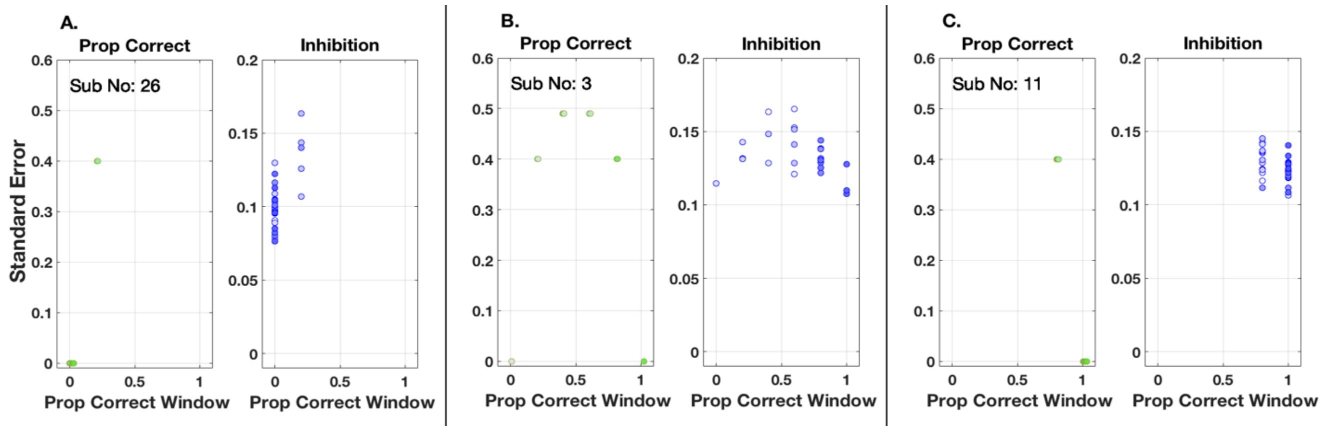

Figure 7 Standard Error values for proportion correct and inhibition for three representative children. Darker shades mean later testing sessions. X axis: Proportion correct for a given moving window.

Temporal Subject Level Model Results Figure 6 shows the results for the Temporal Subject Level Model assuming that *I* for each child is allowed to vary over a moving temporal window. The number of windows differed across subjects, as a function of number of trials (the number of windows = number of trials - window size + 1; where window size is set to 5). For the ease of examination and comparison, in each Panel lines represent 10 different subjects. From left to right, the age at first testing increases, such that the first Panel on the left shows the 10 youngest subjects, and the last Panel on the right is the 2 oldest subjects. The top row of Panels gives each subject's success rate (proportion correct) for each moving window. By examining each line, children's performance trajectory across windows can be followed. Crucially, performance across time is variable within a single subject as well as between subjects, pointing to a 'stably unstable' performance (Baker et al., 2016). The variant trajectories in FB performance again show how a crude age assumption could be masking individual differences both in inhibitory control and in ToM development.

Crucial to the current study, the middle row in Figure 6 depicts the inferred inhibition estimates. The structure of the graphs is the same as the top row, such that each line represents an individual child's trajectory. There are a few important aspects of the inferred inhibition estimates across participants. 1) Individual trajectories are quite variant, with occasional increases and decreases within each individual, 2) individual children vary over time, and 3) some children show gradual increase in inhibitory power over time. Although these children are aging across trials, for the majority of the children neither their performance in the FB task, nor their inferred inhibitory power follows a stably increasing pattern. Figure 6, bottom row shows the posterior predictives for action prediction. Notably, as in the 'subject level' model, action prediction follows inhibition, even across the multiple trials of an individual child.

In Figure 7 the standard error for proportion correct and I is plotted as a function of proportion correct for a given moving window, for three representative children. Panel A is representative of a group of 17 children who showed poor performance even on later testing sessions (as indicated by darker colors), while exhibiting relatively low error on I—meaning that children in this group failed consistently. Panel B represents a group of 12 children who showed improving performance over testing sessions. The inverted U-shape in both graphs illustrate the same pattern of variance and performance as seen in Figure 2 Panels D and E. Finally, the child in Panel C is representative of a group of 15 children who consistently demonstrated high accuracy.

Discussion and General Conclusion

Previously, inhibition's role in ToM reasoning was captured with group averaged data (Wang et al., 2019), accounting for the differences in performance between three- and four-year-olds. Following the same Bayesian modeling approach, *without* any added complexity to the model, we demonstrated how individual differences in inhibitory power can account for the ToM performance, rather than subjective age.

Specifically, we used a novel and rich dataset. This enabled us to capture, first, how performance in ToM reasoning differs among individual children as a function of inhibition, and second, even much more importantly, how differences within a child can directly predict ToM performance over time. We observed that when the overall success is viewed as a function of age, there is only so much that this relationship can explain. The moderate relationship between age and performance is itself not directly predictive of ToM performance, but is instead mediated through the ability to exert executive control (i.e., inhibitory power). Moreover, the cognitive parameter *I* is strongly predictive of performance both within and across individuals. This is in line with the predictions of ToMM+SP model (Leslie et al., 2005), which proposes that the traditional FB tasks require executive control skills, like inhibitory control, to be successful in 'passing' the task, and this is a compelling explanatory factor in preschooler performance, rather than an account that proposes only a lack of conceptual understanding for FB in younger children (e.g. Perner et al., 1987).

We modeled a unique dataset comprising extensive cumulative records of single children in a longitudinal change point study reported in Baker, Leslie, Gallistel, and Hood,

2016. As a first step, we examined how much Inhibitory power varied across individuals in this dataset. A substantial difference between the modeling in Wang et al. (2019), and ours is the nature of this data. In in the Subject Level model we modeled each child separately, allowing us to *infer* different *inhibitory power* values per case.

Further tests of the current model should include longitudinal findings in *i*) low inhibitory demand versions of the FB task, and *ii*) an independent inhibitory control measure. Due to the unavalability of such data, it was not possible to test the current model to observe how well it predicts performance for an individual across time with a low inhibitory demand version. However, although still using aggregate data coming from three and four-year-olds, Wang et al. (2019) showed that the same model was also successful in accounting for, and predicting, performance in the low inhibitory demand version of the standard FB task for a given age group, namely three and four. After future longitudinal research is accomplished, testing the current model with another rich dataset would greatly help in characterizing the individual development trajectories in ToM development, and how inhibitory control plays a role in different versions of the standard task.

The work presented here is a first ever computational analysis of individual differences in children's theory of mind development. Our unique modeling approach has allowed us to infer coginitive parameters central to ToM development, and predict new individual performance trajectories. The combination of longitudinal single case data and computational modeling shows how much our understanding of the true nature of ToM development, and 'development' in general can be expanded, and how such methodology can shed light upon unanswered questions in cognitive developmental science.

References

Baillargeon, R., Scott, R.M. and He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110–118.

Baker, S. T., Leslie, A. M., Gallistel, C., & Hood, B. M. (2016). Bayesian change-point analysis reveals developmental change in a classic theory of mind task. *Cognitive Psychology*, *91*, 124-149.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37-46.

Brooks, S., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434-455.

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development*, *72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The Role of Inhibitory Processes in Young Children's Difficulties with Deception and False Belief. *Child Development*, *69*, 672-691.

Dennet, D. (1989). *The intentional stance.* Cambridge: the MIT press.

Friedman, O., & Leslie, A. (2004). Mechanisms of belief-desire reasoning: inhibition and bias. *Psychological Science*, *15*, 547-552.

Goodman, N. D., Baker, C. L., Bonawitz, E. B., Mansinghka, V. K., Gopnik, A., Wellman, H., & Tenenbaum, J. B. (2006). Intuitive theories of mind: A rational approach to false belief. *Proceedings of the twenty-eighth annual conference of the cognitive science society.*

Gopnik, A., & Astington, J. (1988). Children's understanding of representational change and its relation to the understanding of false belief and appearance-reality distinction. *Child Development*, *59*, 26-37.

Grosso, S., Schuwerk, T., Kaltefleiter, L., & Sodian, B. (2019). 33-month-old children succeed in a false belief task with reduced processing demands: A replication of Setoh et al. (2016). *Infant Behavior and Development*, *54*, 151-155.

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*, 211-238.

Leslie, A., German, T., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, *50*, 45-85.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*, *10*, 325-337.

Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS.* . New York: Wiley.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125-137.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Kang, L. (2006). The Development of Executive Functioning and Theory of Mind: A Comparison of Chinese and U.S. Preschoolers. *Psychological Science*, *17*, 74-81.

Setoh, P., Scott, R., & Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false belief task with reduced processing demands. *PNAS*, *113*, 13360-13365.

Wang, L., & Leslie, A. (2016). Is Implicit Theory of Mind the 'Real Deal'? The Own-Belief/True-Belief Default in Adults and Young Preschoolers. *Mind & Language*, *31*, 147–176.

Wang, L., Leslie, A. M., & Hemmer, P. (2019). A Bayesian framework for the development of belief-desire reasoning: Estimating inhibitory power. *Psychonomic Bulletin & Review*, 26, 205-221.

Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about the false belief. *Child Development*, *72*, 655-684.