CHILD DEVELOPMENT

Child Development, xxxx 2014, Volume 00, Number 0, Pages 1–8

Why We Think We Can't Dance: Theory of Mind and Children's Desire to Perform

Lan Nguyen Chaplin
University of Illinois at Chicago

Michael I. Norton
Harvard Business School

Theory of mind (ToM) allows children to achieve success in the social world by understanding others' minds. A study with 3- to 12-year-olds, however, demonstrates that gains in ToM are linked to decreases in children's desire to engage in performative behaviors associated with health and well-being, such as singing and dancing. One hundred and fifty-nine middle-class children from diverse backgrounds in a Northeastern U.S. metropolitan area completed the study in 2011. The development of ToM is associated with decreases in self-esteem, which in turn predicts decreases in children's willingness to perform. This shift away from performance begins at age 4 (when ToM begins to develop), years before children enter puberty.

Chaperoning a middle school dance—with girls and boys slouched against bleachers, refusing to dance despite booming music—inevitably leads adults to comment: "Why aren't they dancing?" This refusal is particularly notable because these same children, just a few years earlier, were prone to dance, sing, and more generally *perform* constantly—in school, at home, in the backseat of the car, and while watching television—with both relish and confidence in their abilities. Why do people lose this willingness to perform with age? We suggest that one crucial reason is the development of children's awareness that their peers may be critically evaluating their abilities—an offshoot of the ability to understand others' minds—which may decrease their self-esteem and desire to perform. We explored the link between children's ability to understand the minds of observers and their willingness to perform by giving them the choice to sing, dance, or avoid both activities.

Theory of mind (ToM) is generally viewed as a positive development, typically beginning around age 4, with sharp increases between ages 5 and 6 and further development throughout school-age years (Baron-Cohen, Leslie, & Frith, 1985; Perner & Wimmer, 1985; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). The overwhelming consensus is that ToM allows children to achieve success in the social world by interpreting human behavior and understanding cultural meanings and social norms (Bruner & Feldman, 1993; Gauvain, 1998), such that individuals with deficits in ToM have difficulty in social interaction and in determining the intentions of others (Baker, 2003; Frith, Happé, & Siddons, 1994).

ToM is a fundamental, upstream cognitive construct that influences a number of downstream variables—including perspective taking, empathy, and mentalizing more generally—which in turn influence conversation skills, social competence, and communication effectiveness (Begeer, Malle, Nieuwland, & Keysar, 2010; Gallagher & Frith, 2003). While ToM is often conflated with the psychological concepts of empathy and perspective taking, researchers have posited that ToM may be a precursor to these abilities (Howlin, Baron-Cohen, & Hadwin, 1999; Waytz, Gray, Epley, & Wegner, 2010); indeed, Malle (2005) suggests that ToM is fundamental to all aspects of social cognition. Why would increases in ToM—in understanding the minds of others—be linked to a decrease in the desire to perform? Children's concern with garnering desirable social evaluations—and experience of related emotions such as embarrassment—develop

This research was funded by the University of Illinois at Chicago, Villanova School of Business, and Harvard Business School. We are grateful to the children and staff at summer camps in a large Northeastern metropolitan area for their participation. We thank Ayalla Ruvio and Nicholas Sommer for their assistance. We also thank Anne Pick and Kristina Olson for their valuable comments on an earlier version of this manuscript.

Correspondence concerning this article should be addressed to Lan Nguyen Chaplin, Department of Managerial Studies, College of Business Administration, University of Illinois at Chicago, 601 S. Morgan Street, 2206 University Hall, Chicago, IL 60607. Electronic mail may be sent to nguyenl@uic.edu.

 0009-3920/2014/xxxx-xxxx
DOI: 10.1111/cdev.12314

as early as 4 years old, with further development particularly between ages 6 and 11 (Banerjee, Bennett, & Luke, 2012; Banerjee & Lintern, 2000; Watling & Banerjee, 2012). Importantly for our account, ToM is linked to increasing sensitivity to criticism (Cutting & Dunn, 2002; Dunn, 1995); Lecce, Caputi, and Hughes (2011), for example, show that sensitivity to criticism mediates the relation between ToM and academic achievement.

As a result, we hypothesized that improved ToM would predict older children's decreased willingness to perform due to the heightened sensitivity to criticism—and resulting blow to self-esteem—that ToM engenders.

Method

One hundred fifty-nine middle-class children (81 female; 40 African Americans; 29 Asian Americans; 60 non-Hispanic Whites; 15 Hispanic or Latino Americans; and 15 Other) aged 3–12 from eight summer camps in a large Northeastern metropolitan area participated individually in the experiment in 2011. Age groups ranged in size from 12 to 18 children. Note that the gender composition of our sample did not vary by age, $\chi^2(9) = 1.34$, $p = .99$, Cohen's $d = .18$, and gender was not correlated with ToM, self-esteem, or performance choices (rs < .08, ps > .32; Table 1).

Children were interviewed one by one in an unused classroom and were told by the experimenter, "I'm here today to learn more about you and other kids your age." Each child completed a preference task and a ToM task (order counterbalanced). In the preference task, the experimenter presented each child with four options in random order: sing a song of their choosing (performance task), perform a dance of their choosing (performance task; both performance tasks were without music), circle red shapes on a page (nonperformance task), or color in a square (nonperformance task). Children selected two of the options to complete in front of the experimenter. We chose singing and dancing as prototypical performance behaviors that are also subject to scrutiny by peers.

We assessed ToM with three measures intended to capture different aspects and levels of ToM: the Sally and Anne false belief task (Baron-Cohen et al., 1985), the Cookie Box misleading container test (Gopnik & Astington, 1988), and the Duck and Lion social test (Nguyen & Frye, 1999), the latter designed to test a more mature ToM (see Appendix A). Given our interest in the development of ToM across a wide age range (ages 3–12), we chose these measures because they are not only appropriate for use with very young children but also resonate with older children; similar false belief measures have been used with children ages 6–11 (Apperly, Warren, Andrews, Grant, & Todd, 2011; Pellicano, 2010). For each task, participants were required to pass control questions and the test question to score 1 point. The results of each task were summed for a composite measure (range = 0–3, $\alpha = .84$).

Finally, children completed a seven-item self-esteem scale ($\alpha = .95$) using a 5-option "smiley face" scale (see Appendix B). We adapted items from Harter's (1982) Perceived Competence Scale for Children, Harter and Pike's (1984) Perceived Competence and Social Acceptance Scale, and Harter's (1985) Self Perception Profile. Children rated themselves on several domains crucial to childhood self-esteem including cognitive/scholastic compe-

Table 1
Descriptive Statistics and Correlations

	M (*SD*)	1	2	3	4	5	6	7	8	9	10
1. Gender		1.00									
2. Age	7.50 (2.90)	0.00	1.00								
3. Theory of mind	1.92 (1.26)	0.03	0.69**	1.00							
4. Self-esteem	4.35 (0.70)	−0.08	−0.40**	−0.44**	1.00						
5. Picked sing	0.39 (0.49)	−0.09	−0.38**	−0.46**	0.56**	1.00					
6. Picked dance	0.33 (0.47)	0.01	−0.30**	−0.46**	0.34**	−0.12	1.00				
7. Picked sing and dance	0.72 (0.64)	0.01	−0.51**	−0.69**	0.68**	0.68**	0.65**	1.00			
8. Picked circle	0.59 (0.49)	−0.02	0.30**	0.50**	−0.40**	−0.31**	−0.62**	−0.69**	1.00		
9. Picked color	0.69 (0.46)	0.11	0.38**	0.42**	−0.51**	−0.61**	−0.23**	−0.64**	−0.11	1.00	
10. Picked circle and color	1.28 (0.64)	0.06	0.51**	0.69**	−0.68**	−0.68**	−0.64**	−1.00**	0.69**	0.64**	1.00

Note. Numbers represent zero-order correlation coefficients.
**$p < .01$.

tence, social acceptance, physical/athletic competence, physical appearance, and behavioral conduct (e.g., Harter, 1985; Harter & Pike, 1984). Because these measures were initially developed for older children, we checked to ensure that the reliability of the self-esteem scale for our youngest participants (ages 3–5) was acceptable ($\alpha = .89$).

Results

As predicted, age was negatively related to choosing to sing ($\beta = -.38$, $p < .001$, 95% CI [−.40, −.35]) and choosing to dance ($\beta = -.30$, $p < .001$, 95% CI [−.32, −.27]; Figure 1). The percentage of children choosing singing ranged from 75% and 63% (ages 3 and 4, respectively) to 20% and 6% (ages 11 and 12, respectively); similarly, the percent choosing dancing ranged from 50% (ages 3 and 4) to 13% and 12% (ages 11 and 12; Table 2). Conversely, age was positively related to choosing circling ($\beta = .30$, $p < .001$, 95% CI [.27, .33]) and coloring ($\beta = .38$, $p < .001$, 95% CI [.36, .40]). The percentage of children choosing circling ranged from 38% (ages 3 and 4) to 80% and 82% (ages 11 and 12); the percentage of children choosing coloring ranged from 38% to 50% (ages 3 and 4) to 87% and 100% (ages 11 and 12).

Put another way, while 31.2% of 3-year-olds and 18.8% of 4-year-olds chose to *both* sing and dance, not a single child aged 11 or 12 did so. In contrast, only 6.2% of both 3- and 4-year-olds chose to avoid both singing and dancing, compared to 66.7% of 11-year-olds and 82.4% of 12-year-olds. Nor is it the case that older children generally prefer circling and coloring. We offered six circling and coloring tasks to a separate sample ($N = 34$, aged 3–4 or 11–12) as they waited for another experiment. Younger children spontaneously completed more tasks than older children (Ms = 2.21 vs. 1.00), $t(32) = 3.80$, $p < .01$, Cohen's $d = 1.34$, suggesting that older children's desire to complete such tasks in our main experiment is not due to their intrinsically greater liking than younger children. Order of options did not influence our results.

Age was positively related to ToM ($\beta = .69$, $p < .001$, 95% CI [.64, .74]) with low scores at ages 3 and 4 (Ms = .25 and .56) and high scores at ages 11 and 12 (Ms = 2.80 and 2.76), with a particular inflection point between ages 5 and 6 (Figure 1). Finally, age was negatively related to self-esteem ($\beta = -.40$, $p < .001$, 95% CI [−.44, −.36]) and self-esteem was negatively related to ToM ($\beta = -.44$, $p < .001$, 95% CI [.69, −.19]).

We used bootstrapping procedures to confirm the (independent) mediating roles of ToM and self-esteem on the relation between age and performance (choosing to sing and dance). These procedures, preferable for smaller samples, generate

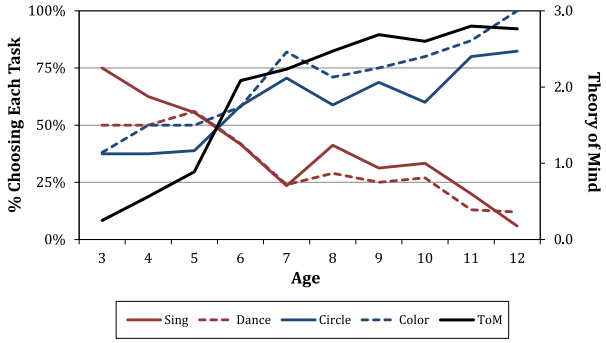


Figure 1. Percentage choosing each task (left axis) and theory of mind (right axis) as a function of age. Theory of mind develops with age (black line); younger children prefer performance tasks (singing and dancing—red lines); older children prefer nonperformance tasks (circling and coloring—blue lines).

Table 2
Theory of Mind, Self-Esteem, and Percentage of Children Selecting Different Tasks as a Function of Age

Age (years)	Theory of mind (*SD*)	Self-esteem (*SD*)	% Sing	% Dance	% Circle	% Color
3	0.25 (0.78)	4.77 (0.45)	75	50	38	38
4	0.56 (0.81)	4.82 (0.24)	63	50	38	50
5	0.89 (1.13)	4.61 (0.50)	56	56	39	50
6	2.08 (1.24)	4.44 (0.58)	42	42	58	58
7	2.24 (1.03)	4.16 (0.72)	24	24	71	82
8	2.47 (0.87)	4.24 (0.80)	41	29	59	71
9	2.69 (0.79)	4.27 (0.65)	31	25	69	75
10	2.60 (0.63)	4.38 (0.65)	33	27	60	80
11	2.80 (0.41)	4.13 (0.92)	20	13	80	87
12	2.76 (0.56)	3.76 (0.67)	6	12	82	100

a 95% confidence interval around the indirect effect where mediation is said to occur if zero falls outside that confidence interval (Preacher & Hayes, 2004, 2008). ToM had a significant mediating effect on the relationship between age and performance (β = −.10, 95% CI [−.13, −.07]), also confirmed by a Sobel test ($z = -6.71$, $p < .01$). Bootstrapping results also showed that self-esteem had a significant mediating effect on the relation between age and performance (β = −.05, 95% CI [−.07, −.03]), also confirmed by a Sobel test ($z = -4.66$, $p < .01$). We note that these results are consistent when we restrict our analyses only to older children (ages 9–12), for whom ToM scores were near ceiling: ToM remains a significant mediator between age and performance (95% CI [−.2103, −.1205]; $z = -6.15$, $p < .01$).

Next, we tested our more nuanced conceptual account—that increases in ToM lead to decreases in self-esteem that affect children's desire to perform—using structural equation modeling. We used Amos 4 software (Byrne, 2001) to test multiple-step, multiple-mediator path models (Hayes, 2009; Iacobucci, Saldanha, & Deng, 2007). The tests of mediation included comparisons of full mediation, nonmediation, and partial mediation models of the relations between ToM, self-esteem, and performance—a total of four models. The full mediation model (Model 1; no relation between ToM and performance) resulted in a poor fit to the data ($\chi^2 = 71.72$, $df = 3$, $p = .00$, comparative fit index [CFI] = .767, normed fit index [NFI] = .765, incremental fit index [IFI] = .772, root mean square error of approximation [RMSEA] = .38, 90% CI [.31, .46]). The nonmediation model (Model 2; only direct relations between ToM and self-esteem on performance) demonstrated a better fit to the data ($\chi^2 = 37.85$, $df = 3$, $p = .00$, CFI = .882, NFI = .876, IFI = .884, RMSEA = .27, 90% CI [.20, −.35]).

In support of our account, however, the partial mediation model (Model 3; with a direct effect of ToM on performance) resulted in the best fit ($\chi^2 = 3.17$, $df = 2$, $p = .00$, CFI = .996, NFI = .990, IFI = .996, RMSEA = .06, 90% CI [.00, −.18]; $\beta_{Age \rightarrow ToM \rightarrow SE \rightarrow Perf.} = -.48$). The model fit was superior to both the full mediation model ($\Delta\chi^2 = 68.55$, $\Delta df = 1$, $p < .001$) and the nonmediation model ($\Delta\chi^2 = 34.68$, $\Delta df = 1$, $p < .001$). Of note, with small samples (such as in our case), the chi-square statistic lacks power and may not discriminate between good-fitting models and poor-fitting models (Kenny & McCoach, 2003). Due to the restrictiveness of the model chi-square, researchers have sought alternative indices to assess model fit. One statistic that minimizes the impact of sample size is Wheaton, Muthen, Alwin, and Summers's (1977) relative/normed chi-square (χ^2/df). Although there is no consensus regarding an acceptable ratio for this statistic, some have recommended that the ratio not exceed 5.0 (Wheaton et al., 1977) while others have recommended that the ratio not exceed 2.0 (Tabachnick & Fidell, 2007). Our χ^2/df ratio is 1.59, suggesting that our model provides an acceptable fit.

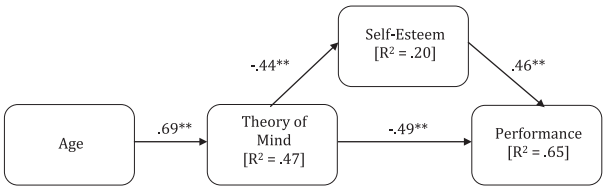


Figure 2. Structural equation model demonstrating that as children age, increases in theory of mind lead to decreases in self-esteem that affect the desire to perform. Performance is coded as the total choices of singing and dancing out of the four options (range = 0–2).
**$p < .001$

In this partial mediation model (Model 3), age was positively related to ToM (β = .69, $p < .001$, 95% CI [.64, .74]) and ToM in turn was significantly and negatively related to self-esteem (β = −.44, $p < .001$, 95% CI [−.69, −.19]) and performance (β = −.49, $p < .001$, 95% CI [−.60, −.36]; Figure 2). Self-esteem partially mediated the relation between ToM and performance, with a significant relation between self-esteem and performance (β = .46, $p < .001$, 95% CI [.33, .57]). This model explained 65% of the variance in performance.

Finally, we tested an alternative partial mediation model (Model 4; with ToM as a mediator between self-esteem and performance). This model resulted in a poor fit to the data ($\chi^2 = 77.80$, $df = 3$, $p = .00$, CFI = .743, NFI = .745, IFI = .749, RMSEA = .49, 95% CI [.40, −.58]). In support of our hypothesis, a chi-square difference test (Kline, 1998) established the superiority of Model 3 (age → ToM → self- esteem → performance) over Model 4 (age → self-esteem → ToM → performance; $\Delta\chi^2 = 34.68, \Delta df = 1, p < .001$).

Discussion

Why do children avoid performing as they grow older? Our results support our account that ToM appears to equip children with the ability to predict that others may not view their performance as favorably as they do, which is associated with decreased self-esteem—and avoiding the spotlight. Note that our data address a salient alternative explanation for our pattern of performance avoidance, one familiar to anyone interacting with

socially awkward adolescents or preteens: As children enter puberty they experience a host of changes that decrease their desire to perform. However, our results show that the shift away from performance begins as early as age 4, years before children enter puberty—suggesting that these changes in later childhood are unlikely to account for our results.

The present research has several limitations that warrant further research. First, although our analyses address several alternative explanations, further experiments are needed to establish a causal impact of ToM on performance. Second, we used ToM measures that were not only appropriate for use with very young children but also resonated with older children; future research is needed to develop measures better suited for wider age ranges. Third, what specific aspects of self-esteem—a multifaceted construct—ultimately link to lower desires to perform? We show that children's explicit self-esteem is linked to avoiding performance, but a fuller understanding of this process could be gained by assessing both implicit and explicit self-esteem (Dunham, Baron, & Banaji, 2007; Greenwald & Banaji, 1995). Fourth, other mediating factors likely play a role in decreasing performance, such as mistaken—and likely erroneously negative—perceptions of peers' opinions of one's abilities, as well as antecedent states such as mood. Finally, future research should examine the generalizability of our results to other kinds of performance behaviors and other kinds of performance situations (e.g., alone vs. in public). Our conceptualization suggests that regardless of the performance behavior—from singing and dancing to playing musical instruments and acting—a more developed ToM is linked to avoidance of performance. At the same time, some evidence suggests that the effect of a peer audience on behavior in younger versus older children is domain specific (Banerjee & Lintern, 2000; Banerjee & Yuill, 1999), suggesting an interesting area for further examination.

Previous research has examined links between development of ToM and increases in negative behaviors such as antisocial deception (Repacholi, Slaughter, Pritchard, & Gibbs, 2003). Our results suggest that ToM may also be linked to decreases in enjoyable behaviors—like singing and dancing—that are associated with activities shown to have benefits for health and happiness (e.g., Bonilha, Onofre, Vieira, Prado, & Martinez, 2008; Brown et al., 2005; Clift & Hancox, 2010). This unwillingness to perform likely perpetuates beyond school dances, affecting the willingness of people older than 12—adults—to engage in such playful behaviors as well.

References

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief-desire reasoning in children and adults. *Child Development*, *82*, 1691–1703. doi:10.1111/j.1467-8624.2011.01635.x

Baker, J. (2003). *Social skills training for children and adolescents with Asperger syndrome and social-communication problems*. Lenexa, KS: Autism Asperger.

Banerjee, R., Bennett, M., & Luke, N. (2012). Children's reasoning about self-presentation following rule violations: The role of self-focused attention. *Child Development*, *83*, 1805–1821. doi:10.1111/j.1467-8624.2012.01813.x

Banerjee, R., & Lintern, V. (2000). Boys will be boys: The effect of social evaluation concerns on gender-typing. *Social Development*, *9*, 397–408. doi:10.1111/1467-9507.00133

Banerjee, R., & Yuill, N. (1999). Children's understanding of self-presentational display rules: Associations with mental-state understanding. *British Journal of Developmental Psychology*, *17*, 111–124.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition*, *21*, 37–46. doi:10.1016/0010-0277(85)90022-8

Begeer, S., Malle, B. F., Nieuwland, M., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Child Development*, *7*, 104–122. doi:10.1080/17405620903024263

Bonilha, A. G., Onofre, F., Vieira, M. L., Prado, M. Y. A., & Martinez, J. A. B. (2008). Effects of singing classes on pulmonary function and quality of life of COPD patients. *International Journal of COPD*, *4*, 1–8.

Brown, W. M., Cronk, L., Grochow, K., Jacobson, A., Liu, C. K., Popovic, Z., & Trivers, R. (2005). Dance reveals symmetry especially in young men. *Nature*, *438*, 1148–1150. doi:10.1038/nature04344

Bruner, J., & Feldman, C. (1993). Theories of mind and the problem of autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 334–366). New York, NY: Oxford University Press.

Byrne, B. M. (2001). *Structural equation modeling with Amos: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Clift, S., & Hancox, G. (2010). The significance of choral singing for sustaining psychological wellbeing: Findings from a survey of choristers in England, Australia and Germany. *Music Performance Research*, *3*, 79–96.

Cutting, A., & Dunn, J. (2002). The cost of understanding other people: Social cognition predicts young children's sensitivity to criticism. *Journal of Child Psychology and Psychiatry*, *43*, 849–860. doi:10.1111/1469-7610.t01-1-00047

Dunham, Y., Baron, A. S., & Banaji, M. R. (2007). Children and social groups: A developmental analysis of

implicit consistency in Hispanic Americans. *Self and Identity*, *6*, 238–255. doi:10.1080/15298860601115344

Dunn, J. (1995). Children as psychologist: The later correlates of individual differences in understanding emotion and other minds. *Cognition and Emotion*, *9*, 187–201. doi:10.1080/02699939508409008

Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, *3*, 108–124.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind." *Trends in Cognitive Sciences*, *7*, 77–83. doi:10.1016/S1364-6613(02)00025-6

Gauvain, M. (1998). Culture, development, and theory of mind: Comment on Lillard. *Psychological Bulletin*, *123*, 37–42. doi:http://dx.doi.org/10.1037/0033-2909.123.1.37

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26–37. doi:10.2307/1130386

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. doi:http://dx.doi.org/10.1037/0033-295X.102.1.4

Harter, S. (1982). The Perceived Competence Scale for Children. *Child Development*, *53*, 87–97. doi:10.2307/1129640

Harter, S. (1985). *Manual for the Self-Perception Profile for Children: Revision of the Perceived Competence Scale for Children*. Denver, CO: University of Denver.

Harter, S., & Pike, R. (1984). The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children. *Child Development*, *55*, 1969–1982. doi:10.2307/1129772

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408–420. doi:10.1080/03637750903310360

Howlin, P., Baron-Cohen, S., & Hadwin, J. (1999). *Teaching children with autism to mind-read*. West Sussex, UK: Wiley.

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A mediation on mediation: evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology*, *17*, 139–153. doi:10.1016/S1057-7408(07)70020-7

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, *10*, 333–3511. doi:10.1207/S15328007SEM1003_1

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Lecce, S., Caputi, M., & Hughes, C. (2011). Does sensitivity to criticism mediate the relationship between theory of mind and academic achievement? *Journal of Experimental Child Psychology*, *110*, 313–331. doi:10.1016/j.jecp.2011.04.011

Malle, B. F. (2005). Folk theory of mind: Conceptual foundations of human social cognition. In R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 225–255). New York, NY: Oxford University Press.

Nguyen, L., & Frye, D. (1999). Children's theory of mind: Understanding of desire, belief and emotion with social referents. *Social Development*, *8*, 70–92. doi:10.1111/1467-9507.00081

Pellicano, E. (2010). The development of core cognitive skills in autism: A 3-year prospective study. *Child Development*, *81*, 1400–1416. doi:10.1111/j.1467-8624.2010.01481.x

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that. . ." Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471. doi:10.1016/0022-0965(85)90051-7

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers*, *36*, 717–731. doi:10.3758/BF03206553

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/BRM.40.3.879

Repacholi, B., Slaughter, V., Pritchard, M., & Gibbs, V. (2003). Theory of mind, Machiavellianism, and social functioning in childhood. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind* (pp. 67–97). New York, NY: Psychology Press.

Tabachnick, G. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Belmont, CA: Duxbury.

Watling, D., & Banerjee, R. (2012). Children's understanding of disclaimers. *Social Cognition*, *30*, 18–36. doi:10.1521/soco.2012.30.1.18

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). The causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*, 383–388. doi:10.1016/j.tics.2010.05.006

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development*, *72*, 655–684. doi:10.1111/1467-8624.00304

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84–136. doi:10.2307/270754

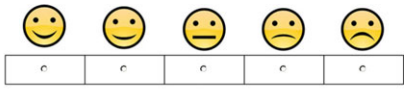
Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. doi:10.1016/0010-0277(83)90004-5

Appendix A. Theory of Mind Tasks

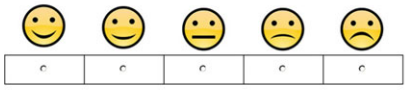
Task	Description	Test and control questions	Scoring
Sally–Anne (Baron-Cohen, Leslie, & Frith, 1985)	In the narrated story (with pictures), a doll (Sally) places a toy car in her basket and leaves the scene. Another doll (Anne) moves the toy car to a box while Sally is away. When Sally returns she goes to look for the toy car.	Test: "Where will Sally look first for her toy car?" Control: 1. "Where is the toy car really?" 2. "Where did Sally put the toy car in the beginning?"	*Pass*-basket *Fail*-box
Misleading container (Gopnik & Astington, 1988)	We filled a familiar cookie box with pencils. All children were first asked to name the expected content (cookies). The researcher then revealed the true content of the box (i.e., pencils), before closing it and asking a control question. A control question was asked immediately after showing true contents of box but before test question.	Control: "What is really in the box?" Test: "Someone else is coming in next. The person hasn't looked inside this box before. When I show it to him/her all closed up like this, what will he/she say is in it?"	*Pass*-cookies *Fail*-pencils
Duck and Lion (Nguyen & Frye, 1999)	Children were shown a story about Duck and Lion, and the scenes were narrated by the researcher. Narrated story: "*Duck and Lion are reading. Duck leaves the room to get another book to read with Lion. Lion decides to stop reading and go to sleep." *Duck and Lion were replaced with David and Liam for ages 6 and older.	Test: "What does Duck (David) think Lion (Liam) is doing in the room?" Control: "What was Lion doing at the beginning?"	*Pass*-reading *Fail*-sleeping

Appendix B. Self-Esteem Scale

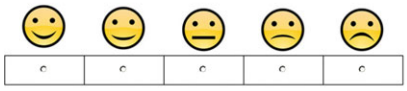
1. When I Think About How Many Friends I Have, I Feel:



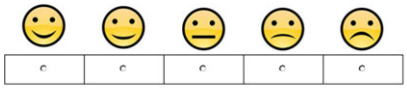
2. When I Think About How Good I Am at Things, I Feel:



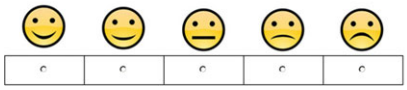
3. When I Think About How Well I Behave, I Feel:



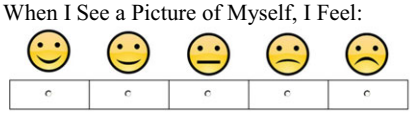
4. When I Think About How Good I Am in School, I Feel:



5. When I Think About How Much Other Kids Like Me, I Feel:



6. When I See a Picture of Myself, I Feel:



7. When I Think About the Things I Can Do, I Feel: