Online Appendix

for

If You Could Read My Mind—An Experimental Beauty-Contest Game with Children

Henning Hermes and Daniel Schunk

A Descriptive Statistics

Table A1: Sample Descriptives

	Mean	SD	N	Min	Max
Female	0.55	0.50	114	0	1
Age in years	9.98	1.04	114	8	12
Class 1	0.22	0.42	114	0	1
Class 2	0.22	0.42	114	0	1
Class 3	0.13	0.34	114	0	1
Class 4	0.13	0.34	114	0	1
Class 5	0.13	0.34	114	0	1
Class 6	0.17	0.37	114	0	1
Fifth Grade	0.44	0.50	114	0	1
Fluid IQ	0.01	1.00	114	-3.4	1.2
Understanding of the Game	19.13	1.59	113	9	20
Perspective-taking	0.22	0.42	112	0	1
Social Appropriateness	0.00	1.00	114	-6	1.2
Interpersonal Reactivity	0.00	1.00	114	-3.1	2.4

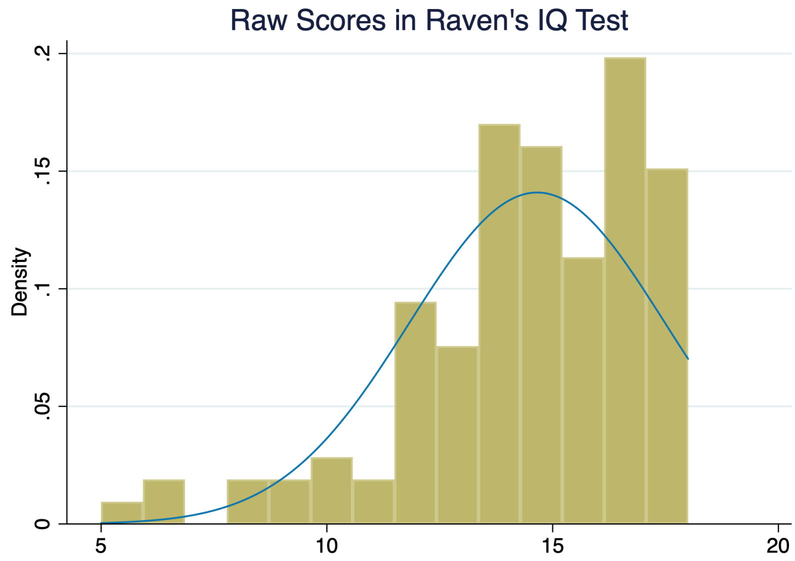
Notes: Sample descriptives for the children sample. We collected data for 23 groups of five children. For one child, we did not obtain parental consent for data use. Therefore, the final number of observations is 114. For one child, we are missing data on the ratings for understanding of the game because the experimenter did not record responses for these children in the protocol sheet.

Table A2: Game Descriptives

	Mean	Median	SD	N	Min	Max
Number Round 1	33.46	28	20.64	114	1	91
Number Round 2	21.73	20	12.17	114	4	56
Number Round 3	14.96	12	11.45	114	1	75
Number Round 4	10.68	8	11.18	114	0	100
Number Round 5	10.04	6.5	14.47	114	0	100
Number Round 6	5.42	3	5.38	114	0	30
Number Round 7	3.89	2.5	4.81	114	0	27
Number Round 8	3.36	2	5.30	114	0	44
Number Round 9	3.33	1	9.77	114	0	100
Number Round 10	2.51	1	5.83	114	0	50
Mean Number 1–10	10.94	10	4.91	114	2.6	26

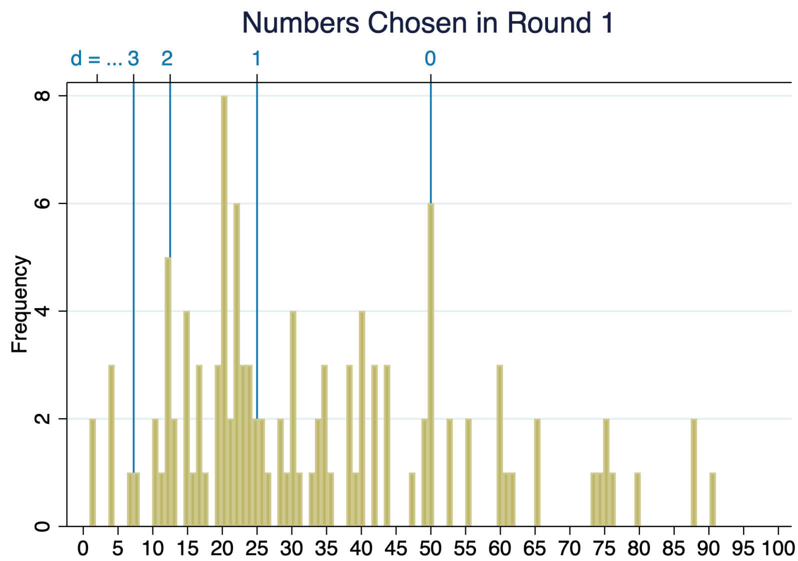
Notes: Descriptives for the numbers chosen in each round of the game for the children sample. We collected data for 23 groups of five children. As we could not obtain parental consent for data use for one child, the final number of observations is 114.

Figure A1: Histogram of Raven's IQ Scores



Notes: This figure shows a histogram of the raw scores in the Raven's IQ test. In principle, scores range from 0–18; in the sample, the minimum score is 5, the maximum score is 18. The average score is 14.7 and the median score is 15. The number of observations is 114.

Figure A2: Histogram of Numbers Chosen in Round 1



Notes: This figure shows a histogram for the numbers chosen in round 1 in the children sample. The blue vertical lines display resulting values for a depth of reasoning (d) of 0, 1, 2, and 3 (starting from an initial reference point of 50). See Section 3.1 for details on depth of reasoning.

B Further Analyses

Table A3: Fluid IQ and Weakly Dominated Choices in the BCG

	DV: Dominated Choice in R1		**DV: Dominated Choice in R1–10**	
	R1 (OLS) (1)	R1 (Probit) (2)	R1–R10 (OLS) (3)	R1–R10 (Probit) (4)
Female	0.169	0.531	0.163	0.504
	(0.122)	(0.380)	(0.131)	(0.389)
Age in years	0.056	0.198	0.005	-0.002
	(0.137)	(0.446)	(0.108)	(0.345)
Fluid IQ	-0.118	-0.401*	-0.106*	-0.364*
	(0.068)	(0.221)	(0.056)	(0.206)
Group FE	Yes	Yes	Yes	Yes
Observations	74	74	89	89

Notes: The results are based on OLS regressions (columns 1 and 3) and probit models (columns 2 and 4). Standard errors in parentheses are clustered at the group level. We cannot use all observations in these models because in round 1, there are eight groups in which no child chose a weakly dominated number, and over all 10 rounds, there are such five groups. Therefore, the variable for weakly dominated choices is perfectly collinear with the group dummy and these groups have to be dropped. * p<.10, ** p<.05, *** p<.01

Table A4: OLS Regressions for Round 1

	DV: Nr. R1 (1)	DV: Nr. R1 $\leq$ 50 (2)	DV: Nr. R1 Age (3)	DV: Coins R1 (4)	DV: Rank R1 (5)	DV: Dist. R1 (6)
Female	4.561	-0.330	5.721	0.107	0.208	0.383
	(3.571)	(2.980)	(4.291)	(0.081)	(0.268)	(0.232)
Age in years	8.639**	7.020**	12.036*	-0.103	0.513	0.606*
	(3.962)	(2.524)	(6.942)	(0.118)	(0.340)	(0.295)
Fluid IQ	-1.878	3.345*	-2.370	0.078	-0.083	-0.275
	(3.762)	(1.841)	(4.198)	(0.065)	(0.227)	(0.299)
Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	114	94	100	114	114	114

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. Column (2) excludes all weakly dominated choices, i.e., numbers larger than 50. Column (3) excludes age outliers, i.e., all children below the 5th and above the 95th percentile within their grade. For details on the outcomes in columns (4)–(6), see Section 2.4.
* p<.10, ** p<.05, *** p<.01

Table A5: OLS Regressions for Depth of Reasoning in R2–10

	DV: Depth of Reasoning R2–R10 (1)
Female	-0.000
	(0.205)
Age in years	0.023
	(0.167)
Fluid IQ	0.059
	(0.086)
Group FE	Yes
Observations	114

Notes: The results are based on OLS regressions. The dependent variable is the average depth of reasoning in rounds 2–10 (see Section 3.1 for details), capped at values -6 and +6. Standard errors in parentheses are clustered at the group level. * p<.10, ** p<.05, *** p<.01

Choices in the First Round. Table A4 reports our analysis of numbers chosen in the first round. In this round, no prior behavior from other players has been observed. There is no significant effect of fluid IQ on first-round choices. The coefficient in column (1) is slightly negative; however, once weakly dominated choices (> 50) are excluded—see column (2)—the coefficient even turns positive (and is significant at the 10% level). Also, for all measures of successful performance, there is no significant link to fluid IQ, as shown in columns (4)–(6). Hence, even for first-round choices, cognitive skills (measured as fluid IQ) are not related to lower entries or successful performance in the new design of the experimental BCG.

Table A4 also reports a significant and large effect of age: older children choose higher numbers in the first round of the game. This is not driven by weakly dominated choices, as column (2) indicates. To rule out potential effects of age outliers, we exclude children below the 5th and above the 95th age percentile within each cohort (third grade and fifth grade). Results in column (3) indicate that the positive link between age and first-round choices persists (the coefficient even increases); therefore, it appears that age outliers are not driving this finding. Columns (4)–(6) demonstrate that these higher numbers translate into worse performance for older children in the first round (but statistically significant only for the distance measure). However, the effect does not transfer to the next rounds (see results from Tables 3; also, when looking at the numbers chosen in round 2, the relationship fades out and becomes insignificant). At first sight, a reasonable explanation would be that older children are (more) familiar with higher numbers. However, we only compare children within groups coming from the same class. Therefore, all of them learn together and have the same experiences in mathematics. We argue that this makes it very unlikely that there are any age-related differences with respect to familiarity with numbers within class (and within groups).[23] Overall, there seems to be a strong link between older children choosing higher numbers that cannot be explained by weakly dominated choices, age outliers, or familiarity with higher numbers.

[23] One could, in principle, think of children repeating a grade but when excluding these age outliers, the effect persists, see column (3).

C Robustness Checks

To support the stability of our findings, we conduct a number of robustness checks. Our tables for robustness checks always report results from re-estimating the models in Table 3.

First, we exclude all children with an experimenter-rated understanding of the game below 19 points (out of 20 points). Results in Table A6 show that our findings are not affected by this sample restriction. Second, we wanted to check whether outliers or extreme choices are driving our results. Therefore, we reproduced our estimations excluding children who chose weakly dominated numbers (see Table A7). Third, in light of the (non-existing) age effects discussed above (see Section 3.3), we wanted to analyze whether the presence of very young or very old children within their grade (e.g., because of grade retention) could influence the results obtained. Therefore, we excluded children below the 5th and above the 95th age percentile (within their cohort) and present the results for these models in Table A8. The coefficients for age remain insignificant, confirming that there is no systematic effect of age on successful performance in the BCG in our age range (however, coefficients now at least point into the hypothesized direction, indicating that in a larger age range one might be able to detect an effect of age). Fourth, all our main findings are conservatively estimated using group-fixed effects, i.e., we only compare children within a group of five kids. The upshot of this approach is that we control for any unobserved heterogeneity across groups. However, we also report estimations using OLS without group-fixed effects in Table A9 in the Appendix. When excluding group-fixed effects, there is still no significant relationship between individual characteristics and performance in the BCG. Finally, to support the notion that age is not substantially related to successful performance and is only linked to choices in the first round, we also present estimations with and without age in Table A10 in the Appendix. Removing age as a control variable essentially does not alter the coefficients for any of our determinants of success in the game.

In addition, we also checked whether choices and performance in the game were related to two external factors, using ANOVAs. First, the color (and also the seating position at the table while playing the game, as this was determined by the color, see Figure A4 in the Appendix) did not influence performance throughout the game. Second, experimenters rotated between colors across groups so that they would explain the game to players with different colors each round.[24] Therefore, we can identify a separate effect of the person who explained the game to the child. There was no significant difference in choices or

[24] Experimenter A explained the game to the child with white in group 1, then to the child with yellow in group 2, and so on, and experimenter B explained the game the child with yellow in group 1, blue in group 2, and so on.

performance with respect to the experimenter who explained the game to the child. Because we use group-fixed effects in all main estimations, other differences such as time of testing, class-, teacher- or school-fixed effects are captured in the group dummy.

Table A6: OLS Regressions with Coins Won, Rank, and Distance Excluding Low Understanding

	DV: Coins R2–10 (1)	DV: Rank R2–10 (2)	DV: Distance R2–10 (3)
Female	-0.090	0.081	-0.392
	(0.464)	(0.205)	(0.496)
Age in years	-0.126	0.060	-0.346
	(0.692)	(0.259)	(0.620)
Fluid IQ	0.305	-0.041	0.203
	(0.306)	(0.155)	(0.287)
Group FE	Yes	Yes	Yes
Observations	90	90	90

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. In this table, children with an experimenter-rated understanding of the game below 19 points (out of 20) are excluded from the analysis. For details on the outcomes, see Section 2.4.
* p<.10, ** p<.05, *** p<.01

Table A7: OLS Regressions with Coins Won, Rank, and Distance Excluding Weakly Dominated Choices

	DV: Coins R2–10 (1)	DV: Rank R2–10 (2)	DV: Distance R2–10 (3)
Female	0.007	0.040	0.067
	(0.478)	(0.204)	(0.447)
Age in years	-0.261	0.111	-0.145
	(0.612)	(0.215)	(0.505)
Fluid IQ	0.436	-0.141	-0.029
	(0.277)	(0.121)	(0.212)
Group FE	Yes	Yes	Yes
Observations	87	87	87

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. In this table, children who chose a weakly dominated number in any of the 10 rounds are excluded from the analysis. For details on the outcomes, see Section 2.4.
* p<.10, ** p<.05, *** p<.01

Table A8: OLS Regressions with Coins Won, Rank, and Distance Excluding Age Outliers

	DV: Coins R2–10 (1)	DV: Rank R2–10 (2)	DV: Distance R2–10 (3)
Female	-0.228	0.123	-0.073
	(0.349)	(0.183)	(0.447)
Age in years	0.456	-0.101	-0.835
	(0.370)	(0.198)	(0.843)
Fluid IQ	0.365	-0.134	-0.076
	(0.268)	(0.106)	(0.246)
Group FE	Yes	Yes	Yes
Observations	100	100	100

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. In this table, children aged below the 5th or above the 95th percentile within their respective grade (third or fifth grade) are excluded from the analysis. For details on the outcomes, see Section 2.4. * p<.10, ** p<.05, *** p<.01

Table A9: OLS Regressions with Coins won, Rank, and Distance (No Group FEs)

	DV: Coins R2–10		DV: Rank R2–10		DV: Distance R2–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.072	-0.148	0.033	0.095	-0.305	-0.239
	(0.370)	(0.363)	(0.177)	(0.147)	(0.426)	(0.373)
Age in years	-0.200	0.171	0.120	-0.009	-0.156	0.010
	(0.506)	(0.219)	(0.203)	(0.062)	(0.490)	(0.250)
Fluid IQ	0.251	0.194	-0.068	-0.067	0.027	0.099
	(0.236)	(0.194)	(0.101)	(0.078)	(0.220)	(0.197)
Group FE	Yes	No	Yes	No	Yes	No
Observations	114	114	114	114	114	114

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. For details on the outcomes, see Section 2.4. * p<.10, ** p<.05, *** p<.01

Table A10: OLS Regressions with Coins Won, Rank, and Distance (Not Controlling for Age)

	DV: Coins R2–10		DV: Rank R2–10		DV: Distance R2–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.072	0.063	0.033	0.038	-0.305	-0.312
	(0.370)	(0.363)	(0.177)	(0.175)	(0.426)	(0.413)
Age in years	-0.200		0.120		-0.156	
	(0.506)		(0.203)		(0.490)	
Fluid IQ	0.251	0.257	-0.068	-0.071	0.027	0.031
	(0.236)	(0.232)	(0.101)	(0.100)	(0.220)	(0.229)
Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	114	114	114	114	114	114

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. Columns (2), (4), and (6) present the main results without controlling for age. For details on the outcomes, see Section 2.4. * p<.10, ** p<.05, *** p<.01

D Exploratory Analysis of Perspective-Taking Abilities in Strategic Interaction

In an attempt to provide some first, exploratory evidence on the importance of *other* abilities than cognitive skills for successful strategic interaction, we also collected data on children's perspective-taking abilities. In contrast to cognitive skills, which we measured as fluid IQ, using an established test (Raven's Matrices), for perspective-taking abilities there is no such an established measure, at least not for children in our age range. Therefore, we decided to measure perspective-taking abilities with different instruments, covering several methods as well as several aspects of perspective-taking abilities. Specifically, we decided to use (i) an easy and quick-to-implement behavioral task measuring perspective-taking ("E on the forehead", Glen 1984), and to use two different self-reported measures, (ii) one focusing on understanding social situations as well as emotions and behavior in third-person situations (Meindl 1998), and (iii) another focusing more on an individual difference perspective along the dimension of cognitive empathy, i.e., the ability to understand and process other people's emotions and perspectives (Garton and Gringart 2005). The latter questionnaire was developed as a child-friendly version of a frequently used tool to measure perspective-taking and empathy in adults, namely the Interpersonal Reactivity Index (IRI, Davis 1983).

Answers for the two questionnaire measures were collected in the first lesson of each school day, together with the whole class (see Section 2.1). The behavioral task was conducted at the end of the explanation session which each child received individually with an experimenter, right before children went to the large table to play the game together with the other children (see Section 2.3 for details). Table A1 provides descriptive statistics for our measures of perspective-taking abilities; the exact instructions for the behavioral perspective-taking task and the wording of the items of the questionnaires are provided in Sections F.2 and G. Specifically, we use the following variables for our subsequent analysis:

Perspective-taking. To provide a behavioral measure of perspective-taking abilities, we adapted the "E on the forehead" task (Glen 1984). This task was designed as a measure of perspective-taking based on self-awareness or self-consciousness and has frequently been used in psychological experiments as a quick and intuitive measure of perspective-taking behavior vs. egocentric responses (see, for example, Steins and Wicklund 1996; Galinsky et al. 2006). In the present study, we asked each child to "trace a capital 'E' with your forefinger on your forehead" and recorded, whether the 'E' was readable from the child's or the experimenter's perspective (or something different was traced). The dummy variable is one if the child traces an 'E' that is readable from the experimenter's perspective and zero if she traces something different (i.e., an 'E' readable from the child's perspective or something different). We were able to collect this task from n = 112 children in our sample; 25 children (22%) traced an 'E' that was readable from the experimenter's perspective.

Social Appropriateness Scale. As one of our two self-reported measures of perspective-taking, we selected six small stories from a questionnaire measuring empathy and appropriate behavior in a social situation that was developed by Meindl (1998). Children receive a point for each correctly solved question. To enable comparison of effect sizes, the variable was standardized to mean = 0 and SD = 1. All children in our sample (n = 114) answered this questionnaire.

Interpersonal Reactivity Index. As a second self-reported measure for perspective-taking abilities, we use a questionnaire adapted for school-aged children developed by Garton and Gringart (2005). This questionnaire is based on one of the most frequently used self-reported measures for perspective-taking and empathy in adults, namely the Interpersonal Reactivity Index (IRI, Davis 1983). We selected four items focusing on cognitive aspects of empathy from the children's version of the questionnaire. In contrast to the Social Appropriateness scale, there is no correct answer, but children rate how much each item applies to them personally. We use the sum of all four items (no reverse-coded item) and standardize the resulting variable to mean = 0 and SD = 1 to enable comparison of effect sizes. All children in our sample (n = 114) answered this questionnaire.

To analyze the relationship between perspective-taking abilities and successful performance in the experimental BCG, we simply add the three measures of perspective-taking abilities to the models we estimated in Table 3. Results are presented in Table A11. Note that our measures for fluid IQ, Social Appropriateness, and Interpersonal Reactivity are standardized to mean = 0 and SD = 1; thus, one can easily compare the size of their coefficients.

As columns (1) and (2) in Table A11 show, all three measures of perspective-taking significantly predict the number of rounds a child wins during the game. When including group-fixed effects, children displaying high perspective-taking abilities in the "E on the forehead" task (labeled "perspective-taking" in the tables) win more than one additional coin (on average, children win 3.8 coins, SD = 1.9). A one standard deviation increase in

Table A11: Exploratory Analysis on the Role of Perspective-Taking Abilities

	DV: Coins R2–10		DV: Rank R2–10		DV: Distance R2–10		DV: Depth of Reas.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.173	-0.068	-0.008	0.068	-0.251	-0.221	0.001
	(0.376)	(0.360)	(0.170)	(0.139)	(0.476)	(0.410)	(2.170)
Age in years	-0.327	0.056	0.175	0.034	-0.066	0.065	0.552
	(0.487)	(0.230)	(0.202)	(0.072)	(0.446)	(0.253)	(1.697)
Fluid IQ	0.137	0.141	-0.013	-0.051	0.179	0.117	-0.297
	(0.260)	(0.211)	(0.110)	(0.084)	(0.283)	(0.219)	(0.843)
Perspective-taking	1.105**	1.149***	-0.532**	-0.388***	-1.652*	-0.219	5.940*
	(0.416)	(0.336)	(0.192)	(0.118)	(0.900)	(0.495)	(3.435)
Social Appropriateness	0.258*	0.345**	-0.112*	-0.102**	-0.292	-0.246	1.418**
	(0.130)	(0.133)	(0.064)	(0.039)	(0.317)	(0.257)	(0.542)
Interpersonal Reactivity	-0.450***	-0.404**	0.200***	0.157**	0.306	0.173*	-2.080**
	(0.155)	(0.182)	(0.065)	(0.057)	(0.214)	(0.091)	(0.982)
Group FE	Yes	No	Yes	No	Yes	No	Yes
Observations	112	112	112	112	112	112	112

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level.
* p<.10, ** p<.05, *** p<.01

scores on the Social Appropriateness scale is linked to 0.26 more coins. Surprisingly, while scores from the Interpersonal Reactivity Index (IRI) significantly predict performance in the experimental BCG, they are *negatively* linked to game performance: a one standard deviation increase in scores from the IRI is related to 0.45 fewer coins during rounds 2–10.

Similarly, perspective-taking abilities are strongly linked to successful performance measured as the rank in distance to the best response (see Section 2.4 for details on the dependent variables), as presented in columns (3) and (4) in Table A11. Controlling for group-fixed effects, children with high perspective-taking abilities in the "E on the forehead task" perform half a rank better (mean rank over rounds 2–10 is 2.5, SD = 0.70). Social Appropriateness is significantly linked to better performance (one SD increase improves the average rank by around 0.1), while the IRI is related to worse performance (a one SD increase worsens the average rank by around 0.2). Columns (5) and (6) show a very similar pattern for the average distance over rounds 2–10 as an outcome, although most relationships are no longer statistically significant. Finally, similar to our findings for measures of actual game *performance*, higher depth of reasoning is systematically associated with perspective-taking abilities (see column (7) in Table A11).

All in all—in contrast to fluid IQ as an important part of cognitive skills—measures of perspective-taking abilities predict successful performance in an experimental BCG: Children showing high perspective-taking abilities in the "E on the forehead" task and scoring high on Social Appropriateness demonstrate better success (and higher depth of reasoning), whereas children with higher scores on the Interpersonal Reactivity Index show worse performance in the game.

But why are children who traced the 'E' so that the experimenter could read it on average so much better in strategic interaction (controlling for other characteristics such as gender, age, and fluid IQ)? Our preferred hypothesis is that the ability to form accurate beliefs about other players' choices and to "put yourself into the other players' shoes" is a key ability to succeed in strategic interaction games. Thus, while we are not aware of any prior study using the "E on the forehead" task to predict strategic interaction behavior, tracing the 'E' in this way seems to be indicative of a very important ability in the area of perspective-taking that makes children successful in strategic interactions. Further studies should shed light on the exact underlying skills or dispositional characteristics that are indicated using this simple behavioral task. Likewise, the Social Appropriateness score, which has mainly been used to assess school-aged children's social and emotional competencies with a focus on empathy (e.g., Schick and Cierpka 2005), is positively associated with successful game performance. However, the relationship is smaller in size. Still, this score seems to correlate reasonably strong with successful performance and could, thus, be of interest for further research with children in this area. The Interpersonal Reactivity Index (IRI) is associated with successful performance to a greater extent and it is highly significant in most of our specifications. Yet, the relationship for the IRI is negative, indicating that high levels of interpersonal reactivity can harm successful performance in the BCG. Why is this the case? In contrast to the other two perspective-taking measures, the IRI score is already significantly linked to choices in the first round and also first-round performance. However, this relationship seems strongly driven by choices above 50: excluding weakly dominated choices eliminates the link between IRI and higher numbers in the first round. In addition, raw correlations between IRI scores and choosing a weakly dominated number in round 1 ($\rho = .25, p = .007$) confirm that children scoring high on the IRI have a higher probability of choosing a number above 50 in the first round. However, this is not the case for higher numbers in general: for example, the correlation between IRI scores with a dummy for choosing numbers higher than 30 is zero. We can only speculate on why the relationship between IRI scores and performance is negative. One potential explanation could be that the IRI is indicative of children who focus on other aspects of the game (or on the other children's behavior in the game) that do not improve performance and occupy cognitive resources. Another hypothesis is that children scoring high on the IRI focus too much on *past* behavior instead of focusing on adjusting the number to be chosen in the right way.

Taken together, our measures of perspective-taking abilities are strongly linked to successful performance in an experimental BCG, although in different directions. Further studies should look more closely at the relationship between this set of abilities and performance in strategic interaction settings, as well as potential interaction effects of cognitive skills and perspective-taking abilities.

E Adult Sample

Table A12: Sample Descriptives (Adults)

	Mean	SD	N	Min	Max
Female	0.60	0.49	120	0	1
Age in years	22.63	4.18	120	18	46
Fluid IQ	0.00	1.00	120	-3.12	1.78

Notes: Sample descriptives for the adult sample. The study was conducted in 12 sessions with 10 participants each (two groups of five adults per session).

Table A13: Game Descriptives (Adults)

	Mean	Median	SD	N	Min	Max
Number Round 1	23.64	21	14.92	120	1	99
Number Round 2	13.47	11	9.00	120	0	47
Number Round 3	8.31	7	7.33	120	0	50
Number Round 4	6.64	4	9.55	120	0	58
Number Round 5	4.83	3	9.21	120	0	90
Number Round 6	3.13	2	6.70	120	0	60
Number Round 7	2.75	1	8.42	120	0	87
Number Round 8	1.59	1	2.74	120	0	17
Number Round 9	1.54	1	5.20	120	0	50
Number Round 10	1.12	0	4.42	120	0	46
Mean Number 1–10	6.70	5.9	4.08	120	1.2	25

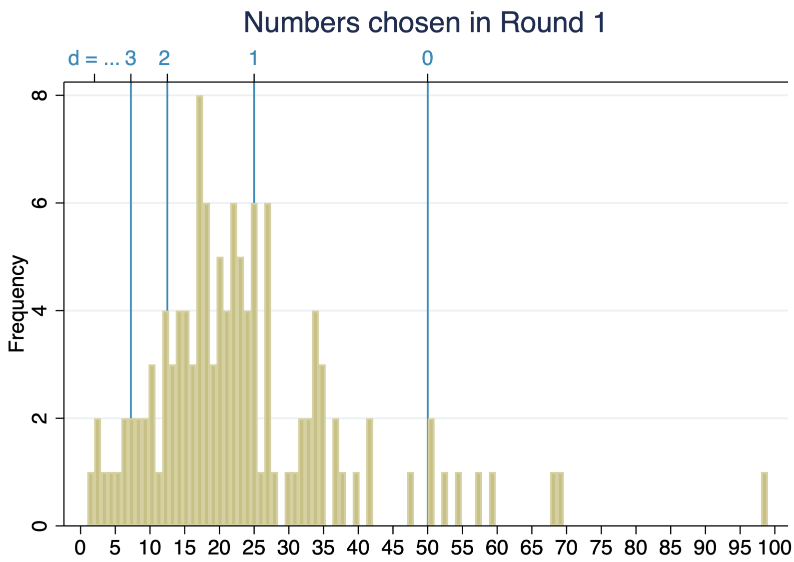
Notes: Descriptives for the numbers chosen in each round of the game for the adult sample.

Table A14: OLS Regressions with Coins Won, Rank, and Distance (Adults)

	Coins R2–10	Rank R2–10	Distance R2–10
Female	-1.282***	0.444**	0.158
	(0.440)	(0.203)	(0.750)
Age in years	-0.040	0.009	0.048
	(0.049)	(0.018)	(0.090)
Fluid IQ	0.183	-0.083	0.026
	(0.276)	(0.122)	(0.295)
Group FE	Yes	Yes	Yes
Observations	120	120	120
R-squared	0.441	0.221	0.180

Notes: The results are based on OLS regressions. Standard errors in parentheses are clustered at the group level. * $p<.10$, ** $p<.05$, *** $p<.01$

Figure A3: Histogram of Numbers Chosen in Round 1 (Adults)



Notes: This figure shows a histogram for the numbers chosen in round 1 in the adult sample. The blue vertical lines display resulting values for a depth of reasoning (d) of 0, 1, 2, and 3 (starting from an initial reference point of 50). See Section 3.1 for details on depth of reasoning.

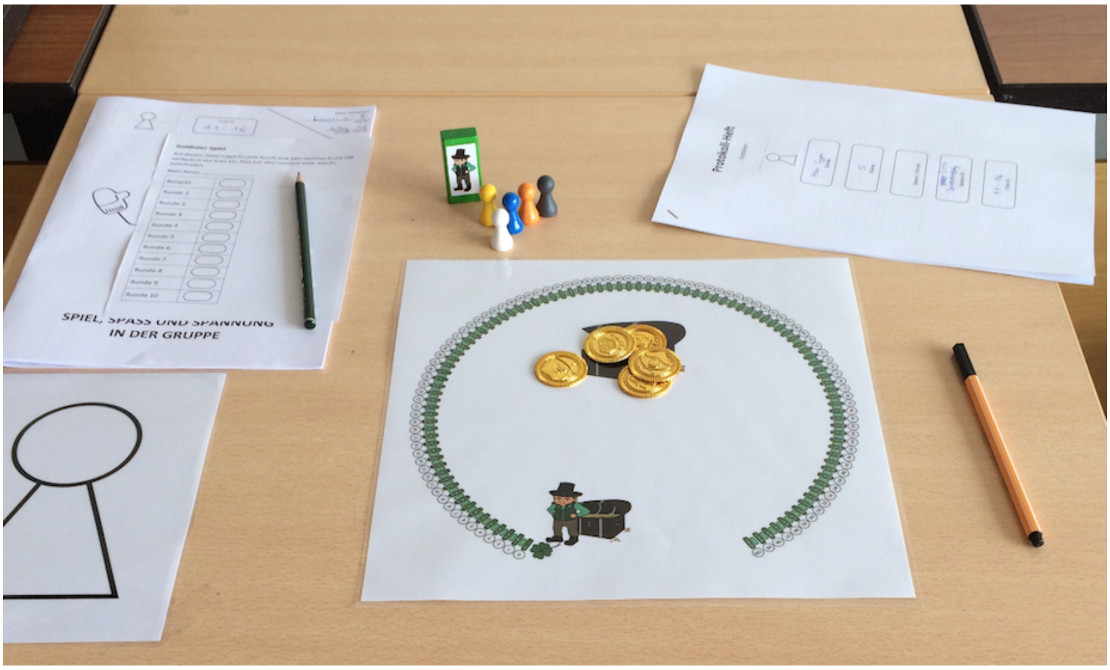
F The Beauty-Contest Game

F.1 Setup of the Game

Figure A4: Experimental Setup of the "Goblin Game"

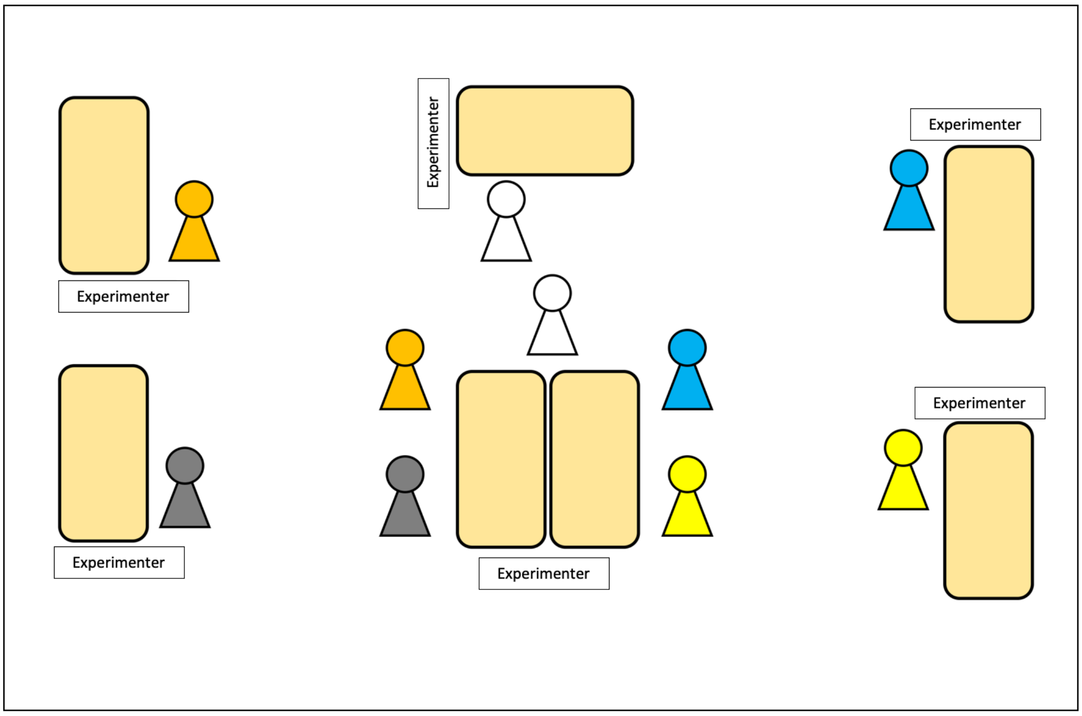
Notes: This is the setup for the main table at which the board game was played. The experimenter guiding the children through the 10 rounds of the game sat at the bottom center of the table. The five other seating positions are marked with one of the five colors (gray, orange, white, blue, and yellow). A note with a reminder about the five steps of the game is placed at every seating position. In the center, there is the actual board game with the goblin, the five colored pawns, and the treasure box filled with gold coins. The board measured 59x59 cm (about 23x23 in), so children could read everything but could also reach everything on the board. In the background, there is one of the five small tables used to explain the game to children in a one-to-one setting (see Figure A5 for details).

Figure A5: Material Used in the "Goblin Game"



Notes: This is the setup for one of the five tables at which the game was explained to the children in a one-to-one setting with an experimenter. In the center, there is a mini-version of the board game used for the explanation, including some gold coins, the green goblin figure, and the five colored pawns. At the bottom left is the indicator that this is the table for the white player (this ensured that the right child would come to the right experimenter—experimenters rotated over colors, so that they would explain the game to a different color for every group). At the top left is the game slip for the child on which he/she wrote down the numbers for each round. Below that is the workbook containing the questionnaires. At the top right is the script used by the experimenter to explain the game to the child (see Section F.2)

Figure A6: Plan for the Room in which the "Goblin Game" Was Played



Notes: This is the plan used to prepare the separate room in which the experimental BCG was played with groups of five children. The separate rooms in all schools were set up following this plan to standardize the setting for all children across schools. There were five small individual tables used to explain the game to children in a one-to-one setting with an experimenter. The main experimenter always sat at the marked position on the large table to guide the children through the game (see Section 2.3 for details about the procedures).

F.2 Instructions for Experimenters

(The following instructions were used as a script by each experimenter during the one-to-one explanation of the game. Each experimenter had these instructions in print in front of him or her and followed the script exactly, step by step; see Figure A5.)

Information for the Experimenter/Instruction Reader

- Show no reactions to the strategies/ideas/suggestions of the children. Only answer specific questions by referring to the rules.
- Example
  Child: "I should never set a figure on numbers greater than 50, right?"
  You: "You can set your figure on any number between 0 and 100. The player who is closest to the goblin at the end of the round will win."
- No explicit suggestions, including nonverbal suggestions, should be given regarding potential game strategies.
- All bold solution parts of the rules should be mentioned, missing aspects should be recorded.

Game Instructions: The Goblin Game

We are going to play a game, in which the goal is to win as many gold coins as possible. At the end of the game you can trade the gold coins for toys. The more gold coins you have at the end of the game, the wider choice of toys you will have. So, try hard! ☺

We will play in groups of five. Every player will get a pawn, you have the color [*NAME THE COLOR and point to the figure*]. There are also the colors [*name the other colors and point to the pawns*].

Other than the five players, there is the goblin [*point to the figure*]. The goblin has hidden gold coins in the forest. But he is a nice goblin and will help you find the gold coins. He always reveals the location of the gold coins to the middle player. However, the goblin is hexed: for every step he takes, he must go back half a step.

The goblin will show you the way to the gold coins. Therefore, you must be **as close as possible to the goblin** after every round. The player who is closest to the goblin after the round [*point to the figure closest to the goblin for clarification*], will win a gold coin [*point to the gold coin in the treasure box*]. If two players are equally far away from the goblin, both of these players will get a gold coin.

In total, we will play 10 rounds, so it is possible to win 10 gold coins. Look, here on your game slip you can see fields for the 10 rounds [*give the players the game slip and show them the 10 rounds*]. Write your name on the top of the page [*let them write their name*].

Good job! Every round will work in the same way, in other words, every round will contain five steps [*accordingly count the steps with your fingers*]:

Step 1:

Every player will write down a secret number between 0 and 100 on his or her paper.

For the first round, this number will be written in the circle after round 1, in the second round in the circle after round 2, and so on. When you are finished writing down your number, you should simply put your playing figure over the number. This way, the other players cannot see what you wrote down.

It is important that the others do not see what number you have written down. Maybe you can hold your hand over what you are writing. Now, write down a number between 0 and

100 in the example field! [*let the child write down a number in the example field of the game slip and write down a number yourself*].

Step 2:

When all players are done writing down a number, everybody can set their figure on the respective field. Of course, you should set your figure on the number you wrote down.

Therefore, we will use the fields on the board game [*point to the fields, so that they can distinguish the goblin fields from the player fields*]. There are fields between 0 and 100 [*point to the fields*]. Therefore, the number that you wrote down in step 1 must be between 0 and 100. Every player will set his/her figure on the respective number that he/she has written down. Let's go! [*now let the child set his / her figure on the field and set the other figures on the fields, 7, 24, 45, and 79—do not let any two figures be set on the same field; if this happens, set them aside from each other*].

Step 3:

Then, the goblin comes. He always runs from his field along the **goblin fields** [*show which ones these are*]. In order to reveal where the gold coins are hidden, he runs to the **middle player**. That is also the third player, if you count from the beginning. Now you can move the goblin! [*count out loud (1, 2, and 3), while the child moves the goblin along the fields*].

Step 4:

Now comes the goblin jump. Because the goblin is hexed, he has to **go back half a step** for every step he takes. So the goblin always jumps back to the field that is **half the number** of the field on which he originally stood. Half of each number is written down on the goblin fields [*point to the respective field, the arrow and the half value*], so that you do not have to calculate. So, the **goblin jumps back by half of the middle player's number** [*move the goblin figure and read the numbers out loud while doing so*].

Step 5:

Now we look: **Who is closest to the goblin?** This player **wins a gold coin**. If two players are the same distance from the goblin, **both** these players win a gold coin. Who wins a gold coin this time? [*let the child identify the winner*].

Very good! Then let's set all the figures back and the next round can begin. Altogether, we will play 10 rounds.

So, again to summarize:

The point of the game is to win as many gold coins as possible. The player who is closest to the goblin after each round wins a gold coin. Each round has five steps [*again count the steps with your fingers while explaining*]: First, each player secretly writes down a number. Then, each player sets his/her figure on that number. Next, the goblin runs to the third player. The goblin then jumps back by half of the number chosen by the third player. The player who is now closest to the goblin gets a gold coin. Then, the next round starts.

Do you have any questions on the rules of the game?

TESTING UNDERSTANDING OF THE GAME

We will go through the game together one more time. You get to explain the steps of the game to me. If you cannot remember something, no problem—I will gladly explain it to you again. So: [*Read out loud (without the headlines "Question X"). If necessary, explain the step* ***one more time****, otherwise don't mind and encourage the child by saying "you will surely see how it works during the game!"*]

Question 1: Can you please explain one more time what happens first in every round?

Answer: All players **secretly** write down, without letting anybody else see, a **number** between **0 and 100** on their game slip. The players **set their pawn on the secret number** that they have written down in order to cover it. This means that the players are done writing down their number.

☐ Immediately completely and correctly explained

☐ With a hint completely and correctly explained

☐ After a repeated explanation, completely and correctly explained

☐ Not completely understood, the following is missing: ...

Question 2: What happens after all players have written down their secret number?

Answer: When all players have written down their number (not before!), **the players simultaneously set their figures on the field** (on the outer fields on the board game) according to **the number written on the game slip**.

☐ Immediately completely and correctly explained

☐ With a hint completely and correctly explained

☐ After a repeated explanation, completely and correctly explained

☐ Not completely understood, the following is missing: ...

Question 3: What happens after all players have set their figures on the board?

Answer: Then the **goblin** runs over the **goblin fields** up to the **third/middle player**.

☐ Immediately completely and correctly explained

☐ With a hint completely and correctly explained

☐ After a repeated explanation, completely and correctly explained

☐ Not completely understood, the following is missing: ...

Question 4: What happens when the goblin is at the middle player?

Answer: The **goblin jumps**. This means, that he jumps back **by half of the number on the field** (the number to which the goblin must jump is indicated on his original field, behind the arrow).

☐ Immediately completely and correctly explained

☐ With a hint completely and correctly explained

☐ After a repeated explanation, completely and correctly explained

☐ Not completely understood, the following is missing: ...

Question 5: What happens after the goblin has jumped back?

Answer: The player who **is closest to the goblin** wins **a gold coin** (if two players are the same distance from the goblin, both these players win a gold coin). Then, **all figures are returned** to the players and **a new round** begins.

☐ Immediately completely and correctly explained

☐ With a hint completely and correctly explained

☐ After a repeated explanation, completely and correctly explained

☐ Not completely understood, the following is missing: ...

In Conclusion: You did a great job! The game is about to start—but first, I will give you **a gold coin** as a thank you!

G Measures for Perspective-taking Abilities

G.1 Perspective-taking Task

(This is the script used by the experimenter to explain the perspective-taking task "E on the forehead". The task is conducted right after the child has received the gold coin for the explanation of the rules of the game; see previous page. The experimenter reads the following text out loud and records the observed behavior right away.)

"Before we start, I have one small task for you.

Please trace, as fast as possible, with your forefinger, the capital letter 'E' on your forehead."

[*Repeat the instructions only once more if necessary. If the child does not understand, encourage him/her to take his/her seat at the group table.*]

The 'E' is ...

☐ From my (experimenter) perspective reversed

☐ From my (experimenter) perspective legible

☐ Neither, the child traced something else

☐ The child did not understand the task, other

"Very good job—now you can go to your spot at the group table. Have fun playing the game!"

G.2 Social Appropriateness Scale

(This questionnaire was adapted from Meindl (1998). It was filled out during the first lesson by all children in the classroom. Questions were read out loud by the experimenter and children filled out the questionnaire in an individual workbook.)

Situation 1: The Camera

Fritz met his friend Jochen on the street and showed him the new camera his parents gave him. Jochen asked Fritz if he could try the camera. While trying to take a picture with the camera, Jochen tripped. The camera fell down and broke.

Question A: How does Fritz feel when he sees that the camera is broken?

☐ 1. He is mad because his camera is broken.

☐ 2. He does not care because he will surely get another camera from his parents.

Question B: How does Jochen feel?

☐ 1. He feels guiltless, because he did not mean to break the camera on purpose.

☐ 2. He is embarrassed that he broke the camera.

Question C: How would you react if you were Fritz?

☐ 1. I would yell at Jochen because he should have been more careful with the camera.

☐ 2. I would tell Jochen that I am upset, but not mad at him, because he did not break the camera on purpose.

Question D: How would you react if you were Jochen?

☐ 1. I would apologize.

☐ 2. I would tell Fritz that he should not be mad because I did not break the camera on purpose.

Situation 2: The Computer

Jürgen wants for a computer for his birthday. However, his parents do not have enough money and give him something else instead.

Question A: How does Jürgen feel when he sees that he did not get a computer?

☐ 1. He does not mind because he received something else instead.

☐ 2. He is disappointed.

Question B: How does not fulfilling Jürgen‘s wish feel to his parents?

☐ 1. They feel sorry that they cannot fulfill his wish.

☐ 2. They do not care because they would not have had enough money to buy the computer anyway.

Question C: How would you react if you were Jürgen?

☐ 1. I would complain loudly to my parents that I would have rather had a computer.

☐ 2. I would try not to show my disappointment and rejoice over the other gift I got.

Situation 3: The Horror Film

Susanne would like to watch a horror film later in the evening. However, her father does not allow this and sends her to bed, with the reasoning that she is still too young.

Question A: What does her father think, when he says that Susanne is still too young?

☐ 1. He thinks that if Susanne watches the film, she would get very scared.

☐ 2. He wants to upset Susanne.

Question B: How does not being allowed to watch the film make Susanne feel?

☐ 1. She feels sad.

☐ 2. She feels mad.

Question C: What would you do if you were Susanne?

☐ 1. I would yell at my father because he is so mean to me.

☐ 2. I would try to talk to him about it again.

Situation 4: The Dishes

Sebastian is about to leave the house because he has arranged to play soccer with his friends. However, his mother asks him to wash the dishes as she still has a lot to do.

Question A: How does Sebastian feel when he hears that he should wash the dishes?

☐ 1. He is sad.

☐ 2. He is mad.

Question B: How does his mother feel?

☐ 1. She is stressed from all the work.

☐ 2. She just does not want Sebastian to go and play soccer.

Question C: What would you do if you were Sebastian?

☐ 1. I would go and play soccer because it was already arranged.

☐ 2. I would wash the dishes and then go to soccer later.

Situation 5: Teasing

Markus constantly gets teased by his classmates because he stutters. Doris joins the class as a new student. She notices that Markus gets teased by everybody and also joins in the teasing.

Question A: Why does Doris tease Markus?

☐ 1. Because Markus stutters.

☐ 2. Because she wants to be accepted by the others.

Question B: How does Markus feel when he gets teased?

☐ 1. He is sad and feels excluded.

☐ 2. He does not take it seriously.

Question C: How would you react if you were Doris?

☐ 1. I also would have joined in the teasing.

☐ 2. I would have refrained from teasing.

Situation 6: The Best Grade

After school, Michael tells his friend Peter that he got an A in math. Peter, who got a C, says to Michael: "You're a stupid nerd".

Question A: Why does Peter say this?

☐ 1. He is envious of Michael.

☐ 2. He does not like Michael.

Question B: How does Michael feel thereafter?

☐ 1. He is hurt, because Peter offended him.

☐ 2. He finds Peter's behavior ridiculous.

Question C: How would you react if you were Michael?

☐ 1. I would tell Peter that he should not exaggerate like that.

☐ 2. I would tell Peter that I am sorry that he only got a C, but that he does not need to offend me.

G.3 Interpersonal Reactivity Index

(This questionnaire uses the items for cognitive empathy from Garton and Gringart (2005). It was filled out during the first lesson by all children in the classroom. Questions were read out loud by the experimenter and children filled out the questionnaire in an individual workbook.)

1. I think people can have different opinions about the same thing.

☐ **does not** apply **at all**

☐ **does not generally** apply

☐ **sometimes** applies

☐ **generally** applies

☐ **fully** applies

2. When I am angry or upset at someone, I usually try to imagine what he or she is thinking or feeling.

☐ **does not** apply **at all**

☐ **does not generally** apply

☐ **sometimes** applies

☐ **generally** applies

☐ **fully** applies

3. When I am arguing with my friends about what we are going to do, I think carefully about what they are saying before I decide whose idea is best.

☐ **does not** apply **at all**

☐ **does not generally** apply

☐ **sometimes** applies

☐ **generally** applies

☐ **fully** applies

4. I sometimes try to better understand my friends by pretending I am them.

☐ **does not** apply **at all**

☐ **does not generally** apply

☐ **sometimes** applies

☐ **generally** applies

☐ **fully** applies

H References Online Appendix

Davis, Mark H. 1983. "Measuring individual differences in empathy: Evidence for a multidimensional approach." *Journal of Personality and Social Psychology* 44 (1): 113–126.

Galinsky, Adam D., Joe C. Magee, M. Ena Inesi, and Deborah H Gruenfeld. 2006. "Power and Perspectives Not Taken." *Psychological Science (0956-7976)* 17 (12): 1068–1074.

Garton, Alison F., and Eyal Gringart. 2005. "The Development of a Scale to Measure Empathy in 8- and 9-Year Old Children." *Australian Journal of Educational & Developmental Psychology* 5:17–25.

Glen, R. 1984. "Perspective taking and self-awareness: Drawing an E on your forehead." *Journal of Personality and Social Psychology* 46 (4): 788–798.

Meindl, C. 1998. "Entwicklung und Validierung eines Fragebogens zur Erfassung von Empathie und seine Anwendung im Rahmen eines Empathietrainings." *mimeo.*

Schick, Andreas, and Manfred Cierpka. 2005. "Faustlos: Evaluation of a curriculum to prevent violence in elementary schools." *Applied and Preventive Psychology* 11 (3): 157–165.

Steins, Gisela, and Robert A. Wicklund. 1996. "Perspective-Taking, Conflict, and Press: Drawing an E on Your Forehead." *Basic and Applied Social Psychology* 18 (3): 319–346.