Supplemental Experiment

In Experiment 1, only 9-year-olds convincingly applied the law of comparative advantage to their strategic division of cognitive labor. However, even the 6-year-olds in the study applied a strategy – one in which they always answer the easiest questions. In this supplemental experiment, we report a study designed to test whether this behavior could still be consistent with the law of comparative advantage.

There will be situations where the optimal solution by the law of comparative advantage is to assign easier questions to the skilled partner, opposite what we expected children to do in Experiment 1. This solution will occur whenever both partners have an approximately equal chance of succeeding on the harder option, but the skilled partner has a better chance of succeeding on the easier option. For instance, let's say a child estimates their chance of success on the easier question to be 60% and their less skilled partner's chances on that question to be 50% (i.e., guessing), but estimates that both partners would be guessing at chance on the harder question in the pair. By the law of comparative advantage, it would be *more* beneficial for the team if the child answers the easier question (increasing the chances of success by 10%) than if they answered the hard question (where there would be no increase).

If this is what children are thinking, then why did only younger children show this alternate solution? Children improve at the numerical discrimination task in this age range (e.g., Halberda & Feigenson, 2008). Children with good metacognition would then very reasonably estimate lower chances of success in general than the older 9-year-olds. Therefore, the younger children in our study may have believed they were guessing despite being the 'better' partner, simply because their number sense is still developing.

Method

Participants

One hundred and eighty-nine children participated in the study (90 girls) between May 2017 and June 2018 in the same manner and geographic location as the other experiments. We primarily recruited 6 and 7-year-old children (with a target sample size of 20 children, but also tested 8-year-olds because of over-recruitment at schools (M = 7;1, range = 6;0 - 8;7). One additional child was excluded from the analyses because of a teacher-reported learning disability.

Materials and Procedures

The procedure was identical to Experiment 1. Children saw 10 warm-up trials of the Number Game, including two impossible trials. Children heard that their partner was either better or worse at the practice questions than they were, and then were asked to divide 14 questions between the partners. At the end of the study, children answered a comprehension question about who was better at the game.

To test if children's strategies reflected a flexible law of comparative advantage, we adapted the stimuli from Experiment 1 to create both 'easy' and 'hard' versions of the task. Easy and hard versions were matched in terms of metaratio (using metaratios 2.65, 2.0, and 1.5 in both versions), so we expected children to be equally capable of discriminating the difficulties in both versions. The critical difference was in the ratios used to create each pair. On easy trials, we used two larger ratios: e.g., a very easy ratio 4 (40 yellow dots for every 10 blue ones) against an easy ratio 2.0, for a metaratio of 2.0 (other pairs were 4.0 vs. 1.5 and 3.0 vs. 2.0). On hard trials, we

used two smaller ratios, like ratio 2.0 against ratio 1.0 (i.e., impossible, other pairs were 3.0 vs. 1.13 and 2.0 vs. 1.33). We hypothesized that if children were accurately predicting their own chances of success, the optimal strategy would flip from assigning the easier question to the less skilled partner (when the chances of success are generally high) to assigning the easier question to the more skilled partner (when the chances of success are generally low).

Further, we tried to increase children's perceived chances of success by boosting their global self-confidence. We reasoned that with an increase in global self-confidence, children might estimate a higher chance of success overall, even on the harder item pairs. With a higher estimate of success, they should then assign harder items to the better partner, as we initially expected.

In one recent study with adults, global self-confidence was increased by providing feedback or by presenting easy questions (Rouault et al., 2019). As we did not want to provide feedback during the task in case it was misinterpreted by children, we decided to organize the easy and hard ratio pairs described above into two blocks. Half of children saw the easy trials first, meant to build up confidence in their ability so that they might estimate higher chances of success on the subsequent hard trials. We expected that these children would assign harder questions to the better of the two partners for both the easy and hard blocks, consistent with 9-year-olds in Experiment 1. In contrast, the other half of children saw the hard trials first, which served as a replication of Experiment 1 where children assigned themselves the easier items. Then, these children saw the easier trials in the second half, for which we expected children to switch their strategy to assign harder questions to the better partner.

Results

First, we examined children's post-test comprehension answers to confirm that the skill manipulation worked. Nearly all children, 91%, passed the comprehension check at the end of the study. Unlike in Experiment 2, this was not dependent on condition, $p = .400$, Fisher's Exact Test (two-tailed). Next, we confirmed that all children understood the Number Game. Children selected the color with more dots above chance of 50%, shown in Table S1.

We conducted an ANOVA with Condition (Partner Better, Partner Worse), Age (6, 7, 8), Trial Difficulty (Easy, Hard), and Order (Easy First, Hard First). The only effect that approached significance was a main effect of Age, $F(2, 177) = 2.47$, $p = .087$, $\eta_p^2 = 0.03$, where older children were generally less likely to assign themselves the easier question (see Figure S1). All other main effects and interactions were non-significant with small effect sizes. We saw no changes from adding easier trials in either order.

Six and seven-year-olds consistently assigned themselves the easier question on most trials in both conditions. Full details about the means and tests against can be found in Table S1, and are shown in Figure S1.

Discussion

Replicating Experiment 1, children generally assigned themselves the easier questions regardless of their partner's skill. This behavior was not modified by the difficulty of the questions, nor by the order in which difficult questions were presented. This study also confirmed that children did remember the relative skill level of their partner, so results cannot be attributed to forgetting or not understanding the skill difference.

While children didn't modify their strategy in response to changes in difficulty, these findings do not rule out an understanding of the law of comparative advantage. The easy/hard

manipulation requires that children attend to their *absolute* likelihood of success on each question (e.g., estimating 60% success). The division of labor task instead emphasizes the *relative* differences in difficulty (this one has a *higher* chance of success). Children at this age are generally overconfident and therefore their absolute judgments of success are not well calibrated to reality (Hagá & Olson, 2017; van Loon et al., 2017). However, children show great precision at differentiating their *relative* chances of success (Baer & Odic, 2019). Therefore, they might not be sufficiently calibrated to display this sophisticated application of the law of comparative advantage, but might still understand the basic principles.

Figure S1

Percentage of Trials Assigning the Easier Question to Self in the Supplemental Experiment

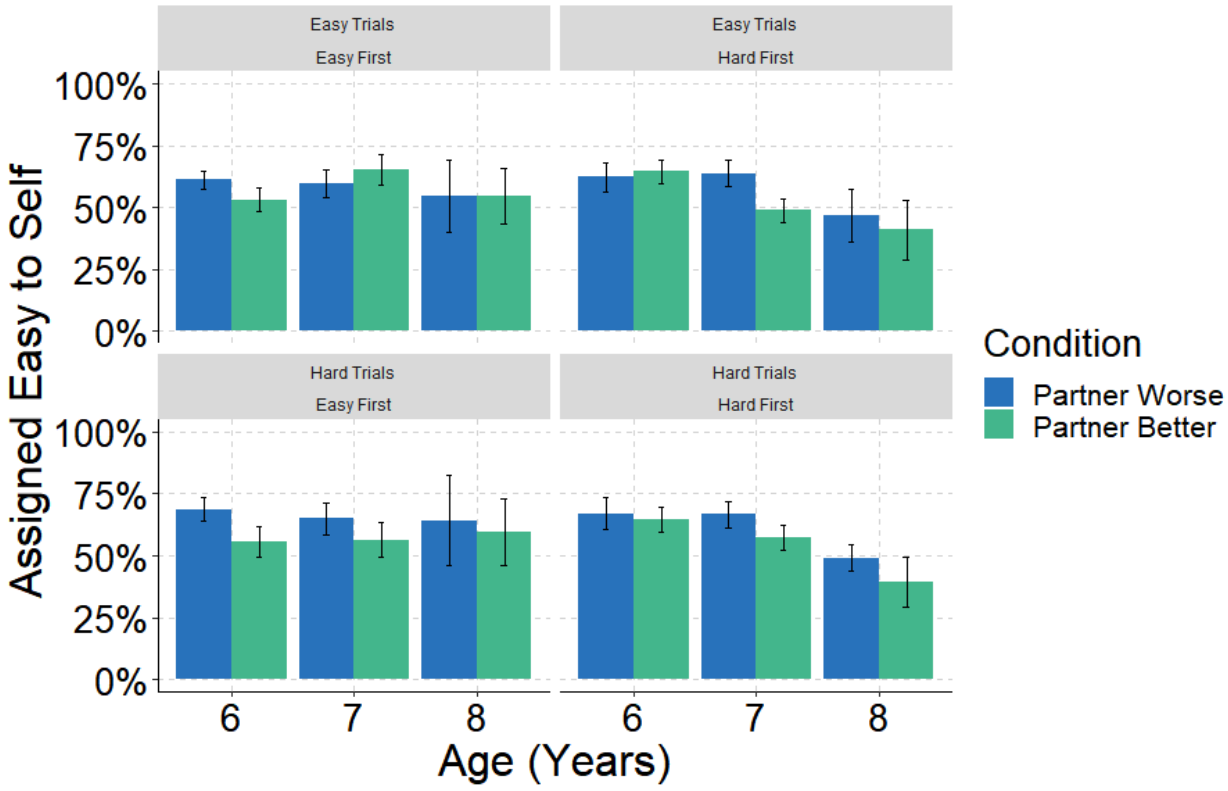


Table S1

Predictions, Means, and Tests Against Chance for the Supplemental Experiment.

	First Block							Second Block						
Age	Prediction	Mean	SD	df	*t*	*p*	*d*	Prediction	Mean	SD	df	*t*	*p*	*d*
Number Game (chose more numerous)														
6		82.70	14.04	78	20.69	< .001	2.33							
7		87.15	9.50	82	35.62	< .001	3.91							
8		88.27	10.40	26	19.47	< .001	3.68							
Division (gave easier to self)														
Partner Better Condition (Hard Trials First)														
6	>50	66.92	28.19	18	2.62	.018	0.60	>50	62.41	25.29	18	2.14	.046	0.49
7	>50	66.67	23.62	20	3.23	.004	0.71	>50	63.95	23.74	20	2.69	.014	0.59
8	>50	48.98	13.94	6	-0.19	.853	0.07	>50	46.94	28.23	6	-0.29	.784	0.11
Partner Worse Condition (Hard Trials First)														
6	>50	64.63	22.87	20	2.93	.008	0.64	<50	64.63	22.87	20	2.93	.008	0.64
7	>50	57.14	21.82	18	1.43	.171	0.33	<50	48.87	19.82	18	-0.25	.807	0.06
8	>50	39.29	28.32	7	-1.07	.320	0.38	<50	41.07	34.52	7	-0.73	.488	0.26
Partner Better Condition (Easy Trials First)														
6	>50	61.22	16.38	20	3.14	.005	0.69	>50	68.71	21.49	20	3.99	.001	0.87
7	>50	59.74	26.32	21	1.74	.097	0.37	>50	64.94	29.48	21	2.38	.027	0.51

8	>50	54.76	35.48	5	0.33	.756	0.13	>50	64.29	44.95	5	0.78	.471	0.32
Partner Worse Condition (Easy Trials First)														
6	<50	53.17	20.09	17	0.67	.512	0.16	<50	55.56	25.41	17	0.93	.367	0.22
7	<50	65.31	28.06	20	2.50	.021	0.55	<50	56.46	31.45	20	0.94	.358	0.21
8	<50	54.76	27.73	5	0.42	.691	0.17	<50	59.52	33.09	5	0.70	.512	0.29

Supplemental Analyses

Age as a Continuous Predictor

In our main analyses, we treat age as a categorical predictor, enabling power for tests against chance alongside clear statements about the ages at which children succeed at this task. Here, we confirm that our analyses are robust to treating age as a continuous predictor.

In each of the following analyses, we conduct analogous ANOVAs to those in the main analyses, but with age as a continuous, centered predictor.

Experiment 1. Children were more likely to take the easier question when they were less skilled than their partner than when they were better, $F(1, 155) = 11.02$, $p = .001$, $\eta_p^2 = .07$. As in the main analysis, this was difference depended on children's age, $F(1, 155) = 10.86$, $p = .001$, $\eta_p^2 = .07$, the details of which can be found in the main text. We further find a main effect of Metaratio, $F(5, 755) = 4.53$, $p < .001$, $\eta_p^2 = .03$, which interacted with Age, $F(5, 755) = 2.56$, $p = .026$, $\eta_p^2 = .02$, and with Age and Condition, $F(5, 755) = 3.15$, $p = .008$, $\eta_p^2 = .02$. There was no main effect of age nor an interaction between Metaratio and Condition.

Experiment 2. Children assigned themselves the easy question more often when their partner was older, $F(1, 162) = 5.73$, $p = .018$, $\eta_p^2 = .03$. Children also assigned themselves the easy question more often as they got older, $F(1, 162) = 5.81$, $p = .017$, $\eta_p^2 = .03$. In this analysis, we also found a significant interaction between Age and Condition, $F(1, 162) = 5.13$, $p = .025$, $\eta_p^2 = .03$, details in the main text. We found no main effect or interaction with Metaratio.

Experiment 3. Older children were significantly more likely than younger children to match skill and difficulty, $F(1, 78) = 15.94$, $p < .001$, $\eta_p^2 = .17$.

First Trial Only

Children saw 14 trials in each of the experiments, which afforded us more power to detect effects. But this may have been tiring for children particularly given that there was no feedback during this phase of the study. This creates a second contrast between our work and that of Magid and colleagues (2018), who used a single trial per child on their physical division of labor task. At the suggestion of a reviewer, we therefore examined just the first trial performance of children in our studies. It also happens that this trial was always the easiest metaratio, so we should expect to see the best possible performance on this trial for all age groups.

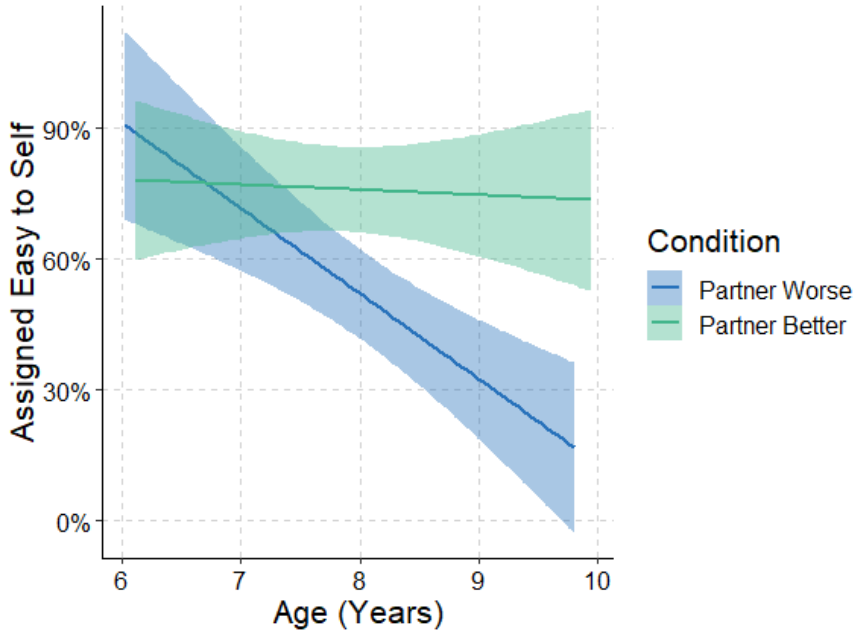
In each of the following analyses, we conducted a logistic regression on choice of the easier trial for self (chose easy = 1, chose hard = 0) with Age as a centered continuous predictor (all experiments) and Condition (Partner Better/Older = 1, Partner Worse/Younger = 0, for Experiments 1 and 2). We thank an anonymous reviewer for suggesting this analysis to us.

Experiment 1. On the very first trial of the experiment, children were more likely to choose the easy question for themselves when their partner was better, OR (Odds Ratio) = 2.74, Wald $\chi^2(1) = 7.7$, $p = .005$. We also saw an effect of age, where older children were more likely to choose hard trials, OR = 0.41, Wald $\chi^2(1) = 12.5$, $p < .001$. Because the Partner Worse

condition is the reference group in this regression, this is largely a reflection of the strong age effect in this condition (the effect is nonsignificant when using Partner Better as the reference condition). This points to the interaction between Age and Condition, replicating what we found when examining all 14 trials, OR = 2.28, Wald $\chi^2(1) = 5.6$, $p = .018$. Average marginal effects revealed a decrease in taking the easy question when paired with a less skilled partner as children got older, AME = -.18, SE = .03, $z = -5.45$, $p < .001$. In contrast, there was no change with age when the partner was better, AME = -.01, SE = .04, $z = -0.27$, $p = .788$. This matches the findings reported in the main analyses (and see Figure S2).

Figure S2

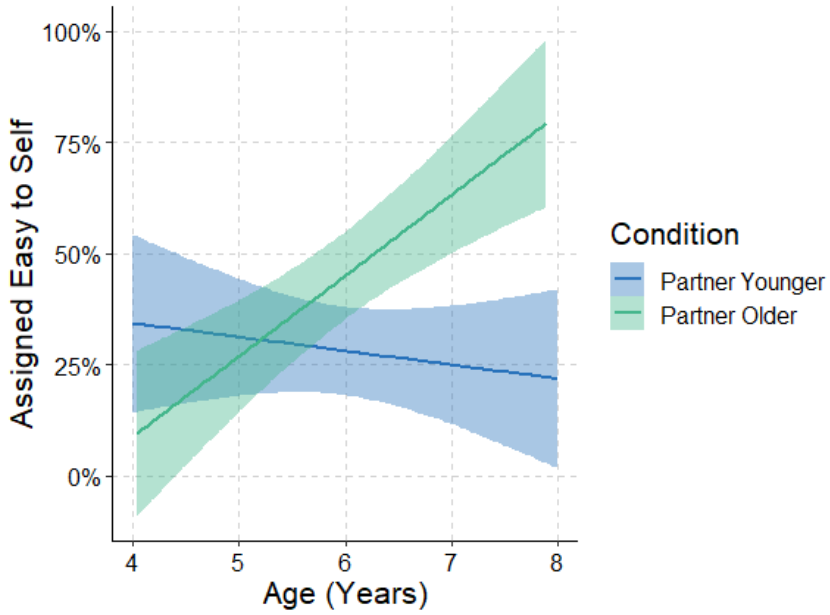
Percentage of Easy Trials Assigned to the Self on the First Trial Only in Experiment 1



Experiment 2. The main effect of condition was on the border of conventional significance levels, OR = 1.95, Wald $\chi^2(1) = 3.7$, $p = .056$, and there was no effect of age, OR = 0.85, Wald $\chi^2(1) = 0.51$, $p = .470$. As in Experiment 1, this is largely a reflection of the reference group, as the age effect was present when using Partner Older as the reference condition. This again points to an interaction, OR = 2.75, Wald $\chi^2(1) = 10.00$, $p = .002$, see Figure S3. The average marginal effects revealed a that older children with an older partner were more likely to assign themselves the easy question than younger children with an older partner, AME = .17, SE = .03, $z = 5.57$, $p < .001$. When children had a younger partner, there was no effect of age, AME = -.03, SE = .04, $z = -.072$, $p = .470$. The presence of this interaction on this first trial, when it did not reach significance in the main analysis, gives some credence to the idea that performance may have waned throughout the study. With a one-shot decision, children may demonstrate more competence on this division of labor task.

Figure S3

Percentage of Easy Trials Assigned to the Self on the First Trial Only in Experiment 2



Experiment 3. Replicating the full analysis of all 14 trials, older children were more likely to correctly pair skill and difficulty than younger children, OR = 1.82, Wald $\chi^2(1) = 7.30$, $p = .007$.

Figure S4

Percentage of Easy Trials Assigned to the Unskilled Partner on the First Trial Only in Experiment 3