Article

Children's arithmetic skills do not transfer between applied and academic mathematics

https://doi.org/10.1038/s41586-024-08502-w

Received: 8 March 2022

Accepted: 5 December 2024

Published online: 05 February 2025

Open access

Check for updates

Abhijit V. Banerjee[1]✉, Swati Bhattacharjee[2], Raghabendra Chattopadhyay[3], Esther Duflo[1]✉, Alejandro J. Ganimian[4], Kailash Rajah[1] & Elizabeth S. Spelke[5]✉

Many children from low-income backgrounds worldwide fail to master school mathematics[1]; however, some children extensively use mental arithmetic outside school[2,3]. Here we surveyed children in Kolkata and Delhi, India, who work in markets (n = 1,436), to investigate whether maths skills acquired in real-world settings transfer to the classroom and vice versa. Nearly all these children used complex arithmetic calculations effectively at work. They were also proficient in solving hypothetical market maths problems and verbal maths problems that were anchored to concrete contexts. However, they were unable to solve arithmetic problems of equal or lesser complexity when presented in the abstract format typically used in school. The children's performance in market maths problems was not explained by memorization, access to help, reduced stress with more familiar formats or high incentives for correct performance. By contrast, children with no market-selling experience (n = 471), enrolled in nearby schools, showed the opposite pattern. These children performed more accurately on simple abstract problems, but only 1% could correctly answer an applied market maths problem that more than one third of working children solved (β = 0.35, s.e.m. = 0.03; 95% confidence interval = 0.30–0.40, P < 0.001). School children used highly inefficient written calculations, could not combine different operations and arrived at answers too slowly to be useful in real-life or in higher maths. These findings highlight the importance of educational curricula that bridge the gap between intuitive and formal maths.

Maths curricula taught in primary school ought to provide children with the concepts and skills that they need both for their daily lives and as a foundation for learning the higher maths required to succeed in school at more advanced levels. Too often, however, formal schooling fails to achieve either of these goals. In India, in 2023, only half of the children enrolled in grades 11 and 12 (16–18 years of age) could divide a three-digit number by a single-digit number[4]. Globally, learning outcomes remain poor despite large increases in school enrolment in many low-to-middle-income countries[1,5].

Moreover, children seem even less likely to be able to use basic arithmetic skills in everyday life. A recent study in India[4], for example, found that only half of the children enrolled in grades 11 and 12 could calculate how many purifying tablets to use in a large pot after they were given the number they needed for a smaller pot. Notably, 35% of the children who were able to solve an abstract division problem failed this verbal exercise[4].

However, many children in low-to-middle-income countries, for instance, those who work in markets, seem to routinely perform more involved calculations as part of their daily jobs. For example, a study in the 1980s of five children who worked as street vendors (mean age of 11 years) in Brazil found that these children ably and flexibly used maths in their work[2]. These findings are in line with ethnographic studies of adults with minimal formal education who also exhibit these abilities[6–8]. Moreover, decades of research[9–12], mostly in Western countries, have shown that children can adeptly combine small sets to create exact larger numbers well before schooling begins[13] and that their facility of learning maths in primary school is associated with their sensitivity to approximate number in later years of schooling[14]. A prominent theory in cognitive psychology suggests that learning maths in meaningful real-world contexts can complement maths learnt in the classroom and provide children with more generalizable and flexible arithmetic skills[15,16]. By learning and practising maths through relevant contexts, children may be better equipped to acquire the flexible cognitive skills they need to transfer maths knowledge across domains.

Under this hypothesis, children who work in markets might be expected to more easily learn the abstract maths taught in school. However, just as children who master abstract maths skills in school may fail to apply them to concrete problems, schools may fail to help children who adeptly use maths concepts in concrete situations in their daily lives to develop more abstract, generalizable skills. In the absence of such help, such skills may not automatically transfer across domains in either direction.

Here we ask whether, in the urban Indian context, the arithmetic skills that are used in market transactions transfer to the more abstract maths

[1]Department of Economics, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Ananda Bazar Patrika, Kolkata, India. [3]Indian Institute of Management, Kolkata, India. [4]Steinhardt School of Culture, Education, and Human Development, New York University, New York, NY, USA. [5]Department of Psychology, Harvard University, Cambridge, MA, USA. ✉e-mail: banerjee@mit.edu; eduflo@mit.edu; spelke@wjh.harvard.edu

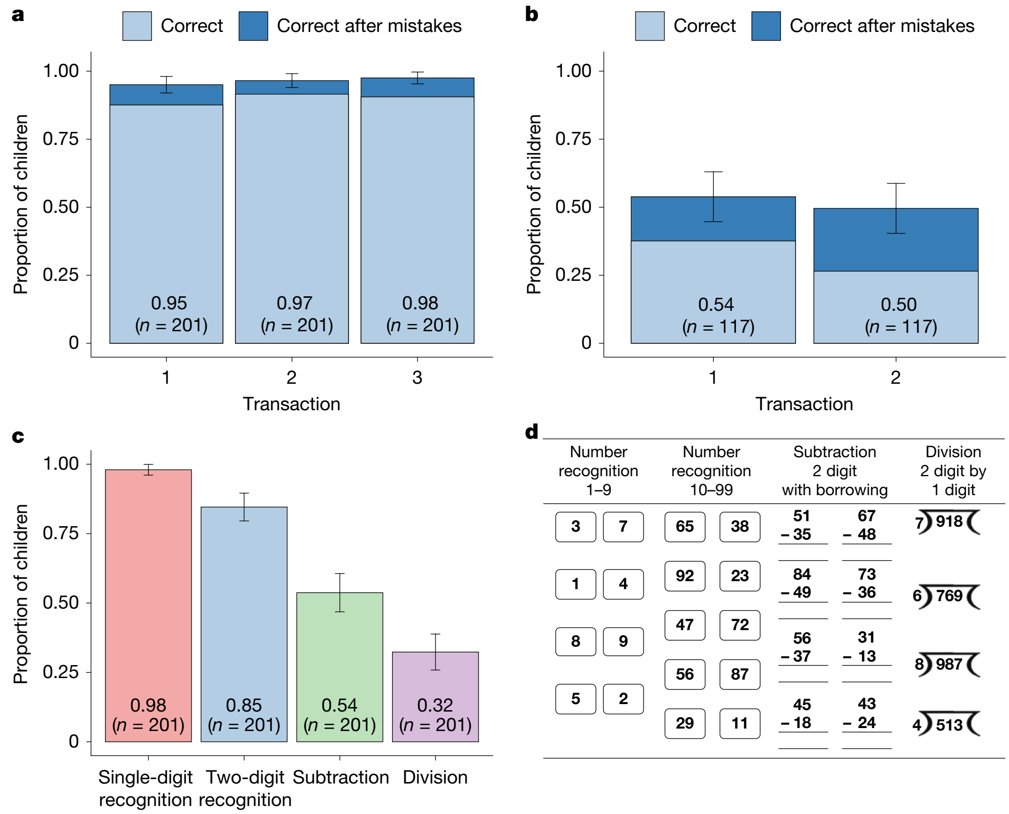


Fig. 1 | Performance of children in Kolkata working in a market (study 1). **a**, Proportion of children who correctly answered the total amount due in transactions involving two goods sold in unusual quantities. **b**, Proportion of children who correctly answered the total amount due in hypothetical transactions. **c**, Proportion of children who were credited with labelling single-digit numbers, labelling two-digit numbers, subtracting and dividing on the ASER test, a tool for assessing numeracy used across India for the ASER. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). **d**, An example of this arithmetic assessment tool. From left to right, each column presents items that respectively task children with identifying single-digit numbers, identifying two-digit numbers, subtracting one two-digit number from another with carrying, and long division of a three-digit number by a single-digit number with a remainder. For grade 5 children, the ASER test begins with the subtraction task and proceeds to the right (division) if children succeed and to the left (identifying two-digit numbers) if they fail. Because success at each level requires mastery of the operations required by the tasks to its left, children who succeed at a given level are credited with mastery of all the levels below it.

skills taught in school. Markets provide a particularly appropriate setting to ask this question because children who work in markets regularly perform maths in real-world activities that are culturally relevant and important to their lives. Moreover, these children are sufficiently engaged in these activities to be able to acquire meaningful maths competencies while often regularly attending school. We also ask whether children who attend school but without real-world experience of using arithmetic skills can translate the maths taught in schools to perform arithmetic operations in more concrete situations such as market transactions.

We surveyed a large number of children in urban India who were either attending or have attended school for years while also selling goods in markets. These children spent part of the day at school and part of the day at the market and therefore received exposure to arithmetic operations in both settings. In these studies, we analysed working children's ability to use maths to compute payments as part of their daily job and their ability to generalize their knowledge, gained either in school or in the market, to new problems. We also surveyed large numbers of school children with no market-selling experience, which enabled us to benchmark working children's performance and to test whether abstract maths knowledge by itself transfers across domains.

Working children effectively perform arithmetic operations

In this study (study 1), we tested whether children working in markets are able to perform complex arithmetic calculations as part of their daily jobs. We surveyed 201 children in Kolkata, West Bengal, working in markets (see the section 'Study 1' of the Methods). Undercover enumerators purchased unusual quantities of two goods from each child working at the store and paid with a bill that was worth more than the required amount. For example, the enumerators might ask how much 800 g of potatoes that the child sells at 20 rupees per kilogram costs and then how much 1.4 kg of onions that the child sells at 15 rupees per kilogram costs. They would then ask for the total cost (37 rupees in this case), before handing the child a 200 rupee note and collecting the change. Each of these steps was designed to measure the children's applied arithmetic skills. Afterwards, the enumerators revealed their identity and asked the participants to complete a set of abstract maths exercises.

Most of these children effectively used arithmetic calculations in their jobs. Figure 1a presents the proportion of children who correctly reported the amount due and change in the three enumerator transactions on both their first and second try (if incorrect on the first try). Overall, 95%, 97% and 98% of children were correct by their second try for the first, second and third transactions, respectively (Fig. 1a and Extended Data Table 1). Most children mentally performed these calculations with no paper aids.

To determine whether these children used mental arithmetic to solve market transactions or had learnt many of the costs for familiar quantities by rote for the items they sold, the enumerators purchased unusual quantities and not round amounts. Moreover, in a subsequent interview, we presented children with a set of two hypothetical market maths problems of similar complexity to those of the market transactions but with different items and prices. Paper and pencil aids were made available to all the children.

Our results suggest that the arithmetic skills acquired in markets extended beyond rote memorization and were transferable to new problems within the realm of market transactions, but with some loss. On average, across both transactions, 52% of the children correctly solved the hypothetical transaction either on the first try or after self-correction despite the challenges of retaining new information about unfamiliar goods, prices and units (Fig. 1b and Supplementary Table 1). They did so without help, without the use of calculators and without using pen and paper.

However, despite their arithmetic competencies, these children struggled with school maths presented in abstract form. We presented the children with problems from India's Annual Status of Education Report (ASER), which classifies children according to their oral answers to written problems of single-digit division with a remainder (taught in grade 4), double-digit subtraction with a remainder (taught in grade 2), recognition of two-digit numbers (taught in grade 1) and recognition of single-digit numbers (taught in kindergarten), administered with paper aids (Fig. 1d). As shown in Fig. 1c, when these children were given the ASER test, only 32% of children could solve a division of a three-digit number by a single-digit number, and 54% of children could solve two subtractions of a single two-digit number from another (Fig. 1c and Supplementary Table 2). This was despite the fact that only 13% of children reported that they had left school before the end of grade 4, when division is taught (Supplementary Table 3), and almost all of them attended grade 2, when subtraction is taught. The performance of these children was similar to the ASER test performance of a representative sample of children who live in rural areas of West Bengal, where 29% of children in grade 5 could solve similar division problems[17]. This result suggests that children who work in markets are not differentially selected on school arithmetic skills.

The marked gap between the strong performance of the working children on real and hypothetical market transactions and their weak performance on the written exercises presented in school is unlikely to be explained by the difficulty of the underlying arithmetic calculations. If anything, the operations required by the market transactions were more difficult than those of the written arithmetic problems on the ASER, as the market transactions involved several operations.

Performance on abstract arithmetic problems remained poor when arithmetic problems with the same structure were orally presented. Only 3% of children could solve a new division problem like those on the ASER test when orally presented, and 43% could solve orally presented subtraction problems after two attempts (Supplementary Table 4). Thus, participants performed worse in school maths problems not just because problems were presented in written form[3] but because of the abstract presentation in itself.

This analysis suggests that children who have flexible arithmetic skills fail to apply them when the problems are presented in an abstract form similar to what is presented in school. Framing arithmetic problems as they are presented in school led them astray, just as an inappropriate or new framing sometimes confounds adults with knowledge of sophisticated maths[6,18].

Working children struggle with unfamiliar problems

We next tested a new group of children from a different region (study 2a), both to examine the generality of the results from study 1 and to assess further potential confounds. In Delhi, we surveyed 400 children who worked among 39 markets. As in Kolkata, we performed three undercover purchases in addition to the survey (see the section 'Study 2' in the Methods and Extended Data Table 2).

The findings in study 2a were consistent with those of study 1. The children we questioned in Delhi correctly calculated the amount due and change on the three successive market maths problems. Overall, 96%, 99% and 97% succeeded in the first, second and third transactions, respectively, by their second try (Fig. 2a and Supplementary Table 5). Only one child used a calculator and 2% received help from others at their store. No child made written calculations (Fig. 3a and Supplementary Table 6). Because each calculation involved four operations, these findings imply both high average accuracy on each operation and excellent working memory. However, like their counterparts in Kolkata, the children in Delhi performed poorly on abstract school maths questions. Only 15% could correctly perform division on the written ASER test (Fig. 2c and Supplementary Table 7), a result similar to published ASER results for Delhi[19], where 18% of children correctly performed division. Similarly, only 11% could do so when the problems were orally presented (Supplementary Table 8).

To test whether children who work performed better on the market maths problems because they were incentivized by market pressure, we randomly assigned children to receive either a fixed payment or a monetary incentive for correctly answering the problems. We did not find evidence that incentives affected their performance ($\chi^2(4, n = 400) = 2.01$, $P = 0.73$; Extended Data Table 3).

To understand more precisely the effect of familiarity on the ability to conduct arithmetic calculations, we asked children five hypothetical market maths problems that became progressively more complex and further removed from their activity in the market that day. Performance declined for problems that required children to hold more new information in mind and to perform a sequence of operations. After one chance to self-correct for initial mistakes, 96% of these children correctly calculated the amount due in a transaction that involved the same goods that they sold at the same price that they sold it, and 88% did so when the price was different ($\beta = 0.09$, s.e.m. = 0.02, 95% confidence interval (CI) = 0.05–0.12, $P < 0.001$; Supplementary Table 9). Children's performance declined when a second item was added at a different price (69% correct, $\beta = 0.19$, s.e.m. = 0.03, 95% CI = 0.14–0.25, $P < 0.001$; Supplementary Table 9), which required multiple operations and provided more opportunities for errors. Performance declined even further when the problems involved two goods sold using an unfamiliar pricing scheme, for example, by units rather than kilograms or vice versa (61% correct, $\beta = 0.08$, s.e.m. = 0.03, 95% CI = 0.01–0.14, $P = 0.02$; Fig. 2b and Supplementary Table 9).

These declines in performance were accompanied by changes in the approach used by the children to solve the problem, with more reliance on pen and paper and techniques taught in school. When they worked with an item sold at their store, only 4% of children used pen and paper. As the transactions became less familiar and involved more operations, children increasingly relied on pen and paper: 29% with two goods sold in familiar units ($\beta = -0.25$, s.e.m. = 0.02, 95% CI = −0.30 to −0.20, $P < 0.001$; Extended Data Table 4), and 41% when it involved two goods sold in unfamiliar units ($\beta = -0.12$, s.e.m. = 0.03, 95% CI = −0.18 to −0.05, $P < 0.001$; Extended Data Table 4 and Supplementary Table 10). Finally, when the problems were presented with no concrete context, as are oral arithmetic operations in school, 58% of them used pen and paper (Fig. 3a).

School children perform poorly on concrete problems

Good arithmetic performance in market maths problems did not translate to good performance in school-like problems. Therefore, we asked whether proficiency in maths taught in schools transfer to real-world situations (study 2b). To that end, we surveyed 200 children attending 20 government schools in Delhi, located in the same zones as the markets from study 2a. These children, matched in age to the children working in markets in Delhi, were given the same written and verbal arithmetic problems as the working children, including the same school maths tests. We also created a play-based pretend market, in which the school children sold items to the enumerator, who presented

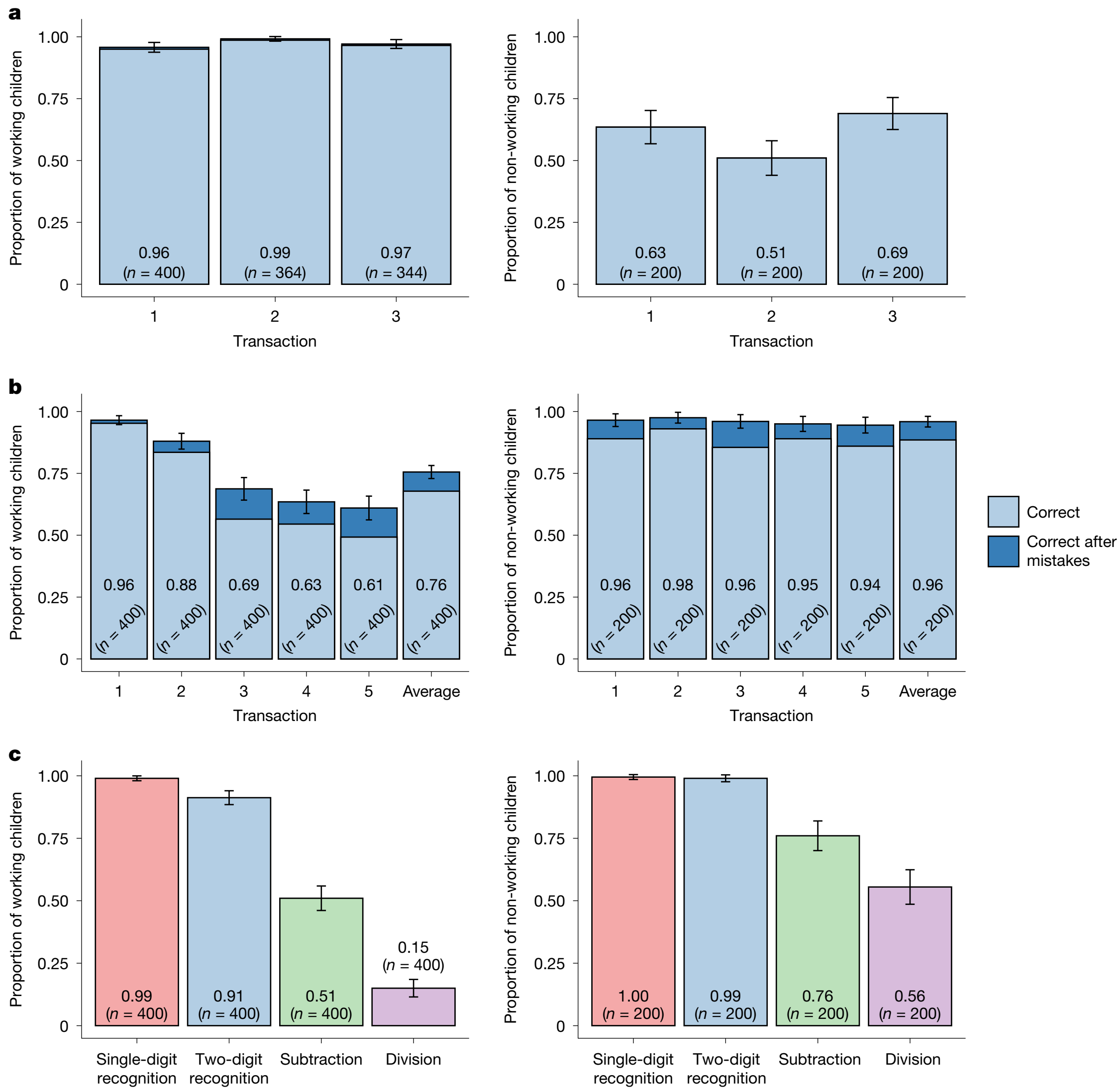


Fig. 2 | Performance of working and non-working children in Delhi (study 2). **a**, The proportion of children who correctly answered the total amount due in three transactions involving two goods sold in unusual quantities. **b**, The proportion of children who correctly answered the total amount due in five hypothetical transactions. **c**, The proportion of children who were credited with labelling single-digit numbers, labelling two-digit numbers, subtracting and dividing on the ASER test. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.).

them with the same hypothetical market maths problems given to the working children in Delhi.

The children in Delhi who were surveyed in schools and in markets differed in many respects, including their schooling history and their experience with markets (Extended Data Table 5). Twenty-two per cent of the non-working school children in study 2 had worked at a market, but only 1% of them had handled transactions. However, non-working children had attained, on average, a full grade level in school above that of the working children. Thus, differences in the arithmetic skills attained by these two groups captured not only the causal effect of attending school or working regularly but also the potential effect of selection for working in markets at a young age and the effect of an additional grade of school instruction, among other factors.

Non-working children solved simple abstract arithmetic problems with high accuracy when they were given unlimited time and pen and paper. Fifty-six per cent of children correctly answered the division portion of the written exercises presented abstractly, as in school (ASER test) (Fig. 2c and Supplementary Table 7). These proficiency rates were below grade-based expectations for grades 8 and 9, but they were above the performance of working children, of whom 15% correctly answered the division portion of the tests ($\beta = 0.41$, s.e.m. = 0.04, 95% CI = 0.34–0.47, $P < 0.001$; Supplementary Table 7). Most of the non-working school children also correctly solved hypothetical transaction maths problems, regardless of whether the problem involved one or two goods. On average, they performed with 96% accuracy with one chance at self-correction (Fig. 2b and Supplementary Table 11).

Thus, non-working school children seemed to master basic arithmetic skills well. However, they performed poorly in the more concrete problems asked in the pretend market. About 60% of them could correctly calculate the amount due in market transactions (63%, 51% and 69% of the non-working school children correctly performed the first, second and third transactions, respectively, on the first or second try; Fig. 2a and Supplementary Table 5). Furthermore, a comparison with working children is somewhat unfair, as the non-working children were able to use pen and paper and had no time limit.

Indeed, non-working school children relied heavily on pen and paper whenever given a chance: 96% of them used pen and paper on the ASER test and 74% of them used pen and paper for the hypothetical market transactions (Fig. 3a and Supplementary Table 12). An examination of the written work of the school children on all these untimed exercises suggested that most school children had not fully mastered the algorithms of elementary arithmetic (Fig. 3c and Extended Data Fig. 1).

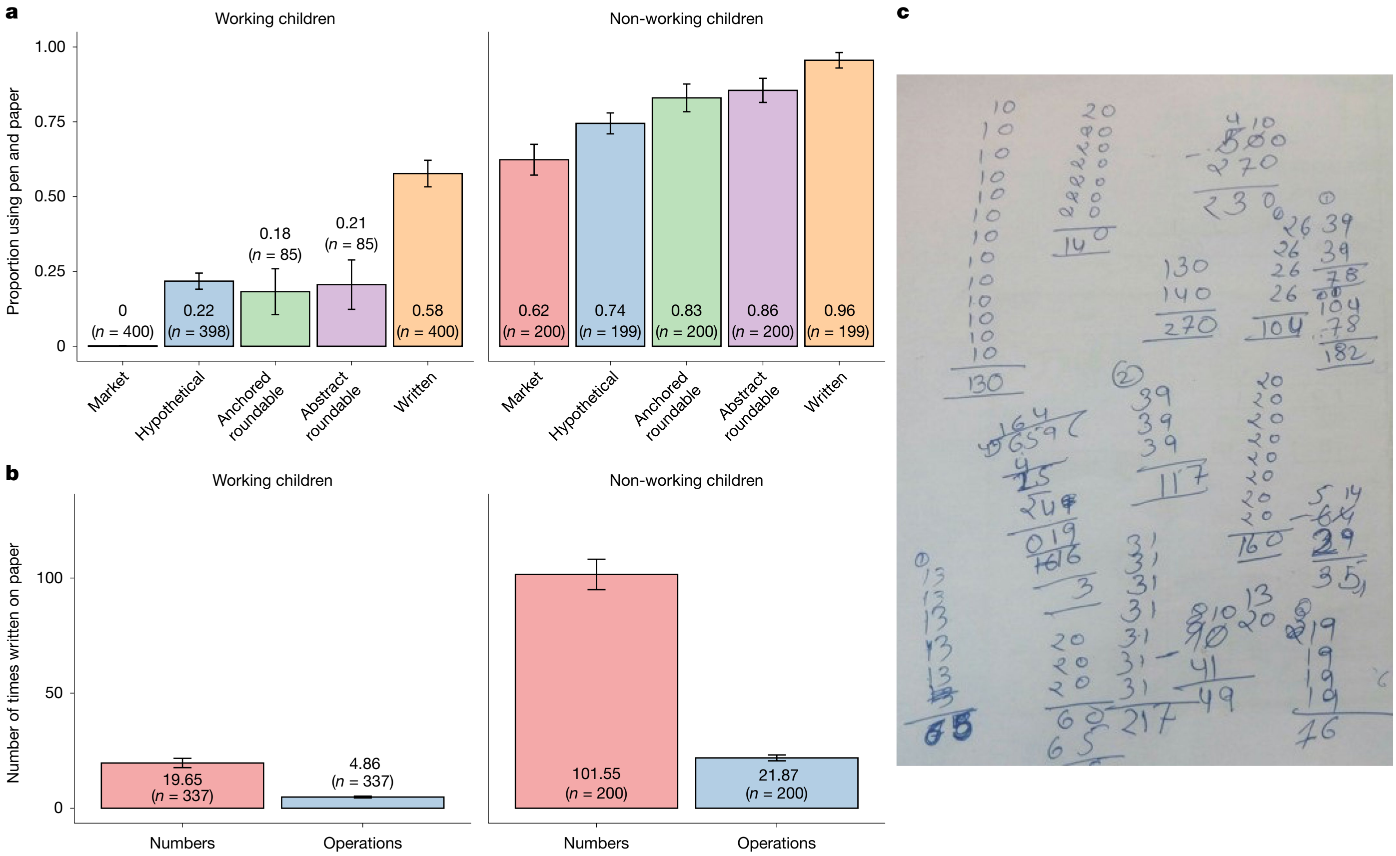


Fig. 3 | Calculation methods of working and non-working children in Delhi (study 2). a, The proportion of children who used pen and paper, classified by respondent type and type of exercise. **b**, The number of times that working and non-working children wrote numbers and operations in the paper given to them for non-oral exercises. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). **c**, Example of calculations on paper written by a non-working child.

On average, a non-working child wrote 102 numbers and 22 operations on their sheets of paper (Fig. 3b and Supplementary Table 13). Nearly half (46%) of these children repeated the same operation and 6% wrote the same operation in two different ways. Close to one-fifth (17%) of these children made extensive use of tally marks and serial numbers, converting addition problems to a process of counting by ones and/or converting multiplication problems to a process of counting by twos or fours (a strategy taught in grade 1). Even with the use of these strategies, non-working school children often had to rewrite or reframe the same problem several times before reaching a solution.

When faced with maths transactions in the simulated market, non-working school children did not change strategy (Fig. 3a). Around two thirds of non-working children used pen and paper (whereas none of the working children did), and across the three transactions, 23% used repeated additions instead of multiplying, which was a significant difference from the working children for matched transactions ($\beta = 0.18$, s.e.m. = 0.02, 95% CI = 0.14–0.23, $P < 0.001$; Supplementary Tables 6 and 14). For the pretend market maths problems, it seemed that these strategies were inadequate, perhaps because they involved several separate steps and the strategies adopted for each step took a long time to perform and combining different steps is yet another skill. However, in this study, we cannot rule out that the school children performed poorly because of the unfamiliar setting of the pretend market (for example, because they were excited by the game aspect).

Limited skill transfer is not explained by confounds

In this study (study 3), we asked why working children and school children performed so differently depending on the context in which a maths problem was presented. We conducted 835 surveys with working children in Delhi over two waves of data collection: March–April 2023 (wave 1) and February 2024 (wave 2). We also surveyed 271 children with no market-selling experience from nearby schools in November–December 2022 (along with wave 1 of the market survey).

Working children again performed well in their jobs. Overall, 85% calculated the correct amount due and change in the market transactions (Supplementary Table 15), whereas non-working children struggled in a pretend market situation. The pretend market was set up in a similar manner as for study 2 but with a time limit, no use of pen and paper and problems matched in difficulty to what children who work in markets do. This set up enabled us to test children's performance in a setting that most closely paralleled that in a market. Only 10% of non-working children could provide the correct change (Supplementary Table 16). Finally, working and non-working children were provided with the same oral abstract subtraction and division problems (with pen and paper allowed). Overall, 59% of non-working children correctly solved these problems compared with 45% of working children ($\beta = -0.14$, s.e.m. = 0.03, 95% CI = −0.20 to −0.08, $P < 0.001$; Fig. 4, left). Thus, the results of study 3 replicated the principal findings of studies 1 and 2 and highlight the contrasts between working and non-working children both in real-life and school-like situations.

To examine the reasons for these differences in performance, new data collection and several experiments were embedded in the data collection effort for both groups of children.

Robustness to alternative explanations

Our leading explanation for these results is that working children do less well on abstract maths questions because they resort to poorly mastered algorithms taught in school. Conversely, school children do poorly in applied problems because they do not know any strategies

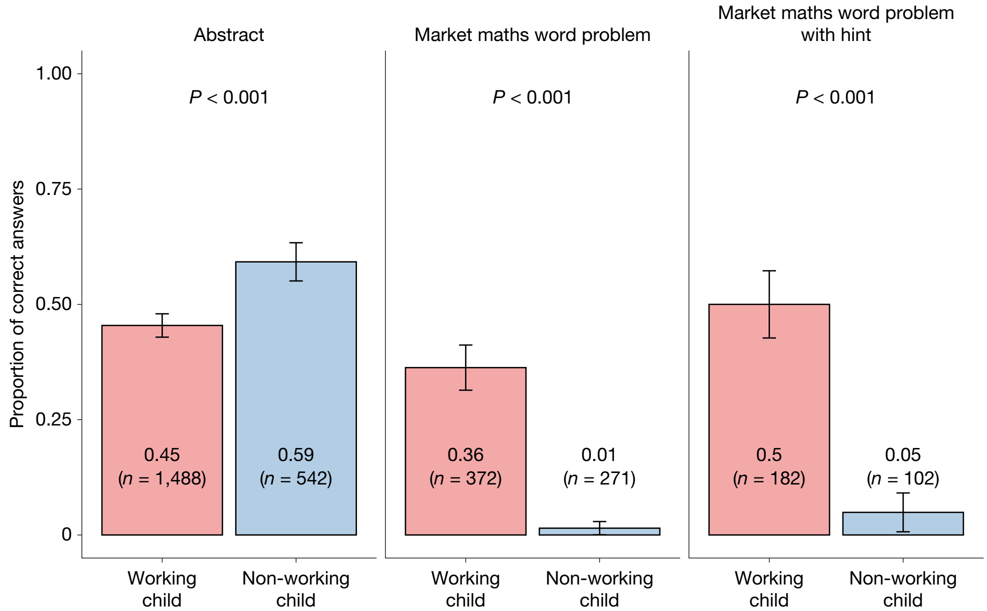


Fig. 4 | Comparison of working and non-working children in oral abstract and market maths word problems (study 3). Left, comparison of the performance of working and non-working children on the same abstract subtraction and division problems. Middle, comparison of the performance of working and non-working children on the market maths word problem after a single attempt. Right, comparison of the performance of working and non-working children on the market maths word problem after receiving a hint to break down the problem and to make use of rounding strategies (process + rounding hint). The oral abstract problems and market maths word problems were not matched in difficulty with one another. A child was coded as correct if they initially correctly answered the question or after receiving the hint. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). *P* values are calculated using two-tailed *t*-tests without adjustments for multiple comparisons.

other than those taught in school and they have not mastered those strategies sufficiently well to solve these more involved problems. It is possible, however, that the children's performance was impaired by stress, stereotype threat or weak incentives to perform accurately, so we tested the effects of these factors.

Familiarity

We asked whether the differences in performance reflect the impact of the presentation by itself or by anxiety induced by the lack of familiarity. It is conceivable that working children were anxious when confronted with a standard abstract problem that they would otherwise be able to solve. Alternatively, school children may have been anxious, or too excited to focus, when performing in the pretend market situation because they had never worked in a market.

To address this issue, we created a question that would feel familiar to both working and non-working children. We devised a concrete maths word problem presented in a format that is familiar to school children but mimicked the activity that working children perform. It is a story of a boy, Vishal, who goes to the market with 200 rupees and buys different quantities of two vegetables. The question is how much money Vishal is left with. Rounding could be used to simplify problems of the cost of both goods. We did not allow pen and paper for this question. This task has a familiar context for school children (who routinely solve word problems in school) and is a familiar task for the working children (who routinely perform this task in their work).

Working children outperformed school children on this market maths word problem. At the first attempt, 36% of working children correctly answered the question compared with 1% of non-working children ($\beta = 0.35$, s.e.m. = 0.03, 95% CI = 0.30–0.40, $P < 0.001$; Fig. 4, centre). When the first answer was incorrect, at the second attempt, a random subset of children were given a hint to break down the problem and to apply a rounding strategy. In such cases, 50% of the working children correctly answered the problem by their second attempt, compared with 5% of the non-working children ($\beta = 0.45$, s.e.m. = 0.05, 95% CI = 0.35–0.55, $P < 0.001$; Fig. 4, right). The working children calculated more fluently than the non-working children despite the attempt to render the problem equally familiar to both groups. We also did not find evidence that working children who also attended school performed better than those out of school on the market maths word problem ($\beta = -0.02$, s.e.m. = 0.05, 95% CI = −0.12 to 0.08, $P = 0.730$; Extended Data Table 6), although the working children who were still in school performed almost twice as well on abstract maths problems ($\beta = 0.29$, s.e.m. = 0.03, 95% CI = 0.22–0.36, $P < 0.001$; Extended Data Table 6). These results provide further evidence of the minimal skill transfer from school to real-world situations.

Stress and stereotype threat

When working children were taken aside to answer questions in a school format, they may have felt threatened by it. It is possible that the situation induced stress or self-stereotyping (for example, 'I am not good at school'). In study 3a, wave 1, we included a randomized experiment with four conditions to test for the impact of stress and stereotype threat on the performance of working children. In the first treatment ('market maths problem first'), to put children at ease, we presented the hypothetical market maths problems before the oral abstract subtraction problems. The second treatment ('encouragement') aimed to vary the level of self-confidence of the children by having the surveyor use encouraging language during the survey versus without. We cross-randomized these treatments to produce four conditions.

At the end of the survey, we measured children's stress levels using a series of Likert-scale questions. Both confidence-related treatments reduced the stress levels of working children but did not improve performance. Relative to the control group who received neither treatment, the market maths problem first group, the encouragement group and the group who received both treatments reported 0.17, 0.32 and 0.23 standard deviations less stress, respectively. However, we did not observe significant improvements in performance of solving abstract subtraction questions across any of the treatments: market

Table 1 | Impact of confidence treatments on stress and maths performance of working children in Delhi

			Market maths problem first only						Encouragement only						Market maths problem first with encouragement					
Outcome	***n***	**Control**	**Treatment**	**Difference**	**95% CI**	**s.e.m.**	***t***	***P***	**Treatment**	**Difference**	**95% CI**	**s.e.m.SE**	***t***	***P***	**Treatment**	**Difference**	**95% CI**	**s.e.m.**	***t***	***P***
Stress index	372	0.19	−0.14	−0.32	−0.62 to 0.02	0.15	−2.09	0.037	0.02	−0.17	−0.47 to 0.13	0.15	−1.12	0.264	−0.05	−0.23	−0.54 to 0.07	0.16	−1.51	0.133
Abstract subtraction index	372	0.11	−0.09	−0.20	−0.49 to 0.08	0.15	−1.40	0.163	0.05	−0.06	−0.34 to 0.22	0.14	−0.40	0.689	−0.05	−0.16	−0.46 to 0.13	0.15	−1.10	0.272
Hypothetical market index	372	−0.07	−0.09	−0.01	−0.32 to 0.29	0.15	−0.10	0.922	0.13	0.21	−0.10 to 0.51	0.16	1.33	0.183	0.04	0.11	−0.18 to 0.40	0.15	0.73	0.465

n, sample size. Control, mean for the control group who received abstract maths problems first (as opposed to hypothetical market maths problems) and no verbal encouragement. Treatment, mean for the respective treatment groups. Difference, the difference between the treatment group and control group (treatment effect). *t* and *P*, *t*-statistics and *P* values, respectively, of the treatment effect using a two-tailed test without adjustments for multiple comparisons. Stress index is a standardized index created from five Likert-scale questions evaluating children's stress levels. In particular, we asked the children the following questions: (1) during this survey, how did you feel?; (2) how confident are you about calculating in your head?; (3) how confident are you about maths at school?; (4) how confident are you in your answer for the last problem? We also asked surveyors: how stressed do you think the child was during the interview? (question 5). Abstract subtraction index is a standardized index of children's performance on two abstract subtraction problems before and after receiving hints. Hypothetical market index is a standardized index calculated from children's performance on hypothetical market maths problems after two attempts.

maths problem first group (β = −0.20, s.e.m. = 0.15, 95% CI = −0.49 to 0.08, P = 0.163); encouragement group (β = −0.06, s.e.m. = 0.14, 95% CI = −0.34 to 0.22, P = 0.689); or market maths problem first with encouragement group (β = −0.16, s.e.m. = 0.15, 95% CI = −0.46 to 0.13, P = 0.272). Therefore we can reject small effects (Table 1).

Incentives

School children may face weaker incentives to correctly answer problems. So, we tested whether providing school children with material incentives improved their performance, by giving a random sample of children tokens to exchange for gifts. We did not find significant differences in performance across groups (β = 0.10, s.e.m. = 0.12, 95% CI = −0.14 to 0.34, P = 0.406; Extended Data Table 7).

Meaningfulness

We randomized children to receive different versions of the survey so that we could test them on oral abstract and oral anchored problems that involved the same numbers. Working children performed significantly better on anchored problems, correctly answering 72% of anchored problems compared with 64% of abstract problems (β = 0.08, s.e.m. = 0.02, 95% CI = 0.03–0.12, P = 0.001; Fig. 5a and Supplementary Tables 17 and 18). By contrast, non-working children performed significantly worse on oral anchored addition and multiplication problems, correctly answering 73% of anchored problems and 78% of abstract problems (β = −0.06, s.e.m. = 0.03, 95% CI = −0.11 to −0.01, P = 0.028; Fig. 5a and Supplementary Table 18). Thus, meaningfulness had differing effects on the performance of the two groups of children, a result consistent with the different contexts in which they learnt arithmetic skills.

Understanding the different processes working children use

Given the large gap between applied and theoretical skills, with no transfer in either direction, we collected additional data on how working children reason and encouraged them to use more efficient algorithms.

Compositional structure

We asked working children how they solved the hypothetical market maths problems involving goods from their store. Many children simplified calculations by using rounding and decomposition techniques that leverage the base-ten structure of the number system (44% when problems involved standard market prices and 36% when they involved new prices) (Supplementary Table 19). They did so by converting complex operations (for example, 11 × 43) into a series of simpler ones (for example, solving 11 × 43 by converting the problem to (10 × 43) + 43). A large proportion (26% at market prices and 16% at new prices; Supplementary Table 19) also used approximations to simplify calculations by rounding subtotals up or down by 1 or 5 (for example, approximating 0.7 kg of cauliflower at 20 rupees per kilogram to be 15 rupees). These methods may explain how working children solve complex problems so quickly and accurately without the need for pen and paper.

We then tested whether children leveraged these strategies on school-like abstract and anchored problems. We compared their performance on problems with an operand ending in 1 or 9 (roundable problems) with their performance on problems presenting numbers of similar sizes lacking those endings (non-roundable problems such as 12 × 42). Working children performed slightly better on roundable problems (67% correct) than on non-roundable problems, although the difference was not significant at the 5% level (64% correct, β = 0.03, s.e.m. = 0.02, 95% CI = 0.00–0.07, P = 0.083; Fig. 5b and Supplementary Table 20). For non-working children, we did not see evidence of a difference: 75% answered roundable arithmetic questions correctly compared with 74% of non-roundable ones (β = 0.00, s.e.m. = 0.02, 95% CI = −0.04 to 0.05, P = 0.868; Fig. 5b and Supplementary Table 20). It was notable that non-working school children did not use this structure, particularly as it forms the foundations of the arithmetic algorithms taught in school. Moreover, working children, who apparently discovered this structure over the course of their work in markets, had not mastered the school-taught algorithms that were based on it.

Hints

In further experiments, we asked whether the differences between working and non-working school children are easily overridden or difficult to overturn. Because working children performed better on problems that were anchored to a familiar and concrete context, we asked whether they could improve their performance in abstract arithmetic problems by reframing them in such a context. Because school children did not seem to effectively use the base-ten system to simplify problems, and market children did so only to a limited extent, we asked whether children in both groups could be induced to use rounding to improve their calculations. We addressed these questions by including an experiment in which children were provided various hints.

For working children, we tested rounding and anchoring hints in two separate scenarios. Neither anchoring nor rounding hints had a significant impact on how accurately the children answered these problems, with treatment effects of −2 percentage points (β = −0.02, s.e.m. = 0.04, 95% CI = −0.09 to 0.05, P = 0.583; Extended Data Fig. 2a and Supplementary Table 21) and 1 percentage point, respectively (β = 0.01, s.e.m. = 0.03, 95% CI = −0.06 to 0.07, P = 0.866; Extended Data Fig. 2b and Supplementary Table 22). For non-working children, the hints were given in the context of the market maths word problem. They received hints to break down the problem and to use rounding strategies. Again, we did not find evidence that hints improved performance relative to those who only received a second chance (β = 0.00, s.e.m. = 0.03, 95% CI = −0.06 to 0.06, P = 0.881; Extended Data Table 8 and Extended Data Fig. 2c). We conclude that by the time

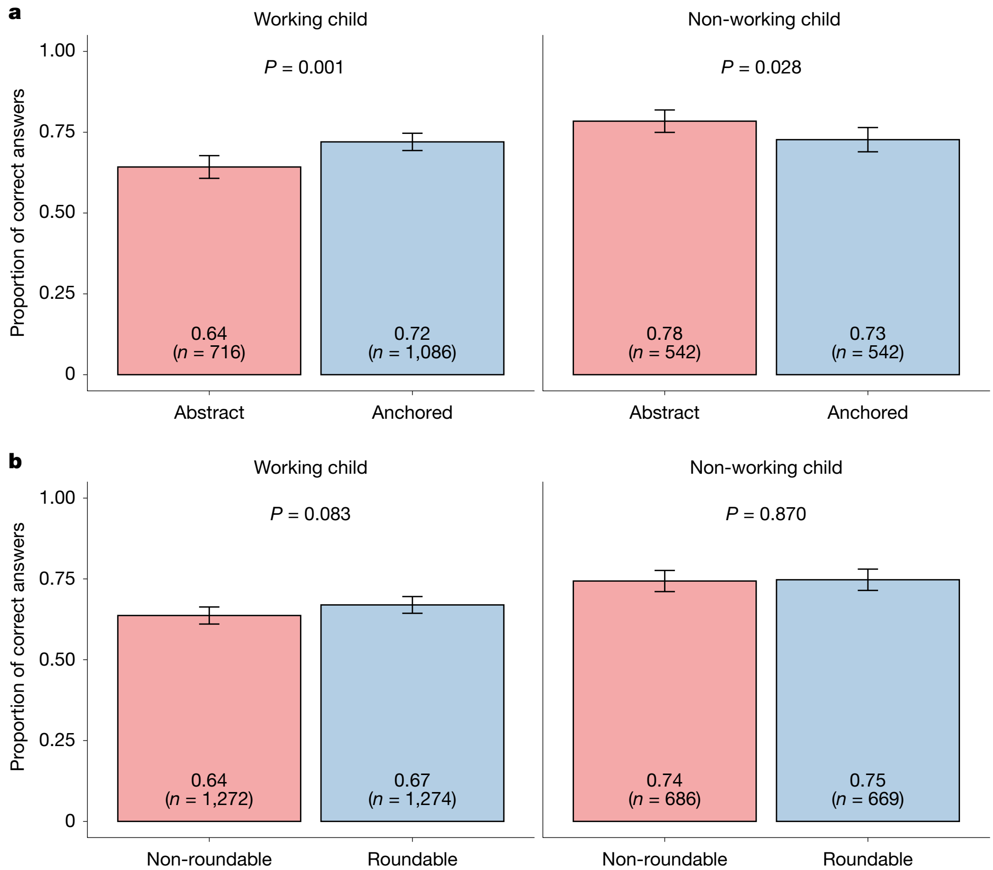


Fig. 5 | Comparison of working and non-working children on abstract versus anchored and roundable versus non-roundable problems (study 3). **a**, The proportion of children who correctly answered abstract versus anchored problems. We tested children on oral problems involving addition and multiplication. **b**, The proportion of children who correctly answered roundable versus non-roundable problems. We tested children on oral problems involving abstract addition, abstract subtraction, abstract multiplication, anchored addition and anchored multiplication. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). *P* values were calculated using two-tailed *t*-tests without adjustments for multiple comparisons.

they reach adolescence, the cognitive differences between working and non-working children are not easily overturned.

Schools must bridge the gap between formal and intuitive maths

These studies provide evidence from three large groups of children from India that the majority of children working in markets are skilled in the arithmetic calculations used in their daily work. Moreover, they flexibly perform market transactions such that they rapidly adjust to correctly deal with transactions involving unusual quantities of the goods they sell. When tasked with a hypothetical transaction involving a change in price or items, most children performed correctly on the first try, and many of those who did not were correct on the second try. Even when presented with a complex market-inspired word problem involving four different operations, 35% correctly answered on the first try (and 50% succeeded after a chance at self-correction and a hint), which implied high levels of success at performing, sequencing and remembering each of the four arithmetic operations. By contrast, only 1% ($\beta = 0.35$, s.e.m. = 0.03, 95% CI = 0.30–0.40, $P < 0.001$; Fig. 4, centre) of the non-working children could solve the word problem in one try (5% with hints; $\beta = 0.45$, s.e.m. = 0.05, 95% CI = 0.35–0.55, $P < 0.001$; Fig. 4, right). We rule out confounding factors such as the possibility that the working children were chosen for their maths skills, that they received help from others, used calculation aids or had stronger incentives. It seems plausible that experience in markets, where goods change with seasonal factors, where prices change regularly across or within days and where there is always time pressure, leads children to discover flexible arithmetic strategies.

By contrast, we found little evidence that these real-world competencies helped children to perform school-based maths. Working children, most of whom were still in school (and the rest had attended school in the past) were not able to solve elementary written or oral abstract problems of subtraction and division as presented in school. This result is notable, particularly given that the algorithms taught in school are also based on the same base-ten compositional structure that working children use in solving the market maths problems that can be simplified by rounding to the nearest ten.

Conversely, non-working school children performed more accurately than working children in solving simple abstract problems taught in schools but poorly on concrete problems involving several simple operations, such as the ones routinely solved by working children. Moreover, their accuracy came at an extreme cost in speed. That is, many school children solved simple multiplication problems, such as 24 × 8, by rewriting them as longer addition problems (here, a problem with eight two-digit addends) or still longer counting problems (here, by tallying and counting the numbers of twos and ones in the written addition problem). Thus, neither group of children had developed the maths skills required for further study in maths, science and many other disciplines.

These findings point to a broader failure of the pedagogical practices in India to make usable connections between intuitive and formal understanding of maths ideas. The pedagogy in place does not teach school students strategies to do maths in real-world settings. It also does not take advantage of the fact that working children have developed such strategies on their own to help them bridge their specific skill sets, well-tailored to the markets, to the skills needed to succeed in a school-based setting. This failure echoes the finding from a randomized control trial (RCT) of low-income preschool children in Delhi[20], whereby primary schools were not able to build on intuitive maths skills that children had acquired by playing math games in preschool. Although children's intuitive sense of maths durably improved following an intervention in preschool, their performance in formal maths did not when they encountered the grade 1 maths curriculum over the subsequent year[20].

These findings call for a maths pedagogy that explicitly addresses these translational challenges through curricula that connect abstract maths symbols and concepts to intuitively meaningful contexts and problems. Consistent with that call, a RCT found that introducing financial education to public high school students in Brazil was effective not only in improving financial literacy but also in reducing school failure rates, with qualitative interviews suggesting that students felt more engaged with maths presented in familiar contexts[21].

The low performance of working children and school children in primary school maths problems also calls for changes in how maths is introduced to children, in particular to better synergize intuitive knowledge and training in symbolic maths. There is encouraging evidence that it might be possible: two RCTs of students in preschool, kindergarten and grade 1 in Delhi showed that a curriculum that pairs intuitive and abstract maths materials, and leads children to exercise both in alternation in group games, had durable impacts on both their school maths and their intuitive maths skills (Dean, J. T. et al., manuscript in preparation).

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08502-w.

1. Angrist, N., Djankov, S., Goldberg, P. K. & Patrinos, H. A. Measuring human capital using global learning data. *Nature* **592**, 403–408 (2021).
2. Carraher, T. N., Carraher, D. W. & Schliemann, A. D. Mathematics in the streets and in schools. *Br. J. Dev. Psych.* **3**, 21–29 (1985).
3. Carraher, T. N., Carraher, D. W. & Schliemann, A. D. Written and oral mathematics. *J. Res. Math. Edu.* **18**, 83–97 (1987).
4. ASER. *Annual Status of Education Report 2023. Beyond Basics: A Survey of Youth in Rural India* (ASER Centre, 2024).
5. World Bank. *World Development Report 2018: Learning to Realize Education's Promise* (World Bank, 2017).
6. Reed, H. & Lave, J. Arithmetic as a tool for investigating relations between culture and cognition. *Am. Ethnol.* **6**, 568–582 (1979).
7. Pica, P., Lemer, C., Izard, V. & Dehaene, S. Exact and approximate arithmetic in an Amazonian Indigene group. *Science* **306**, 499–503 (2004).
8. Gordon, P. Numerical cognition without words: evidence from Amazonia. *Science* **306**, 496–499 (2004).
9. Fuson, K. C. *Children's Counting and Concepts of Number* (Springer, 1988).
10. Groen, G. J. & Resnick, L. B. Can preschool children invent addition algorithms? *J. Edu. Psychol.* **69**, 645–652 (1977).
11. Carpenter, P., Moser, J. M. & Romberg, T. A. (eds) *Addition and Subtraction: a Cognitive Perspective* (Erlbaum, 1982).
12. Resnick, L. B. in *Addition and Subtraction: a Cognitive Perspective* (eds Carpenter, T. P. et al.) Ch. 10 (Erlbaum, 1982).
13. Feigenson, L. & Carey, S. Tracking individuals via object-files: evidence from infants' manual search. *Dev. Sci.* **6**, 568–584 (2003).
14. Halberda, J., Mazzocco, M. M. M. & Feigenson, L. Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* **455**, 665–668 (2008).
15. Brown, J. S., Collins, A. & Duguid, P. Situated cognition and the culture of learning. *Educ. Res.* **18**, 32–42 (1989).
16. Curtis, P. C., Labov, J. B., Bertenthal, M. W. & Gollub, J. P. (eds) *Learning and Understanding: Improving Advanced Study of Mathematics and Science in US High Schools* (National Academies Press, 2002).
17. ASER. *Annual Status of Education Report 2016* (ASER Centre, 2016).
18. Ginsburg, H. P. in *Addition and Subtraction: a Cognitive Perspective* (eds Carpenter, T. P. et al.) Ch. 14 (Erlbaum, 1982).
19. ASER. *Urban Ward Survey 2014: Delhi* (ASER Centre, 2014).
20. Dillon, M. R., Kannan, H., Dean, J. T., Spelke, E. S. & Duflo, E. Cognitive science in the field: a preschool intervention durably enhances intuitive but not formal mathematics. *Science* **357**, 47–55 (2017).
21. Bruhn, M. et al. The impact of high school financial education: evidence from a large-scale evaluation in Brazil. *Am. Econ. J. Appl. Econ.* **8**, 256–295 (2016).

Methods

All materials and methods were reviewed and approved by the Institutional Review Boards at the Massachusetts Institute of Technology (MIT) (IRB1504007082, IRB1612798217 and IRB2208000727) and the Institute for Financial Management and Research (IRB00007107). We obtained informed consent from all study participants. The studies involved randomization, and we used power calculations to determine our target sample sizes. We include a more detailed summary of our methods in the Supplementary Methods.

Statistical analysis and code

In all analyses, we calculated group means and compared them using two-tailed *t*-tests. In Extended Data Table 6 and Supplementary Tables 21 and 22, we compared groups using linear regressions. We present these results with and without controls. To calculate standardized indices (for example, when measuring stress), we first standardized the individual outcomes that make up a given index, averaged them and then standardized the final variable to have a mean of 0 and a standard deviation of 1. We conducted all of our analyses in Stata and plotted figures in R.

Study 1

Sample. There is no list of all informal (wholesale and retail) markets that employ children in India because child labour is illegal. Therefore we drew a convenience sample of markets in all studies. In study 1, we identified potential markets and then visited them to verify that they employed children. This process identified 92 markets. We drew a convenience sample of working children at each market. An enumerator walked around each market and noted where children were located. We only approached children who appeared to be 16 years of age or younger, were selling goods and sold more than one good. We did, however, stop the survey if we discovered that the child was 18 years of age or older. This led us to include some children who were 17 years of age. In study 1, 201 out of 285 working children whom we approached agreed to participate. Given that the vast majority of children consented to participate, selection on who consents is unlikely to be driving our results, although we cannot rule this out entirely. Of these children, 141 were enrolled in school (average grade attained at the time of testing, 6.24) and 60 were not (average grade attained, 3.67) (Supplementary Table 3).

Survey structure. For study 1, our survey consisted of the following sections: (1) market transactions; (2) consent; (3) ASER written assessment; (4) oral abstract and anchored assessment; (5) hypothetical transactions; and (6) demographics.

Market transactions. In the market transactions section, each working child was approached by two enumerators dressed as regular buyers who purchased two goods (for example, 300 g of beans and 200 g of peas). When goods were sold by the unit (for example, bananas), enumerators bought them by the unit. When they were sold by the kilogram (for example, tomatoes), enumerators bought them by the kilogram or gram. In both cases, enumerators avoided purchasing common quantities (for example, 1 kg or 500 g or 1–2 units) to minimize the probability that children had memorized the prices of the quantities requested.

When a child told the enumerators the amount due, they were asked how they had arrived at that total (for example, by an enumerator who said, 'I expected it to be less'), which prompted the child to explain their calculation. If the calculation was correct, the enumerators paid and left. If it was incorrect, the enumerators asked the child to verify their calculation (for example, saying 'Are you sure? How did you get that? Do it again and see'). Regardless of whether the child's calculation of the amount due was correct, the enumerators offered more money than was due, took the change and left. The use of calculators, paper aids and help from adults was permitted but the enumerator kept a record of each.

In studies 1 and 2a, two other pairs of undercover enumerators repeated the same process, resulting in three transactions per child. One of the three pairs of enumerators revealed their identity and invited the child to participate in the study.

ASER. The ASER is a survey of learning outcomes for school-aged children conducted annually by Pratham, a non-governmental organization in India. ASER is representative of rural India, although in some years, it was also conducted in select urban areas.

Study 2a

Sample. We followed a similar process for identifying children and markets as in study 1. In this study, we conducted the main survey after the first market transaction. We later returned with undercover enumerators to complete the second and third transactions. Of the 442 working children approached, 400 agreed to participate and 344 completed all three transactions (Extended Data Table 2)

Survey structure. The survey consisted of the following sections: (1) market transactions (as in study 1); (2) consent; (3) ASER written assessment (as in study 1); (4) oral abstract and anchored assessment (as in study 1); (5) hypothetical transactions; (6) roundable problems for a subsample of 85 children; and (6) demographics.

Rounding prices in the market. It is common practice for shopkeepers in markets in India to round prices up or down by one or to the nearest multiple of five. In studies 2 and 3, for all market transactions (for working children), pretend market transactions (for school children) and hypothetical transactions (for working and non-working children), we considered the price quoted for an individual good to be correct if it was rounded up or down by one from the exact answer, rounded up or down by five from the exact answer or the exact answer itself. For example, if the exact answer was 35.5 rupees, we would treat 30, 35, 35.5, 36 and 40 as correct answers. We did this because such rounding is common practice and does not reflect the inability of the child to solve problems.

For the combined price of goods 1 and goods 2, we considered any answer that adds the correct (rounded or unrounded) prices of individual goods to be correct. For example, if carrots cost 18 rupees and potatoes cost 32 rupees, we would consider a total price of 52 rupees to be a correct answer; 18 rupees can be correctly rounded to 20 rupees, and 20 rupees plus 32 rupees gives a total of 52 rupees.

For the change provided, we calculated the correct answer by subtracting the true cost of each good from the cash provided (not by subtracting the amount quoted by the child from the cash provided). For example, if we purchased 600 g of apples at 30 rupees per kg and 200 g of potatoes at 20 rupees per kg and gave the child 200 rupees, then we would calculate the exact answer as $200 - (0.6)(30) - (0.2)(20) = 200 - 18 - 4 = 178$, irrespective of what the child quoted for the total amount due. We also allowed the same rounding as described for the combined price of goods 1 and goods 2. For example, we would also treat $200 - 20 - 5 = 175$ to be correct.

Hypothetical transactions. In study 2a, there were five hypothetical transactions that varied one aspect of the transaction at a time. For example, for a child who sold carrots by the kilogram that day, the first would be a hypothetical problem focused on selling carrots at the familiar price by kilogram. The second problem would be selling carrots at a new price by kilogram, the third, selling carrots at the same new price by kilogram and a second unfamiliar good sold at a new price by kilogram. The fourth problem would be selling a new good, for example bananas, at a new price by units, and finally, selling bananas at the same new price by units and a second unfamiliar good at a new price by units (see Supplementary Methods, study 2a, for more details and examples of the problems). We provided children with pen and paper.

Incentive randomization. Each child was randomly assigned to two versions of the exercises: a version that offered them 10 rupees for every exercise that they answered correctly (in addition to the 200 rupees for participating) and a version that did not offer any incentives.

Roundable problems. In study 2, a subset of 85 working children were given four exercises explicitly designed to encourage the use of decomposition and repeated grouping. To give children maximum latitude to reveal their sensitivity to the base-ten structure, paper and pencil aids were made available to them. In all other respects, the problems mirrored those of the abstract and anchored arithmetic problems from the preceding test. They were not provided any hint that this was a possible strategy to follow. We present the results in Supplementary Table 23.

Solution methods. We collected all the sheets of paper that working children used during the written exercises, hypothetical transactions and roundable problems. We developed a coding system to characterize the calculation strategies used by these children (for example, the number of numbers that children wrote on the page, whether they used tallies or serial numbers, among others) and asked data-entry operators to use these codes to enter data on every sheet of paper. Every sheet was coded twice by two different data-entry operators to identify and correct any inconsistencies in data entry.

Study 2b

Sample. In study 2b, in July–August of 2017, we surveyed 200 children attending 20 government schools located in a 2-km radius of the markets we identified. We selected ten children at random from grade 8 and ten children from grade 9. All school children agreed to participate.

Survey structure. Study 2b consisted of the following sections: (1) consent; (2) pretend market transactions; (3) ASER written assessment (as in study 2a); (4) oral abstract and anchored assessment (as in study 2a); (5) hypothetical transactions (as in study 2a) and roundable problems (as in study 2a); and (6) demographics. We recorded solution methods as in study 2a.

Pretend market transactions. We set up a pretend market in classrooms with plastic goods and money. We told children to imagine they were selling items in a shop and that the surveyor was a customer. We randomized children to sell by kilograms or units. We asked the child to go to the desk at the front of the classroom and select any two items that they wanted to sell. We then checked the goods that they selected had different prices and explained to the child that the prices written were sold per kilogram or by unit. Children were then given plastic money of 300 rupees in small change and told that they could use pen and paper during the transaction if needed. We then followed the same process we used with working children, performing three separate transactions.

Study 3a, wave 1

Sample. We followed a similar sampling strategy to study 2a. Of the 489 working children who completed a market transaction with one of our surveyors, 372 agreed to participate and completed the full survey. We present the characteristics of our study 3 sample in Supplementary Table 24.

Survey structure. Study 3a, wave 1 consisted of the following sections: (1) market transactions (as in studies 1 and 2a); (2) consent; (3) hypothetical transactions and calculation explanations; (4) oral abstract subtraction and hints (with the order of (3) and (4) randomized); (5) market maths word problem; (6) market maths word problem with hints; (7) oral abstract division; and (8) demographics and stress measures.

Hypothetical transactions and calculation explanations. We asked children two hypothetical transaction questions. They were allowed to attempt the question twice if they were initially incorrect. After each question, we asked children how they solved the problem. We audio-recorded their response, transcribed it and then categorized the solution method they used.

Market maths word problem. We gave children a problem that was presented in a format that was familiar to school children but mimics the activity that working children do. Rounding could be used to simplify the cost of both goods. We did not allow pen and paper. The problem read: "Vishal went to the market with 200 rupees. He bought 450 grams of peas at 100 rupees a kilogram, and 200 grams of tomatoes at 90 rupees a kilogram. How much money does he have left?".

Confidence treatments. In previous work, we found that children who performed well in arithmetic operations in the course of their job did not do well when solving standard arithmetic calculations when presented in a school-maths formulation. Two possible reasons for this are that children become stressed or lose confidence when they are presented with problems in a setting they are not familiar with. To test this hypothesis, we introduced two cross-randomized treatment variations. The first treatment varied whether oral abstract subtraction problems were presented before or after the hypothetical market maths problems. If abstract problems cause stress and stress impairs performance, then children who completed the abstract problems first will be more stressed while completing the abstract and hypothetical market maths problems and perform less well. The second treatment attempted to vary the level of self-confidence of the children by having the surveyor use more or less encouraging language during the survey. Specifically, we began by telling children we were interested in how children who work in markets 'manage to be so exceptionally good at doing mental calculations' versus 'manage to do mental calculations in the course of their work, but often do poorly in school'.

Stress measures. As a first-stage measure for the confidence treatments, we measured children's stress, confidence and attitudes to maths at the end of the survey.

Study 3a, wave 2

Sample. We returned to 34 out of the 38 markets we visited in study 3a, wave 1. We did not conduct a real market transaction with children in this study. However, we used a similar strategy to recruit children to ensure a comparable sample. As in previous studies, surveyors spent time observing children to determine whether they were handling transactions at their shop (calculating prices and providing change). If they appeared to be doing so, surveyors approached the child and enquired about the price of a good at a specific quantity. For example, asking 'How much do 300 g of apples cost?'. If the child did not answer the question and had someone else in the shop answer on their behalf, we assumed that the child did not handle transactions and did not begin the survey with them. As a final check, we also directly asked children at the beginning of the survey whether they handled transactions and screened them out of the study if they did not. Of the 566 working children who were approached by one of our surveyors, 463 passed these screens, agreed to participate and completed the full survey.

Survey structure. Our survey consisted of the following sections: (1) consent; (2) simple arithmetic and confidence building; (3) hypothetical transaction; (4) warm-up abstract and anchored addition and multiplication arithmetic calculations; (5) oral abstract subtraction with anchoring hints; (6) oral abstract and anchored multiplication with rounding hints; (7) process hints; and (8) demographics.

Warm-up abstract and anchored addition and multiplication arithmetic operations. We asked children a series of oral addition and multiplication problems. Some of the problems were abstractly framed whereas others were anchored to a concrete context. This section served several purposes. First, it put children further at ease before they began the hint sections of the survey. Second, we randomized children into five groups who were given differing numbers of problems (ranging from two problems to six problems). This provided random variation for the position that questions appeared within the survey, which controlled for question order effects in our subsequent hint experiments. Finally, we compared performance on abstract and anchored problems. We randomized children into two versions of the test to exactly balance problem difficulty on abstract and anchored problems. We also provided a mix of roundable and non-roundable problems of similar difficulty.

Hint experiments. We include a detailed summary of the anchoring, rounding and process hint experiment methods in the Supplementary Methods (study 3a), with results in Extended Data Table 8, Extended Data Fig. 2 and Supplementary Tables 21, 22 and 25.

Study 3b

Identification of schools. We followed a similar sampling strategy to study 2b. We surveyed a total of 300 children, half in grade 7 and half in grade 8. All children we approached agreed to participate. Our main analysis focused on the 271 school children with no market-selling experience.

Survey structure. For study 3b, our survey consisted of the following sections: (1) consent; (2) ASER (as in studies 1 and 2; Supplementary Table 26); (3) written abstract, oral abstract and oral anchored assessment; (4) market maths word problem (as in study 3a, wave 1); (5) pretend market problem (as in study 3a, wave 1); and (6) market maths word problem with hints.

Written abstract, oral abstract and oral anchored assessment. We tested how school children's arithmetic performance changed depending on how questions were delivered by asking questions in written abstract, oral abstract and oral anchored forms. For each of these sections, we tested children on one addition, subtraction, multiplication and division question. Other than the key feature, we balanced the complexity of the problems across these sections by randomizing children into one of four versions of the survey. For example, across the versions, children were tested on oral abstract and oral anchored problems with the exact same numbers. We also balanced the difficulty of roundable and non-roundable problems by ensuring they had similar numerical quantities and the same requirements to carry over numbers.

Pretend market transactions. We again set up a pretend market in children's classrooms but made several changes relative to study 2 to test children's performance in a setting that most closely parallels selling in a market. In particular, we designed the problem to be of comparable difficulty to the real-world market transactions, did not allow children to use pen and paper and required children to respond in 4 min or less.

Market maths word problem with hints. We tested whether providing a hint to use rounding strategies improves performance for non-working children, who may underutilize flexible arithmetic algorithms. There were two treatment groups: one received a hint to break down the problem; the other received a hint to both break down the problem and to use rounding strategies. Finally, the control group could re-attempt the problem without a hint.

Incentives. All children received tokens on the basis of whether they correctly answered the questions. We randomized whether children could exchange these tokens for non-financial gifts (incentives treatment) or whether children received a fixed gift (no incentives treatment).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data are publicly available without restriction from the J-PAL Dataverse on the Harvard–MIT data archive (https://doi.org/10.7910/DVN/VPVH0O). Source data are provided with this paper.

Code availability

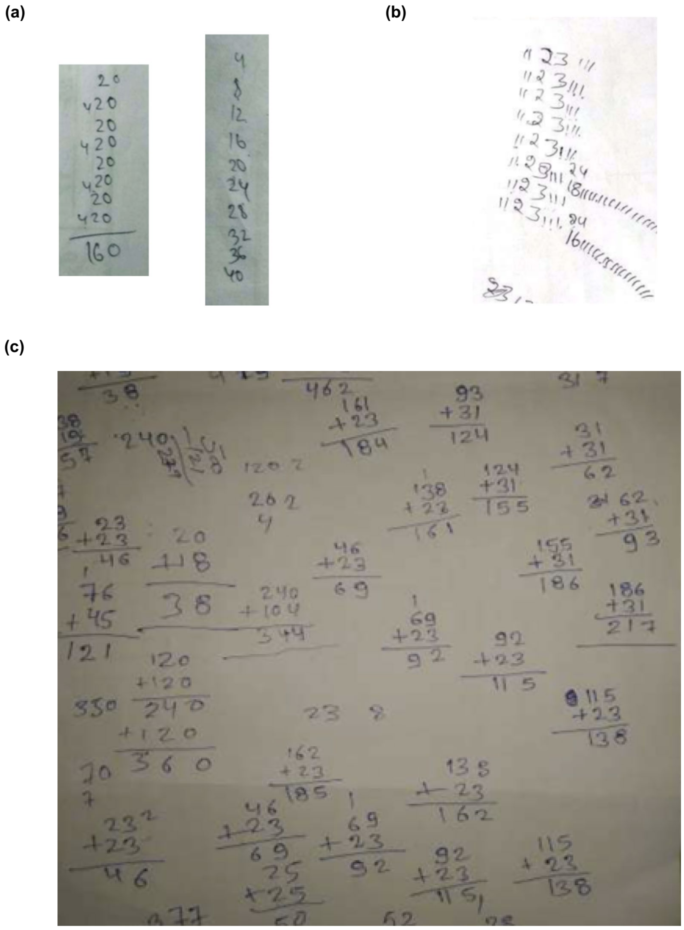
All code is publicly available without restriction from the J-PAL Dataverse on the Harvard–MIT data archive (https://doi.org/10.7910/DVN/VPVH0O).

Acknowledgements This project was funded by the Post-Primary Education Initiative of the Abdul Latif Jameel Poverty Action Lab (J-PAL), the Foundation Blaise Pascal and the Fonds AXA pour la Recherche. The Delhi Directorate of Education authorized the conduct of two of the studies in this article. We thank R. Banerji, P. Bergman, N. Buchbinder, A. de Barros, Y. Dhavala, W. Easterly, A. Eble, H. Hill, H. Kannan, R. Murnane, J. Rockoff, J. Star, M. Urquiola, A. Venkatachalam, A. Wuermli, H. Yoshikawa and seminar participants for comments; and M. Goenka, S. Mallick, R. Menon, R. Pugalia, G. Sharma, A. Srinivasan and P. Tiwari for their research assistance.

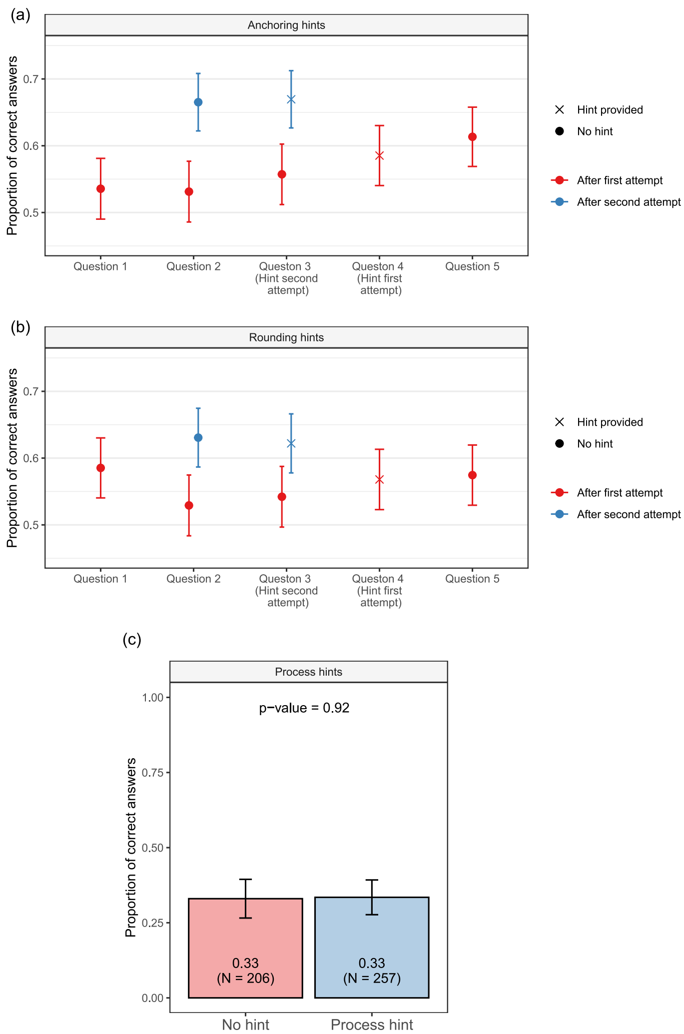
Author contributions The following authors contributed to the research design and management of data collection: A.V.B. and E.D. to studies 1–3; S.B. and R.C. to study 1; A.J.G. and E.S.S. to studies 2 and 3; and K.R. to study 3.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08502-w.
Correspondence and requests for materials should be addressed to Abhijit V. Banerjee, Esther Duflo or Elizabeth S. Spelke.
Peer review information *Nature* thanks Thomas Breda, Stanislas Dehaene, Zach Parolin and Harry Patrinos for their contribution to the peer review of this work.



Extended Data Fig. 1 | An example of non-working children's scrap paper in study 2, when solving the problem: "You want to buy 8 carrots at 23 rupees each; what do you pay?". (a) Multiplication converted to repeated numbers by writing down the memorized series for counting by fours. (b) An unsuccessful attempt to solve a problem by tallying. (c) Multiplication converted to repeated addition.



Extended Data Fig. 2 | Impact of hints on working children's performance in study 3 wave 2. (a) Proportion of children who correctly answered abstract questions in the anchoring hints section of the survey. (b) Proportion of children who correctly answered abstract and anchored multiplication questions in the rounding hints section of the survey. (c) Proportion of children who correctly answered a complex abstract or anchored question in the process hints section of the survey. Error bars show 95% confidence intervals around the mean (mean ± 1.96 × s.e.m). P-values are calculated using two-tailed t-tests without adjustments for multiple comparisons.

Extended Data Table 1 | Proportion of children who calculated the amount due and change correctly in market transactions in study 1

	(1) All	(2) Out of school	(3) In school	(4) Col. (3)-(2)
A. Transaction 1				
Amount due and change correct on first try	0.88	0.87	0.88	0.01
SD/SE	(0.33)	(0.34)	(0.33)	(0.05)
t-statistic				0.25
CI				[-0.09, 0.11]
p-value				0.80
Amount due and change correct on second try	0.07	0.03	0.09	0.06
SD/SE	(0.26)	(0.18)	(0.29)	(0.04)
t-statistic				1.45
CI				[-0.02, 0.14]
p-value				0.15
B. Transaction 2				
Amount due and change correct on first try	0.92	0.90	0.92	0.02
SD/SE	(0.28)	(0.30)	(0.27)	(0.04)
t-statistic				0.51
CI				[-0.06, 0.11]
p-value				0.61
Amount due and change correct on second try	0.05	0.05	0.05	-0.00
SD/SE	(0.22)	(0.22)	(0.22)	(0.03)
t-statistic				-0.01
CI				[-0.07, 0.07]
p-value				0.99
C. Transaction 3				
Amount due and change correct on first try	0.91	0.85	0.93	0.08
SD/SE	(0.29)	(0.36)	(0.26)	(0.04)
t-statistic				1.76
CI				[-0.01, 0.17]
p-value				0.08
Amount due and change correct on second try	0.07	0.13	0.04	-0.09
SD/SE	(0.26)	(0.34)	(0.20)	(0.04)
t-statistic				-2.33
CI				[-0.17, -0.01]
p-value				0.02
N	201	60	141	201

Columns 1-3 show means and standard deviations for each group. Column 4 shows differences in means and the associated standard errors, t-statistics, confidence intervals, and p-values, using two-tailed t-tests without adjustments for multiple comparisons.

Extended Data Table 2 | Number of children working in markets approached and surveyed

	(1) Study 1	(2) Study 2	(3) Study 3: Wave 1	(4) Study 3: Wave 2
Children approached	285	442	489	561
Children dropped	84 (29%)	42 (10%)	117 (23%)	98 (17%)
Children did not consent	84	37	77	81
Children consented, but refused midway	0	5	40	17
Children surveyed	201	400	372	463
Children with all transactions	201 (71%)	344 (86%)	372 (100%)	-

Columns 1-4 show the number and percentage of working children in each category. In studies 1 and 2, we attempt to perform three market transactions with children. In study 3 wave 1, we perform one market transaction. In study 3, wave 2, we do not perform a market transaction with children but instead ask the child about the cost of a good at a particular quantity to check they are indeed handling transactions.

Extended Data Table 3 | Proportion of children in each category of the ASER test by incentive group in study 2

	(1) No incentive	(2) Incentive	(3) Col. (2)-(1)
Division level	0.14	0.16	0.01
SD/SE	(0.35)	(0.37)	(0.04)
t-statistic			0.37
CI			[-0.06, 0.08]
p-value			0.71
Subtraction or higher level	0.50	0.53	0.03
SD/SE	(0.50)	(0.50)	(0.05)
t-statistic			0.65
CI			[-0.07, 0.13]
p-value			0.52
Two-digit number or higher level	0.92	0.90	-0.01
SD/SE	(0.27)	(0.29)	(0.03)
t-statistic			-0.51
CI			[-0.07, 0.04]
p-value			0.61
One-digit number or higher level	0.99	0.99	0.01
SD/SE	(0.12)	(0.07)	(0.01)
t-statistic			0.79
CI			[-0.01, 0.03]
p-value			0.43
Chi-square test-statistic			2.01
d.f.			4
p-value			0.73
N	222	178	400

Columns 1 and 2 show means and standard deviations for each group. Column 3 shows differences in means and the associated standard errors, t-statistics, confidence intervals, and p-values, using two-tailed t-tests without adjustments for multiple comparisons. The chi-square test row tests the equality of distribution of the ASER test scores in both samples, showing the chi-square statistics and associated p-values (in bracket). All levels are cumulative. Children are classified in division level if they answer two subtractions and a division correctly. They are in subtraction level if they answer two subtractions correctly but cannot answer the division correctly. They are classified in two-digit number recognition level if they can answer one or neither of two subtractions correctly and can recognize at least five two-digit numbers. They are classified in one-digit number recognition level if they can answer one or neither of two subtractions correctly, can recognize fewer than five two-digit numbers, and can recognize at least five one-digit numbers. They are classified as beginners if they cannot do any of the tasks listed above. All children could choose the division and subtractions problems that they wanted to solve and the numbers that they wanted to recognize. Working children were randomly assigned to performance incentives. For more details, see Supplementary Methods and the links included there to the assessment protocols.

Extended Data Table 4 | Comparison of calculation methods across hypothetical transactions for working children in study 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Hypothetical Transaction					Differences				
	First	Second	Third	Fourth	Fifth	Col. (1) - (2)	Col. (2) - (3)	Col. (3) - (4)	Col. (4) - (5)	Col. (3) - (5)
Used pen and paper	0.04	0.04	0.29	0.30	0.41	-0.00	-0.25	-0.01	-0.11	-0.12
SD/SE	(0.19)	(0.20)	(0.46)	(0.46)	(0.49)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)
t-statistic						-0.18	-10.19	-0.31	-3.20	-3.51
CI						[-0.03, 0.02]	[-0.30, -0.20]	[-0.07, 0.05]	[-0.17, -0.04]	[-0.18, -0.05]
p-value						0.85	<0.01	0.76	<0.01	<0.01
Used fingers to count	0.11	0.17	0.23	0.21	0.23	-0.06	-0.06	0.02	-0.02	0.00
SD/SE	(0.32)	(0.38)	(0.42)	(0.41)	(0.42)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
t-statistic						-2.34	-2.13	0.68	-0.60	0.08
CI						[-0.11, -0.01]	[-0.12, -0.00]	[-0.04, 0.08]	[-0.08, 0.04]	[-0.06, 0.06]
p-value						0.02	0.03	0.50	0.55	0.93
Solved with easier numbers	0.01	0.03	0.12	0.20	0.12	-0.02	-0.10	-0.08	0.09	0.01
SD/SE	(0.10)	(0.16)	(0.33)	(0.40)	(0.32)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)
t-statistic						-1.83	-5.18	-3.08	3.41	0.33
CI						[-0.04, 0.00]	[-0.13, -0.06]	[-0.13, -0.03]	[0.04, 0.14]	[-0.04, 0.05]
p-value						0.07	<0.01	<0.01	<0.01	0.74
Added instead of multiplying	0.34	0.37	0.35	0.31	0.31	-0.03	0.03	0.04	-0.01	0.03
SD/SE	(0.47)	(0.48)	(0.48)	(0.46)	(0.46)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
t-statistic						-0.89	0.74	1.21	-0.23	0.98
CI						[-0.10, 0.04]	[-0.04, 0.09]	[-0.02, 0.11]	[-0.07, 0.06]	[-0.03, 0.10]
p-value						0.38	0.46	0.23	0.82	0.33
N	400	400	400	400	400	800	800	800	800	800

Columns 1-5 show means and standard deviations for each transaction. Columns 6-10 show the differences in solution methods between problems and their associated standard errors (in parentheses), t-statistics, 95% confidence intervals, and p-values using two-tailed t-tests without adjustments for multiple comparisons.

Extended Data Table 5 | Proportion of study participants with selected characteristics in study 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Working children					
	All	Out of school	In school	Col. (3)-(2)	School children	Col. (5)-(1)
Female	0.16	0.06	0.23	0.17	0.09	-0.08
SD/SE	(0.37)	(0.23)	(0.42)	(0.04)	(0.28)	(0.03)
t-statistic				4.74		-2.62
CI				[0.10, 0.25]		[-0.14, -0.02]
p-value				<0.01		0.01
Age	14.01	14.96	13.39	-1.57	13.68	-0.34
SD/SE	(2.15)	(1.88)	(2.09)	(0.21)	(1.34)	(0.17)
t-statistic				-7.63		-2.03
CI				[-1.97, -1.17]		[-0.66, -0.01]
p-value				<0.01		0.04
Attends school	0.61	0.00	1.00	1.00	1.00	0.40
SD/SE	(0.49)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
t-statistic				-		11.41
CI				-		[0.33, 0.46]
p-value				-		<0.01
Current grade (if in school)			7.42		8.50	1.08
SD/SE			(2.23)		(0.50)	(0.16)
t-statistic						6.71
CI						[0.76, 1.39]
p-value						<0.01
Attends tuition	0.31	0.08	0.45	0.38	0.58	0.28
SD/SE	(0.46)	(0.27)	(0.50)	(0.04)	(0.49)	(0.04)
t-statistic				8.76		6.85
CI				[0.29, 0.46]		[0.2, 0.36]
p-value				<0.01		<0.01
Never enrolled in school	0.08	0.19	0.00	-0.19		
SD/SE	(0.26)	(0.39)	(0.00)	(0.03)		
t-statistic				-7.51		
CI				[-0.24, -0.14]		
p-value				<0.01		
Highest grade		6.02				
SD/SE		(2.76)				
t-statistic						
CI						
p-value						
Dropped out before completing grade 4		0.14				
SD/SE		(0.35)				
t-statistic						
CI						
p-value						
Years selling in market	1.81	1.58	1.96	0.39		
SD/SE	(2.24)	(2.14)	(2.30)	(0.23)		
t-statistic				1.69		
CI				[-0.06, 0.84]		
p-value				0.09		
Sells by kg/ltr	0.45	0.54	0.38	-0.16		
SD/SE	(0.50)	(0.50)	(0.49)	(0.05)		
t-statistic				-3.18		
CI				[-0.26, -0.06]		
p-value				<0.01		
Worked at market before					0.22	
SD/SE					(0.42)	
t-statistic						
CI						
p-value						
Handled transactions before					0.01	
SD/SE					(0.12)	
t-statistic						
CI						
p-value						
N	400	158	242	400	200	600

Columns 1-3 and 5 show means and standard deviations for each group. Columns 4 and 6 show differences in means and the associated standard errors, t-statistics, confidence intervals, and p-values, using two-tailed t-tests without adjustments for multiple comparisons. Cells left blank indicate that data on that variable were not collected for a given group of participants. In Delhi study 2, children were asked their current grade if enrolled and their highest grade otherwise.

Extended Data Table 6 | Comparison of performance of working and non-working children on the exact same problems in study 3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Group means			Difference without controls			Difference with controls		
		Working children		Non-working vs in school working	Non-working vs out of school working	In school working vs out of school working	Non-working vs in school working	Non-working vs out of school working	In school working vs out of school working
	Non-working	In school	Out of school	Col. (1) - Col.	Col. (1) - Col.	Col. (2) - Col.	Col. (1) - Col.	Col. (1) - Col.	Col. (2) - Col.
Abstract	0.59	0.56	0.27	0.03	0.32	0.29	-0.01	0.31	0.25
CI	[0.55, 0.63]	[0.53, 0.60]	[0.23, 0.31]	[-0.04, 0.09]	[0.25, 0.39]	[0.22, 0.36]	[-0.09, 0.06]	[0.21, 0.41]	[0.17, 0.33]
SE	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)
t-statistic				0.88	9.49	8.37	-0.40	6.24	6.00
p-value				0.381	0.000	0.000	0.692	0.000	0.000
N	542	932	556	1474	1098	1488	1474	1098	1488
Market word problem	0.01	0.36	0.37	-0.34	-0.36	-0.02	-0.31	-0.35	0.02
CI	[0.00, 0.03]	[0.29, 0.42]	[0.29, 0.46]	[-0.40, -0.28]	[-0.44, -0.28]	[-0.12, 0.08]	[-0.39, -0.23]	[-0.47, -0.23]	[-0.10, 0.13]
SE	(0.01)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.06)	(0.06)
t-statistic				-10.58	-8.60	-0.35	-7.92	-5.59	0.30
p-value				0.000	0.000	0.730	0.000	0.000	0.766
N	271	233	139	504	410	372	504	410	372
Market word problem with hint	0.05	0.48	0.53	-0.43	-0.48	-0.05	-0.44	-0.42	0.02
CI	[0.01, 0.09]	[0.39, 0.58]	[0.41, 0.65]	[-0.54, -0.33]	[-0.60, -0.35]	[-0.20, 0.10]	[-0.55, -0.32]	[-0.61, -0.23]	[-0.15, 0.20]
SE	(0.02)	(0.05)	(0.06)	(0.05)	(0.06)	(0.08)	(0.06)	(0.10)	(0.09)
t-statistic				-8.32	-7.52	-0.61	-7.19	-4.38	0.28
p-value				0.000	0.000	0.545	0.000	0.000	0.781
N	102	112	70	214	172	182	214	172	182

Columns 1-3 show the proportion of correct answers for a given problem type along with the associated 95% confidence interval, standard error, and sample size. Columns 4-9 show differences in means and their associated 95% confidence interval, standard errors (in parentheses), t-statistic, p-values, and sample size using two-tailed t-tests without adjustments for multiple comparisons. Abstract shows children's performance on abstract subtraction problems on their first attempt. Market word problem shows children's performance on their first attempt of the market word problem. Market word problem with hint shows children's performance on the market word problem after two attempts when they receive a hint to breakdown the problem and make use of rounding strategies if they are initially incorrect.

Extended Data Table 7 | Impact of incentives on non-working children's performance in study 3b

Outcome	(1) N	(2) Control Mean	(3) Treatment	(4) CI	(5) SE	(6) t	(7) p
Standardized Index	271	-0.08	0.10	[-0.14, 0.34]	0.12	0.83	0.406
IRT: All questions	271	-0.07	0.09	[-0.13, 0.31]	0.11	0.82	0.415
IRT: Abstract arithmetic questions	271	-0.07	0.11	[-0.09, 0.32]	0.10	1.07	0.287
IRT: Anchored arithmetic questions	271	-0.06	0.08	[-0.10, 0.27]	0.10	0.87	0.388
IRT: Market questions	271	-0.03	0.03	[-0.09, 0.15]	0.06	0.54	0.586

Columns 1-7 show the sample size, control group mean, treatment effect, 95% confidence interval, t-statistics, standard error, and p-value (respectively) using two-tailed t-tests without adjustments for multiple comparisons. We show performance across using a standardized index with all questions, as well as a 1 parameter logit item response theory model index with all questions, abstract arithmetic questions, anchored arithmetic questions, and market questions. For market questions, we include the market word problem with hints, the market word problem without hints, and the pretend market question.

Extended Data Table 8 | Impact of process and rounding hints on non-working children's performance on the market word problem in study 3b

	(1)	(2)	(3)	(4)	(5)	(6)
	Group means			Differences		
	No hint	Process hint	Process + rounding	Process hint vs no hint Col (2) - Col (1)	Process + rounding hint vs no hint Col (3) - Col (1)	Process + rounding hint vs Process hint Col (3) - Col (2)
Market word problem	0.04	0.04	0.05	-0.01	0.00	0.01
CI	[0.00, 0.09]	[-0.01, 0.08]	[0.01, 0.09]	[-0.07, 0.05]	[-0.06, 0.06]	[-0.05, 0.07]
SE	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
t-statistic				-0.21	0.15	0.36
p-value				0.834	0.881	0.718
N	90	79	102	169	192	181

Columns 1-3 show the proportion of correct answers after two attempts for the market word problem along with their associated 95% confidence interval, standard errors, and sample size. Columns 4-6 show differences in means and their associated 95% confidence intervals, standard errors, t-statistics, p-values, and sample size using two-tailed t-tests without adjustments for multiple comparisons. In no hint children were allowed to reattempt the problem (if they were initially incorrect) without a hint. In process hint, they were given a hint to breakdown the problem into smaller parts. And in process + rounding hint, they were also given hints to use rounding strategies in addition to the process hints.

Corresponding author(s): Esther Duflo

Last updated by author(s): Oct 16, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Computer code was not used to collect data.
Data analysis	Stata 18 and R 4.4.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data is publicly available without restriction on the J-PAL dataverse on the Harvard-MIT data archive - https://doi.org/10.7910/DVN/VPVH00

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Our survey team recorded the gender of the child. We do not disaggregate our results by gender.
Reporting on race, ethnicity, or other socially relevant groupings	We do not report on race or ethnicity.
Population characteristics	We report the age, gender, education and market handling experience of all studies in the Supplementary Tables (S3, Extended Data Table 5, Table S24)
Recruitment	In both cities, a list of markets where children worked was drawn. In each market, all children who appeared to be 16 or younger and sold at least 2 different goods were approached. To establish the sample of school children, with visited nearby schools and interviewed children of a similar age. We present the proportion of children approached and the proportion who select into the study in Extended Data Table 2.
Ethics oversight	Committee on the Use of Humans as Experimental Subjects (COUHES) at Massachusetts Institute of Technology (MIT) (IRB1504007082, IRB1612798217, and IRB2208000727) and the Institutional Review Board at the Institute for Financial Management and Research (IFMR)(IRB00007107).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study uses quantitative methods to report on the ability of children working in markets and enrolled in school to complete different arithmetic tests.
Research sample	201 working children in Study 1 400 working children in Study 2 200 school children in Study 2 835 working children in Study 3 271 school children in Study 3 The samples are not representative. For working children, we recruited a sample of children working in informal markets in Delhi and Kolkata. The school children were selected from schools located close to the informal markets. They were selected to be of a similar age to the working children.
Sampling strategy	First, in both cities, a list of markets where children worked was drawn. In each of these markets, all children who appeared to be 16 or younger and sold at least 2 different goods were approached. To establish the sample of school children, with visited nearby schools and interviewed children of similar age.
Data collection	The market surveys in both studies had two main parts: market transactions and exercises. Market transaction data was collected by "mystery shoppers" who purchased two goods, handed a bill, and asked for the change. The children were approached and consented to participate in the exercise part which was conducted as a one and one interview. For school children, we conducted a single survey in their classroom.
Timing	Study 1: June 2015-May 2016 Study 2: February-May 2017 Study 3: November 2022 - February 2024
Data exclusions	Consistent with our analysis plan, we exclude school children in study 3b with market selling experience. A total of 27 children were excluded. We do this to compare math abilities of children with and without market-selling experience.

Non-participation	Study 1: 84 working children (of 285) working children refused to participate, Study 2: 42 working children refused (of 442) to participate. Study 3: 215 working children refused (of 1050). All school children agreed to participate.
Randomization	In study 1, subjects were not allocated to experimental groups. In study 2, we randomize the order of the presentation of (i) ASER written assessment, with the (ii) oral abstract, oral anchored, and hypothetical transaction. We also randomly assigned children to receive either a fixed payment or a monetary incentive for solving the problems correctly. In Study 3a wave 1, we randomized,(i) the order in which we presented the hypothetical and oral abstraction questions, (ii) whether we used encouraging language when introducing the survey (to boost the confidence of children), (iii) the oral abstract subtraction problems that children solved (iv) whether we provided children with a hint to solve abstract subtraction problems, (v) whether we provided children with a hint to solve a market word problem. In Study 3a wave 2, we randomized (i) which oral anchored and abstract problems children solved, (ii) the number of warm-up problems children solved, (iii) the numbers we used in the anchoring hint and rounding hint sections of the survey, (iv) whether we provided children with a hint to break down a complex problem into smaller parts. In Study 3b, we randomized (i) which written abstract, oral abstract, and oral anchored problems children solved, (ii) whether or not we provided children with incentives based on their performance, (iii) whether or not we provided children with a hint to solve a market word problem.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*