Children are less able than they used to be

That's the startling finding of research into how they develop maths and science skills. John Crace reports

Tuesday January 24, 2006
The Guardian

It has become an annual rite of summer. Out come the Sats/GCSE/A-level results - take your pick - and up pops a government minister to say that grades are higher than ever, teachers and schools have done a fantastic job, but there's still room for improvement. Not everyone takes this at face value and there are a few grumbles about exams becoming easier. But even if there are suspicions that standards have dropped, no one has ever seriously suggested that children's cognitive abilities have deteriorated. Until now.

New research funded by the Economic and Social Research Council (ESRC) and conducted by Michael Shayer, professor of applied psychology at King's College, University of London, concludes that 11- and 12-year-old children in year 7 are "now on average between two and three years behind where they were 15 years ago", in terms of cognitive and conceptual development.

"It's a staggering result," admits Shayer, whose findings will be published next year in the British Journal of Educational Psychology. "Before the project started, I rather expected to find that children had improved developmentally. This would have been in line with the Flynn effect on intelligence tests, which shows that children's IQ levels improve at such a steady rate that the norm of 100 has to be recalibrated every 15 years or so. But the figures just don't lie. We had a sample of over 10,000 children and the results have been checked, rechecked and peer reviewed."

To understand both the science and its implications, we need to step back 30 years, to when Shayer was part of a six-strong team of academics - including Margaret Brown, Geoffrey Matthews and Philip Adey - engaged in research at Chelsea College on concepts in secondary science and mathematics. "We realised that no one had actually bothered to investigate how children learned maths and science, or where the difficulties lay," he says. "So the Social Science Research Council (SSRC) funded a five-year project - the longest ever research programme of its kind - to find out."

As the time frame suggests, it was a slow process and Shayer has clear memories of a young, blue-suited academic - one Ted Wragg - being sent round after two years had elapsed to check up that the SSRC's money was being well spent. Wragg gave the Chelsea College team the thumbs up and in 1979 the research was published.

One of Shayer's main difficulties had been to establish a benchmark of ability. The psychometric tradition had obvious disadvantages. For one thing, the Flynn effect implied that an absolute scale of mental age was impossible, but there were other problems. A score of 105 might tell you that a child is slightly above average, but it does not tell you what maths he or she can or can't understand. For this reason, Shayer decided that using the developmental model of the Swiss psychologist Jean Piaget was a better bet.

"Although controversial, Piaget's descriptions do provide an underlying, logic-based, theoretical model to differentiate different levels of complexity," he says. "It describes the same behaviours - for example, the ability to control variables in experimenting - whether

the subject is nine or 16." Crucially, the model met the statistical demands of being criterion-referenced and could be given equal interval properties.

Four main stages

According to Piaget's model, children go through four main stages of development - sensorimotor (infancy), pre-concrete (up to age 5), concrete (5-11) and formal (11-16) - each of which are divided into several sub-groups. Shayer's first task was to check this model against a broad cohort of 14,000 schoolchildren. "We conducted a wide range of tests on all the secondary-age year groups over the course of a year," says Shayer. "These were designed to assess a child's exact ability on the Piagetian scale."

Shayer's work naturally focused on the different sub-groups of the concrete and the formal. The concrete stage, in regard to maths, meant testing a child's ability to put things in order, use descriptive models and plot simple graphs. The formal stage involved testing more abstract concepts and the ability to predict.

His results showed that Piaget had only described the top 20% of the population. "Like many scientists, Piaget picked the best specimens, so his results were weighted in favour of the most able children," says Shayer. "We took a broad section of the population and found that, far from being at the early formal level (3A) as Piaget had predicted, the average 11-year old was firmly back in the centre of the middle concrete level (2B)."

Not everyone was overjoyed by these findings. Many educationists found it hard to accept that children were less able than previously thought, and were reluctant to admit that there were huge differences in development that weren't purely attributable to environmental factors. To Shayer, though, it was no great surprise. "You would expect children of bright parents to be brighter than average," he says. "Similarly, you would expect children whose parents played with them regularly in a creatively challenging way to do better on developmental tests."

The main objection to Shayer's research came from those who argued that the Piagetian tests described only a child's ability to perform those particular functions and were of no predictive value with regard to general level of performance. "Shayer disproved this with his subsequent work in the 1980s," says Paul Black, emeritus professor of education at King's College and chair of the 1988 National Curriculum Task Group on Assessment and Testing (TGAT), whose report formed the basis for the implementation of Sats.

"He helped to develop two-year intervention programmes for those children who had been identified by the Piagetian model as being below average in year 7. Science and maths were the contexts through which the programmes were taught, but the prime focus was on general developmental skills.

"These programmes [Cognitive Acceleration through Science Education and Cognitive Acceleration through Maths Education] both significantly increased the children's Piagetian scores and markedly improved their maths and science GCSE grades from those predicted at entry level testing. More important, these children also showed an improvement on predicted grades in other subjects, such as English and history. This showed the programmes had a generic impact, rather than just a specific effect."

Shayer's work was subsequently validated by similar research in Greece, Pakistan and Australia. It also managed to free itself of its purely Piagetian approach by assimilating some of the properties of psychometric testing into a unified developmental test. It became

one of the criteria by which age-related attainment targets were benchmarked when the national curriculum was introduced in 1988.

And there the story would have ended were it not for the fact that Shayer's wife, scientist Denise Ginsburg, was regularly employed by schools to run their Year 7 maths and science developmental testing to see which children needed the Case or Came programmes.

"She reported to me that she had begun to notice a significant falling off in children's abilities," Shayer says, "and, because of this, I decided to investigate further." His research project was undertaken last year and involved the assessment of 10,000 year 7 children's performance on developmental volume and heaviness (VH) tests.

Objective method

VH, which concerns the conservation of liquid and solid materials, internal volume and intuitive density, was chosen partly because it has substantial predictive validity for both science and mathematics achievement and is an effective way of alerting teachers to their students' range of abilities, but also because it is recognised as a test that measures abilities that are not directly teachable. As such, it was an objective research method, free from any process of adaptation to changing circumstance.

"Similar tests conducted in the 70s showed a big difference between boys and girls," says Shayer, "with boys scoring noticeably better than girls. The new research reveals that the gender gap has disappeared, with both sexes deteriorating significantly. Boys have fallen by more than one Piagetian sub group - from the middle of 2B [mature concrete] to below the middle of 2A/2B [middle concrete]. By any standards, this is a huge and significant statistical change."

For the same reasons that he stood by Shayer's original research, Black believes no one should dismiss these current findings. "There are bound to be those who would prefer to ignore these results," he says, "because they find them politically unacceptable or inconvenient. But Shayer has a proven track record and you have to respect his science."

Those likely to be particularly discomforted by Shayer's findings are people who swear by the validity of GCSE and Sats results. The idea that most children are achieving the government level 4 targets in maths and science at key stage 2 is clearly anomalous with Shayer's findings, as is the notion that secondary schools are now taking children who are two years behind developmentally and still getting them up to GCSE speed in just five years.

So how does Shayer explain this? "The Qualifications and Curriculum Authority obviously insists that standards haven't dropped," he says, "but this doesn't fit all the evidence. A-level maths and science teachers often report that their students don't know as much as they used to. And some parts of the GCSE science syllabus, such as density, have been dropped. Examiners may well be asking easier questions and marking more leniently. These things can happen unconsciously.

"There is some evidence that the extra hour allocated to maths in primary schools under the numeracy initiative has had some impact on Sats scores, but there is greater evidence of teachers teaching to the tests. This means students can perform well in the tests without necessarily understanding the underlying concepts."

Black does not demur. "Research from around the world shows that, when the stakes are high, teachers teach to the tests," he says. "This produces a short-term, three-year uplift in results before they plateau. We also get to see some artificial results, such as in the US, where every state is above the national average in its test scores. In the UK, the National Audit Office has questioned the validity of some Sats scores.

"We also need to ask what the tests test. Do they measure what's important or what doesn't matter?"

In the end, as Black points out, it comes down to a value judgment as to which tests are most worthwhile. Shayer believes that his have the objective science to back them up. But, if he's right, then the question that must be answered is why children's developmental skills have fallen off so much. Shayer is reluctant to be drawn into specifics.

"We can speculate," he says, "but there's no hard evidence. I would suggest that the most likely reasons are the lack of experiential play in primary schools, and the growth of a video-game, TV culture. Both take away the kind of hands-on play that allows kids to experience how the world works in practice and to make informed judgments about abstract concepts."

This research project is Shayer's swansong; after a long and distinguished career, he's heading for retirement and his prime concern is a forthcoming holiday to Egypt. Having opened the can of worms, he's happy to stand back and leave the field clear for others to seek solutions.