Learning and Scientific Reasoning

Comparisons of Chinese and U.S. students show that content knowledge and reasoning skills diverge.

Lei Bao,[1]* Tianfan Cai,[2] Kathy Koenig,[3] Kai Fang,[4] Jing Han,[1] Jing Wang,[1] Qing Liu,[1] Lin Ding,[1] Lili Cui,[5] Ying Luo,[6] Yufeng Wang,[2] Lieming Li,[7] Nianle Wu[7]

The development of general scientific abilities is critical to enable students of science, technology, engineering, and mathematics (STEM) to successfully handle open-ended real-world tasks in future careers (*1–6*). Teaching goals in STEM education include fostering content knowledge and developing general scientific abilities. One such ability, scientific reasoning (*7–9*), is related to cognitive abilities such as critical thinking and reasoning (*10–14*). Scientific-reasoning skills can be developed through training and can be transferred (*7*, *13*). Training in scientific reasoning may also have a long-term impact on student academic achievement (*7*). The STEM education community considers that transferable general abilities are at least as important for students to learn as is the STEM content knowledge (*1–4*). Parents consider science and mathematics to be important in developing reasoning skills (*15*).

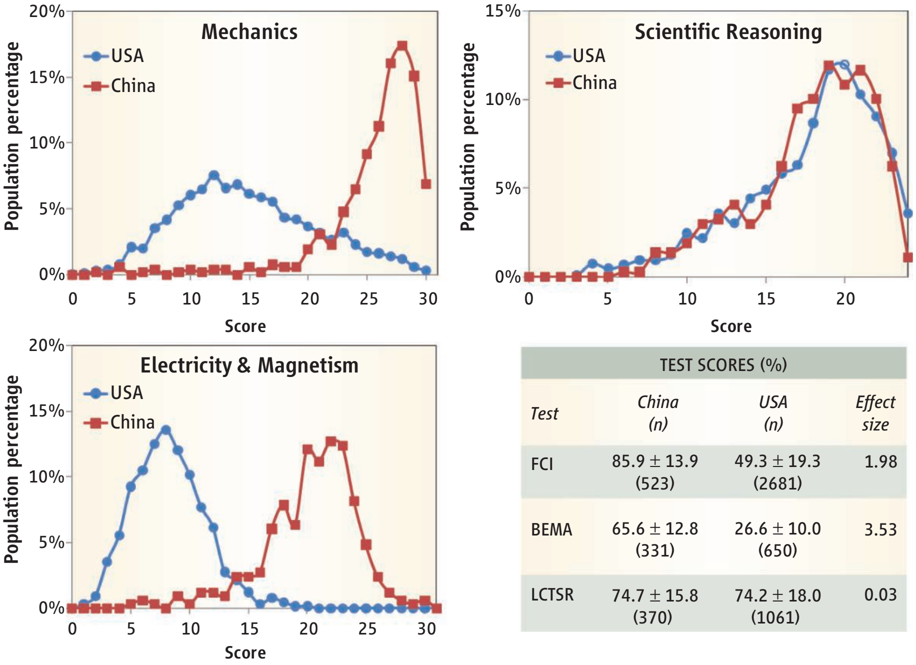
We therefore asked whether learning STEM content knowledge does in fact have an impact on the development of scientific-reasoning ability. The scientific-reasoning ability studied in this paper focuses on domain-general reasoning skills such as the abilities to systematically explore a problem, to formulate and test hypotheses, to manipulate and isolate variables, and to observe and evaluate the consequences.

Research Design

Students in China and the United States go through very different curricula in science and mathematics during their kindergarten through 12th grade (K–12) school years. This provides systemically controlled long-term variation on STEM content learning, which we used to study whether or not such learning has any impact on the development of scientific-reasoning ability. Scientific reasoning is not explicitly taught in schools in either country.

In China, K–12 education is dominated by the nationwide college admission exam given at the end of grade 12. To comply with the requirements of this exam, all Chinese schools adhere to a national standard within all courses. In physics, for example, every student goes through the same physics courses, which start in grade 8 and continue every semester through grade 12, providing 5 years of continuous training on introductory physics topics (*16*). The courses are algebra-based with emphasis on development of conceptual understanding and skills needed to solve problems.

In contrast, K–12 physics education in the United States is more varied. Although students study physics-related topics within other general science courses, only one of three high school students enrolls in a two-semester physics course (*17*). As a result, the amount of instructional time and the amount of emphasis on conceptual physics understanding and problem-solving skills are very different in the two countries. Similar curriculum differences between the United States and China are reflected in other STEM areas such as chemistry, biology, and mathematics (*16*).



TEST SCORES (%)			
Test	*China* *(n)*	*USA* *(n)*	*Effect size*
FCI	85.9 ± 13.9 (523)	49.3 ± 19.3 (2681)	1.98
BEMA	65.6 ± 12.8 (331)	26.6 ± 10.0 (650)	3.53
LCTSR	74.7 ± 15.8 (370)	74.2 ± 18.0 (1061)	0.03

Content knowledge and reasoning skills diverge. Comparisons of U.S. and Chinese freshmen college students show differences on tests of physics content knowledge but not on tests of scientific reasoning.

Chinese students go through rigorous problem-solving instruction in all STEM subject areas throughout most of their K–12 school years and become skillful at solving content-based problems. It remains unclear, however, whether this training is transferable beyond the specific content areas and problem types taught.

We used quantitative assessment instruments (described below) to compare U.S. and Chinese students' conceptual understanding in physics and general scientific-reasoning ability. Physics content was chosen because the subject is conceptually and logically sophisticated and is commonly emphasized in science education (*15*). Assessment data were collected from both Chinese and U.S. freshmen college students before college-level physics instruction. In this way the data reflect students' knowledge

---

[1]Department of Physics, The Ohio State University, Columbus, OH 43210, USA. [2]Department of Physics, Beijing Jiaotong University, Beijing 100044, China. [3]Department of Physics, Wright State University, Dayton, OH 45435, USA. [4]Department of Physics, Tongji University, Shanghai 200092, China. [5]Department of Physics, University of Maryland, Baltimore County, Baltimore, MD 21250, USA. [6]Department of Physics, Beijing Normal University, Beijing 100875, China. [7]Department of Physics, Tsinghua University, Beijing 100084, China.

*Author for correspondence. E-mail: bao.15@osu.edu

Published by AAAS

and skill development from their formal and informal K–12 education experiences.

Data Collection and Analysis

From the early 1980s, researchers and educators in psychology and cognitive science (*11–14*) have developed many quantitative instruments that assess reasoning ability. Some are included as components in standard assessments such as the Graduate Record Examination, whereas others are stand-alone tests such as Lawson's *Classroom Test of Scientific Reasoning* (LCTSR) (*8*, *9*). We used the LCTSR because of its popularity among STEM educators and researchers. Common categories of reasoning ability assessments include proportional reasoning, deductive and inductive reasoning, control of variables, probability reasoning, correlation reasoning, and hypothesis evaluation, all of which are crucial skills needed for a successful career in STEM.

Research-based standardized tests that assess student STEM content knowledge are also widespread. For example, in physics, education research has produced many instruments. We used the *Force Concept Inventory* (FCI) (*18*, *19*) and the *Brief Electricity and Magnetism Assessment* (BEMA) (*20*). These tools are regularly administered by physics education researchers and educators to evaluate student learning of specific physics concepts.

Using FCI (mechanics), BEMA (electricity and magnetism), and LCTSR (scientific reasoning), we collected data (see figure, page 586) from students ($N = 5760$) in four U.S. and three Chinese universities. All the universities were chosen to be of medium ranking (*15*). The students tested were freshmen science and engineering majors enrolled in calculus-based introductory physics courses. The tests were administered before any college-level instruction was provided on the related content topics. The students in China used Chinese versions of the tests, which were first piloted with a small group of undergraduate and graduate students ($n = 22$) to remove language issues.

The FCI results show that the U.S. students have a broad distribution in the medium score range (from 25 to 75%). This appears to be consistent with the educational system in the United States, which produces students with a blend of diverse experiences in physics learning. In contrast, the Chinese students had all completed an almost identical extensive physics curriculum spanning five complete years from grade 8 through grade 12. This type of education background produced a narrow distribution that peaks near the 90% score.

For the BEMA test, the U.S. students have a narrow distribution centered a bit above the chance level (chance 20%). The Chinese students also scored lower than their performance on the FCI, with the score distribution centered around 70%. The lower BEMA score of students in both countries is likely due to the fact that some of the topics on the BEMA test (for example, Gauss's law) are not included in standard high school curricula.

The FCI and BEMA results suggest that numerous and rigorous physics courses in the middle and high school years directly affect student learning of physics content knowledge and raise students to a fairly high performance level on these physics tests.

The results of the LCTSR test show a completely different pattern. The distributions of the Chinese and U.S. students are nearly identical. Analyses (*15*) suggest that the similarities are real and not an artifact of a possible ceiling effect. The results suggest that the large differences in K–12 STEM education between the United States and China do not cause much variation in students' scientific-reasoning abilities. The results from this study are consistent with existing research, which suggests that current education and assessment in the STEM disciplines often emphasize factual recall over deep understanding of science reasoning (*2*, *21–23*).

What can researchers and educators do to help students develop scientific-reasoning ability? Relations between instructional methods and the development of scientific reasoning have been widely studied and have shown that inquiry-based science instruction promotes scientific-reasoning abilities (*24–29*). The current style of content-rich STEM education, even when carried out at a rigorous level, has little impact on the development of students' scientific-reasoning abilities. It seems that it is not what we teach, but rather how we teach, that makes a difference in student learning of higher- order abilities in science reasoning. Because students ideally need to develop both content knowledge and transferable reasoning skills, researchers and educators must invest more in the development of a balanced method of education, such as incorporating more inquiry-based learning that targets both goals.

Our results also suggest a different interpretation of assessment results. As much as we are concerned about the weak performance of American students in TIMSS and PISA (*30*, *31*), it is valuable to inspect the assessment outcome from multiple perspectives. With measurements on not only content knowledge but also other factors, one can obtain a more holistic evaluation of students, who are indeed complex individuals.

References and Notes

1. R. Iyengar *et al.*, *Science* **319**, 1189 (2008).
2. A. Y. Zheng *et al.*, *Science* **319**, 414 (2008).
3. B. S. Bloom, Ed., *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain* (David McKay, New York 1956).
4. National Research Council (NRC), *National Science Education Standards* (National Academies Press, Washington, DC, 1996).
5. NRC, *Learning and Understanding: Improving Advanced Study of Mathematics and Science in U.S. High Schools* (National Academies Press, Washington, DC, 2002).
6. H. Singer, M. L. Hilton, H. A. Schweingruber, Eds., *America's Lab Report* (National Academies Press, Washington, DC, 2005).
7. P. Adey, M Shayer, *Really Raising Standards: Cognitive Intervention and Academic Achievement* (Routledge, London, 1994).
8. A. E. Lawson, *J. Res. Sci. Teach*. **15**, 11 (1978).
9. Test used in this study was *Classroom Test of Scientific Reasoning*, rev. ed. (2000).
10. P. A. Facione, *Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment* (California Academic Press. Millbrae, CA, 1991).
11. H. A. Simon, C. A. Kaplan in *Foundations of Cognitive Sciences*, M. I. Posner, Ed. (MIT Press, Cambridge, MA, 1989), pp. 1–47.
12 R. E. Nisbett, G. T. Fong, D. R. Lehman, P. W. Cheng, *Science* **238**, 625 (1987).
13. Z. Chen, D. Klahr, *Child Dev.* **70**, 1098 (1999).
14. D. Kuhn, D. Dean, *J. Cognit. Dev.* **5**, 261 (2004).
15. See Supporting Online Material for more details.
16. This is based on the Chinese national standards on K–12 education (www.pep.com.cn/cbfx/cpml/).
17. J. Hehn, M. Neuschatz, *Phys. Today* **59**, 37 (2006).
18. D. Hestenes, M. Wells, G. Swackhamer, *Phys. Teach.* **30**, 141 (1992).
19. The test used in this study is the 1995 version.
20. L. Ding, R.Chabay, B. Sherwood, R. Beichner, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105 (2006).
21. M. C. Linn *et al.*, *Science* **313**, 1049 (2006).
22. A. Schoenfeld, *Educ. Psychol.* **23**, 145 (1988).
23. A. Elby, *Am. J. Phys*. **67**, S52 (1999).
24. C. Zimmerman, *Dev. Rev.* **27**, 172 (2007).
25. P. Adey, M. Shayer, *J. Res. Sci. Teach.* **27**, 267 (1990).
26. A. E. Lawson, *Science Teaching and the Development of Thinking* (Wadsworth, Belmont, CA, 1995).
27. R. Benford, A. E. Lawson, *Relationships Between Effective Inquiry Use and the Development of Scientific Reasoning Skills in College Biology Labs* (Arizona State University, Tempe, AZ, 2001); Educational Resources Information Center (ERIC) accession no. ED456157.
28. E. A. Marek, A. M. L. Cavallo, *The Learning Cycle and Elementary School Science* (Heinemann, Portsmouth, NH, 1997).
29. B. L. Gerber, A. M. Cavallo, E. A. Marek, *Int. J. Sci. Educ.* **23**, 5359 (2001).
30. Trends in International Mathematics and Science Study (TIMSS), http://nces.ed.gov/timss/.
31. Programme for International Student Assessment (PISA), www.pisa.oecd.org/.

32. We wish to thank all the teachers who helped with this research.

Supporting Online Material

www.sciencemag.org/cgi/content/full/323/5914/586/DC1

10.1126/science.1167740