

## The Health Sciences and Technology Academy: Utilizing Pre-college Enrichment Programming to Minimize Post-secondary Education Barriers for Underserved Youth

SHERRON BENSON MCKENDALL, PRISCAH SIMOYI, ANN L. CHESTER, and JAMES A. RYE

West Virginia is considered one of the most rural states in the nation, with over 60% of its population classified as rural.<sup>1</sup> The state experiences relatively high unemployment, and it ranks among the lowest (49th) of all states in median household income.<sup>2</sup> Fifty-eight percent of the students in West Virginia counties are eligible for free or reduced-price lunch.<sup>3</sup> Furthermore, only 14.7% of adult residents 25 years and over have attained a bachelor's degree or higher,<sup>4</sup> putting the state 50th in higher education.

The rural nature of the state coupled with economically depressed communities has limited the availability of secondary-level science courses required for health sciences majors in college. Additionally, most counties in West Virginia are considered medically underserved, and therefore it is important to increase the number of health care providers in rural areas of the state.<sup>5</sup> However, if the state's under-represented students do not receive adequate preparation in pre-college math and science, the proportion who can attend college and succeed will continue to be limited,<sup>4</sup> and the pool for the health professions will be too small.

To overcome some of these barriers, West Virginia University and 21 West Virginia counties have come together in the Health Sciences and Technology Academy (HSTA) in a community-campus partnership. Its web site is (<http://www.wv-hsta.org>). A pre-college enrichment program,<sup>6</sup> HSTA helps students learn tools to enable them to progress through high school, college, and professional school. The HSTA program consists of an on-campus (WVU) Summer Institute at West Virginia University where students and science teachers are engaged in learning activities facilitated by science and education faculty. These science teachers also facilitate HSTA community-based science clubs during the school year. The HSTA model uses the inquiry-based theory that encompasses problem posing, problem solving, and persuasion.<sup>7,8</sup> Research suggests that inquiry activities emphasizing problem solving enhance middle-level students' self-confidence in mastering science and their attitudes towards the discipline.<sup>9</sup> Furthermore, inquiry-based learning is considered fundamental to students' understanding of science concepts and processes. The National Science Education Standards (NSES) call for greater emphasis on "inquiry into authentic questions generated from student experiences [which] is the central strategy for teaching science."<sup>10</sup> As a follow up to the NSES, a practical guide has been developed for educators who wish to emphasize inquiry-based instruction.<sup>11</sup> A principal thrust within the community science clubs is inquiry-based learning of science through extended investigations and community service projects.<sup>12</sup> The model also engages students in authentic learning processes (i.e., real-world problem-solving circumstances), which are both fun and challenging.<sup>13,14</sup> Students' projects often target health-related topics and may potentially inform and benefit various communities through dissemination at local and state levels.<sup>15</sup>

### Methods

HSTA's effect on the academic success of its graduates and their decisions to pursue post-secondary studies and/or health sciences majors was assessed using quantitative and qualitative methods. HSTA participants are selected based on at least two of the following criteria: African American, financially disadvantaged, rural, and first generation aiming for higher education. Participants are ad-

mitted to HSTA during the ninth grade and participate in various activities until they graduate from high school, at which time they are considered HSTA graduates. There are 35 and 61 HSTA graduates for the 1998 and 1999 academic terms, respectively.

Telephone interviews were conducted in the fall and spring of the 1999–2000 academic term. Graduates were asked a series of questions that employed a Likert-type scale regarding HSTA's impact on pursuit of post-secondary study (1 = no impact to 5 = very high impact), choosing a health sciences major (1 = no impact to 5 = very high impact), preparation for college (1 = not at all prepared to 3 = extremely prepared), and preparation for major (1 = not at all prepared to 3 = extremely prepared). The participants were also asked to briefly explain why they had rated the program's impact and preparation levels as such.

In the fall of 1999, HSTA's impact on graduates' college performances was assessed with an independent *t*-test comparing the mean ( $\mu$ ) grade-point average (GPA) of 40 HSTA students (experimental) with that of 120 non-HSTA students at West Virginia University (WVU) in Morgantown, West Virginia. The 120 non-HSTA students were randomly selected from those enrolled at WVU with the same status (e.g. freshman), declared major, and residency. In order to achieve an effective sample size based on these characteristics (i.e., status, major, WV residency), three controls were matched to each experimental case.

### Results and Discussion

*Interviews.* The return rates for questionnaires on pursuing post-secondary study and preparation level were 97% (93 students) and 80% (77) for the combined cohorts (1998 and 1999).

*Post-secondary Study.* The graduates' responses about HSTA's impact on their decisions to pursue post-secondary study indicate a strong impact, 3.88 and 3.96 (5 = very high) for the 1998 and 1999 graduates, respectively. The graduates provided a variety of reasons. One stated, "Before HSTA I didn't even know I could go to college because I'm from a poor family, and they gave me the chance to go to college." Another graduate affirmed that HSTA was the reason for her being in college. She posited that

First of all, I didn't think I would be able to go to college and I really—I didn't have anybody in my family that said okay here is where I'm going, you ought to check this stuff out. When I came up here, I fell in love with the [WVU] campus. . . . And I got in and it's nice to have contacts . . . I knew a lot of the teachers and a lot of the faculty through HSTA and it really helped me out a lot.

A graduate who intended to pursue a nursing career reported that "when I was in high school and worked with HSTA for the summer, we got to work with the cadavers. That's hands-on experience that I'd never have had." Essentially, HSTA provides students with tangible experiences that bring excitement to the learning process. Not only is the program a pipeline for participants who wish to pursue post-secondary study, it also provides financial support for students who would not have had the opportunity to attend college.

*College Major.* Approximately 66% (23) of the 1998 graduates and 80% (49) of the 1999 graduates chose health sciences majors.

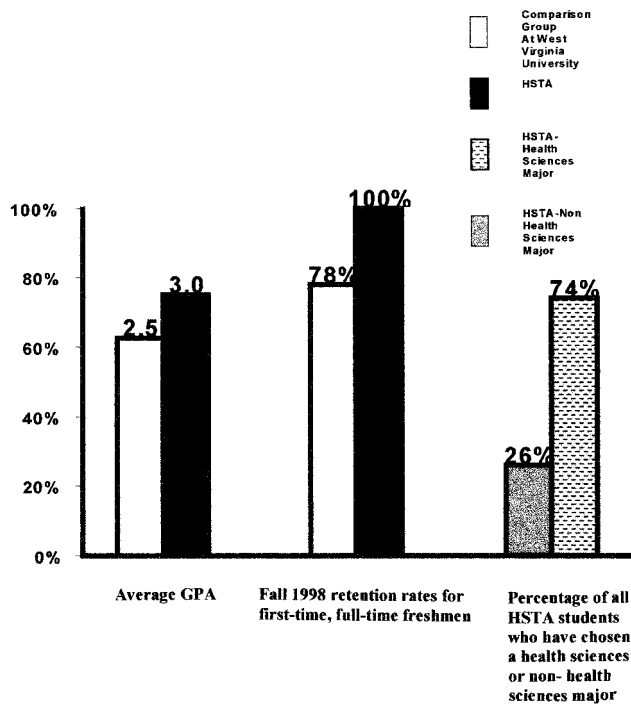


Figure 1. The Health Sciences and Technology Academy (HSTA) graduates' GPAs, their retention rates at West Virginia University compared with a non-HSTA group at the university, and their choices of major in higher education, fall and spring 1999–2000.

Among these graduates, the impact of HSTA on this decision ranged from moderate to high, 3.60 (1998 cohort) and 3.74 (1999 cohort) (5 = very high). The graduates rated the program highly because of the hands-on learning experiences it had afforded them. For example, one graduate stated

Whenever we . . . would do hands on experiences, it just made me more interested, especially in Psychology because when we go to mess with the brains . . . it just made me more interested. It made it not seem as hard or as bad as what people think it is.

Another graduate reported that, "HSTA allowed me to see a lot of different areas in the health field that I wouldn't have seen otherwise. It kind of gave me a taste of everything and just sort of oriented me." Overall, these experiences not only expose students to various occupations of which they would otherwise have no knowledge, but it provides the opportunity to explore horizons within the realm of health sciences.

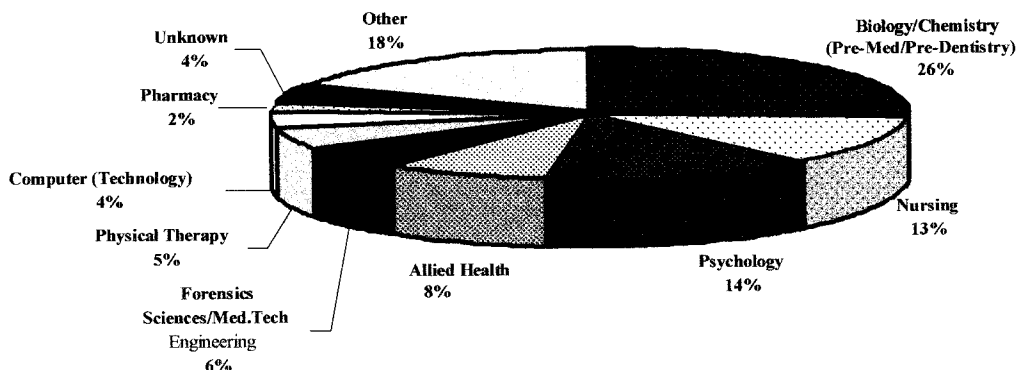


Figure 2. Percentages of Health Sciences and Technology Academy (West Virginia University) graduates choosing different higher education majors, 1998 and 1999.

**College Preparation.** The graduates responded positively regarding level of preparation for college and major as a result of their participation in HSTA. Of both cohorts, 98% (94) were pursuing post-secondary study. In response to questions about college preparation, the mean responses were 2.45 and 2.46 (3 = extremely prepared) for the 1998 and 1999 graduates, respectively.

The graduates rated the program's preparation for their majors as 1.95 (1998 cohort) and 2.27 (1999 cohort) (3 = extremely prepared). Thus, the overall perception is that HSTA prepared them at least moderately for their majors. The higher rating given by the 1999 graduates may be due to the higher percentage of them who intended to pursue health sciences majors.

**College Performance.** The HSTA graduates had a significantly higher undergraduate GPA (population mean  $\mu$ ) undergraduate GPA of 3.00) than the non-HSTA control group's mean GPA of 2.51. An independent *t*-test comparing the mean GPA of HSTA graduates with that of non-HSTA students at West Virginia University proved that there is a statistically significant difference between the GPAs ( $\alpha = .05$ ,  $p = .0014$ ,  $t = 3.2495$ ). The result exemplifies HSTA's impact on those who matriculate to and graduate from the program. After the *t* test was performed, a 99% confidence interval of the true  $\mu$  of the non-HSTA population was determined. The true  $\mu$  of the non-HSTA population is between 2.31 and 2.71. Thus, we are 99% confident that the true mean GPA for the non-HSTA student population lies in the interval [2.31, 2.71]. Therefore, the true  $\mu$  GPA (3.00) of HSTA students who attend WVU is not only higher than that of the control group (2.51) but also higher than that of the total non-HSTA population.

**Retention.** All HSTA's graduates who enrolled at WVU during the fall of 1998 were retained, compared with a rate of 78% for non-HSTA first-time freshmen (see Figure 1). Furthermore, an overwhelming 74% of 1998 and 1999 HSTA graduates are pursuing health sciences majors, compared with 26% of the graduates who have chosen other fields of study (see Figure 1). The graduates majoring in health sciences are particularly drawn to fields such as biology/chemistry, nursing, psychology, and allied health (see Figure 2).

## Conclusion and Implications

The Health Sciences and Technology Academy provides a pipeline for underrepresented youth to pursue their higher education goals. Through pre-college enrichment measures, HSTA gives students multifaceted opportunities for academic enrichment, which help them to realize that they can become accomplished individuals in their communities and in society at large. Pre-college programs such as HSTA can provide enriching experiences for underprivileged students who may not foresee the importance of completing high school and going to college.

Although HSTA has not provided its graduates with enrichment

experiences beyond high school, it can be assumed that the success of these students at WVU can be attributed, in part, to the pre-college enrichment provided by HSTA. Many of the graduates at WVU and other higher education institutions express a deep sense of fulfillment as a result of their participation in HSTA. Furthermore, many have expressed that their desires to pursue health sciences as well as technologic careers are due, in part, to the HSTA program. Their performance relative to that of non-HSTA students with similar interests is extremely encouraging. We believe that the HSTA model provides an exciting opportunity to extend inquiry-based learning, via longitudinal science projects, beyond what otherwise would be possible in the science classroom. All evidence indicates that the long-term benefits of this pre-college enrichment program will be positive.

Correspondence: Sherron Benson McKendall, PhD, Health Sciences and Technology Academy, PO Box 9026, Robert C. Byrd Health Sciences Center, Morgantown, WV 26506. Reprints are not available.

---

#### References

1. Table 1. Urban and Rural Population: 1990 to 1990. (<http://www.census.gov/population/censusdata/urpop0090.txt>). Accessed 3/99. Washington, DC: U.S. Census Bureau.
2. Table D. Median Income of Households by State: Income 1998. (<http://www.census.gov/hhes/income/income98/in98med.html>). Accessed 3/99. Washington, DC: U.S. Census Bureau.
3. West Virginia Child Nutrition Program: Percentage of Needy Students (By School) October 1999. Office of Child Nutrition, West Virginia Department of Education, Morgantown, WV.
4. Table B-13. Educational Attainment of Persons 25 Years Old and Over, for States: March 1997. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Digest of Education Statistics, 1998.
5. Rye JA, Chester AL. WVU-community partnership that provides science and math enrichment for underrepresented high school students. *Acad Med.* 1999;74:352-5.
6. Piehl DH, Mihm JC. Winning pre-college education programs. *Res Technol Manage.* 1995;38:34-40.
7. Hinman RL. Content and science inquiry: what inquiring minds need to know. *The Science Teacher.* 1998;65:25-7.
8. Peterson N, Jungck J. Problem-posing, problem-solving, and persuasion in biology. *Acad Comput.* 1988;2:14-14, 48-50.
9. Helgeson S. Research in problem solving: middle school. In: Gabel D (ed). *Handbook of Research on Science Teaching and Learning.* New York: MacMillan, 1994, 248-68.
10. National Research Council. *National Science Education Standards.* Washington, DC: National Academy of Sciences, 1996:31.
11. Olson S, Loucks-Horsley S. (eds). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning.* (<http://www.nap.edu/catalog/9596.html>). Accessed 5/00. Washington, DC: National Academy Press, 2000.
12. Rye JA, Bardwell G, Hu J. Connecting science, health, and technology through authentic investigations. *Science Educator.* 1999;8:19-24.
13. Cronin JE. Four misconceptions about authentic learning. *Educational Leadership.* 1993;50:78-81.
14. Gordon R. A curriculum for authentic learning. *Education Digest.* 1998;63:4-9.
15. Magnusson SJ. The learning environment as a site of science education reform. *Theory Into Practice.* 1995;34:43-50.

## The Mount Sinai Humanities and Medicine Program: An Alternative Pathway to Medical School

MARY R. RIFKIN, KENNETH D. SMITH, BARRY D. STIMMEL, ALEX STAGNARO-GREEN,  
and NATHAN G. KASE

In 1984 the AAMC report of the Panel on the General Professional Education of the Physician<sup>1</sup> recommended that students preparing for medical school should strive for a curriculum that provides a broad study in both the sciences and the humanities and that required courses should be kept to a minimum. One way to encourage premedical students to follow a truly broad liberal arts education would be to accept students to medical school early in their college careers, thereby alleviating the pressure to focus excessively on the traditional science-based curriculum. Because there is no evidence to suggest that science majors are necessarily more qualified for medical school, we initiated an experimental program that encouraged humanities and social science majors to pursue their individual interests in college and to obtain a broad, maturing, liberal arts education. Such students might be expected to be less focused on the technology of medicine, bring different perspectives to the practice of medicine, and simultaneously diversify the student body.

In 1989 the Mount Sinai School of Medicine (MSSM) started the Humanities and Medicine (H&M) Program, an early-assurance-of-admission program designed for humanities and social science majors at a targeted group of five liberal arts colleges and universities (Amherst, Brandeis, Princeton, Wesleyan, and Williams).<sup>2</sup> Students in this program are selected during the first semester of their sophomore year in college. Admission into the program is based on a written application with personal essays, verbal and math SAT scores, high school and college transcripts, letters of recommendation, and personal interviews. The students are required to major in the humanities or social sciences and are required to complete only one year of college biology and one year of college chemistry with a grade of B or better.

Admission to MSSM is contingent upon successful completion of undergraduate studies, provided the GPA does not drop below a minimum of 3.0. MCAT scores are not required. In addition, students are required to spend an eight-week summer term at Mount Sinai after their junior year, during which they are exposed to clinical activities and complete a much abbreviated course on the principles of organic chemistry and physics relevant to medicine. Housing and a stipend are provided. Students admitted to the H&M program are under no obligation to attend Mount Sinai should their career choices change or another medical school appear more attractive. Also, the students have the option of deferring their admission to medical school for one year after obtaining the undergraduate degree.

This study reports the outcomes of ten years' experience with the H&M Program. Our experience shows that although students in this program have more academic difficulties in the preclinical years, they excel in the clinical/community setting and have greatly enriched the medical school environment. This program demonstrates that success in medical school does not depend on a traditional premed science curriculum.

### Method

The achievements of all H&M students ( $n = 85$ ) matriculating at MSSM between 1991 and 1997 have been compared with those of two matched cohorts of students who had been accepted through the standard admission process and had completed all standard pre-

med science requirements. Students in each cohort were matched to the H&M students on the basis of year of matriculation, gender, age (within three years), category of educational institution (top 30 liberal arts colleges or universities, taken from the 1998 *US News & World Report* Survey), and, when possible, ethnicity, and were either humanities/social science majors or science majors. The groups of 85 students included students at different stages of their medical school careers and five classes of graduates (1995–1999).

For each group, academic performance in medical school in both basic science courses and clinical clerkships and performance on the USMLE Step 1 examination were analyzed. In addition to these quantitative indicators of performance, we performed an analysis of the students' overall medical school achievements and contributions to the medical school environment in terms of extracurricular activities, student leadership, and service, by evaluating their election to AOA and receipt of special awards. *P* values were determined using the  $\chi^2$  test.

### Results

The undergraduate science/math background of students entering MSSM through the H&M program consists of one year each of biology and chemistry and a short summer course at MSSM, "Physics and Organic Chemistry Relevant to Medicine." This differs from the premed science/math requirements for all other students matriculating at MSSM, namely one year each of biology, chemistry, organic chemistry, physics, and math. The data in Table 1 show that a significantly higher proportion of H&M students had at least one course failure in the basic science years than did the students with traditional premed science backgrounds, who were either humanities majors or science majors. Over 75% of the course failures of H&M students occurred in the first semester of year one, where there were nine failures in biochemistry, six in embryology, six in cell biology, and five in gross anatomy (data not shown). Among the 20 H&M students who failed one or more courses, nine students failed multiple courses, with the range being up to four courses. In the second basic science year, the proportion of H&M students with at least one course failure decreased, with no single course having a disproportionate number of failures.

Compared with their classmates, the H&M students had a higher failure rate on the USMLE Step 1 examination (Table 1), although all these students eventually passed it (data not shown). In an attempt to determine whether failure on the Step 1 examination could be predicted from data available at the time of acceptance into the H&M program, we analyzed the correlation of these students' SAT scores with their performances on the Step 1 examination. Neither Verbal SAT ( $R^2 = 0.08$ ) nor Math SAT ( $R^2 = 0.07$ ) scores correlated with the Step 1 examination score. However, all students who failed the Step 1 examination had Verbal SAT scores  $\leq 650$ .

In the clinical years of medical school, statistically significant differences in performance between the H&M students, when compared with the matched cohorts, were less evident. The failure rate of H&M students in clinical clerkships (Table 1), the garnering of clerkship honors, and election to AOA (Table 2) were not significantly different from those of either matched humanities majors or matched science majors. In fact, the H&M students with mul-

**TABLE 1. Performance of Humanities and Medicine Students Compared with Two Matched Cohorts in Preclinical Courses, in Clinical Clerkships, and on the USMLE Step 1 Examination, Mount Sinai School of Medicine, 1991–1998**

	Humanities and Medicine Students	Matched Regular Premed Students, Humanities Majors	Matched Regular Premed Students, Science Majors	<i>p</i>
<i>Basic science year one</i> : students with at least one course failure	20 (85)*	11 (85)	2 (85)	<.001
<i>Basic science year two</i> : students with at least one course failure	10 (76)	3 (77)	3 (77)	<.03
<i>Clinical clerkships</i> : students with at least one clerkship failure	6 (76)	2 (77)	1 (77)	<.09
<i>USMLE Step 1</i> : students failing on first try	10 (76)	2 (77)	3 (77)	<.02

\* Number in parentheses indicates total number of students analyzed.  
*p* values were determined by chi-squared analysis of the data.

multiple clerkship honors were often the same students who had had academic difficulty in the basic science years or who had failed the Step 1 examination. Analysis of specific clerkships indicated that the H&M students excelled in the psychiatry and pediatrics clerkships (data not shown).

In the preclinical years, Book Awards are given to those students who have performed outstanding extracurricular activity within the community or who have contributed time and energy in service to the institution. Over half the Book Awards were awarded to H&M students (Table 2). H&M students are also disproportionately represented on various subcommittees of the Student Council and other institutional committees, as well as serving in large numbers as student group representatives to national organizations such as the American Medical Student Association, American Medical Women's Association (AMWA), and Students for Equal Opportunity in Medicine (SEOM). Furthermore, a greater proportion of H&M students than students in the two matched cohorts received prizes and awards at graduation (Table 2).

Additional data, not shown, indicate that the H&M students completed medical school at the same rate and did not have a higher attrition rate than students entering medical school with more traditional premed backgrounds. Analysis of residency placements indicated that 77% of the H&M students placed in university hospital-based programs, as opposed to affiliate hospital-based programs, as did 74% of the science majors cohort and 69% of the humanities majors.

## Discussion

The Humanities and Medicine (H&M) Program challenges the long-standing belief that there is a necessary relationship between undergraduate science preparation and the successful completion of medical school and physician excellence. Students in this program are encouraged to use their time in college to pursue in depth their individual interests in their particular majors, which must be in the humanities or social sciences. They often spend considerable time in study abroad, independent research projects in their major fields, or extracurricular activities on campus, such as creative or performing arts or journalism. These students thereby avoid premature specialization and can obtain a broad, maturing, liberal arts education.

The academic performance of H&M students at MSSM has been compared with the performances of two matched cohorts: matriculated students with the standard, required science course background who majored either in the humanities/social sciences or in science. Since the medical school basic science courses are all graded by a norm-referenced rather than criterion-referenced system, and all the other students had had at least two more years of science, including organic chemistry, it is not surprising that the H&M students had more academic difficulties in the preclinical years than did the traditional premed students. However, in the clinical years and in the community setting, the H&M students were similar to the traditional premed students in garnering clerkship honors, institutional awards and prizes, and election to AOA.

**TABLE 2. Numbers of Honors and Awards Given to Humanities and Medicine Students and to Two Matched Cohorts, Mount Sinai School of Medicine, 1991–1999**

	Humanities and Medicine Students	Matched Regular Premed Students, Humanities Majors	Matched Regular Premed Students, Science Majors	<i>p</i> †
Clerkship honors/students				
0 honors grade	7	11	16	.12
1–5 honors grades	57	51	44	.06
6–10 honors grades	12	15	15	.77
Alpha Omega Alpha	14 (76)*	9 (77)	15 (77)	.08
Book awards				
First year	5 (85)	2 (85)	1 (85)	.21
Second year	9 (77)	2 (77)	4 (77)	.06
Graduation awards/prizes, classes of 1995–1999	21 (61)	10 (61)	13 (61)	.03

\* Number in parentheses is total number of students.  
† *p* values were determined by chi-squared analysis of the data.

All the H&M students who failed clinical clerkships ( $n = 6$ ) also had course failures in *both* of the basic science years, whereas none of the students in the cohort groups ( $n = 3$ ) who failed clinical clerkships had course failures in both of the first two years of medical school. While the numbers are small, these data, together with other information about these students' career goals and motivation, suggest that this subset of H&M students may represent students not wholly committed to the study of medicine. There was no evidence in the undergraduate records of these students that could have predicted this pattern of failure.

Although previous reports by others<sup>3,4</sup> indicate that there is no significant correlation between medical school performance and undergraduate major, the students in those studies had completed the required science courses of a traditional premedical undergraduate education. Our report on the performance of the H&M students, who have majored in the humanities or social sciences and who have had minimal science education in college, indicates that, as might be expected, these students have significantly more academic difficulty in the basic science years in medical school than matched classmates who have completed the traditional premedical curriculum. Moreover, we found that all H&M students who failed the USMLE Step 1 exam had verbal SAT scores equal to or less than 650. Thus, in an effort to minimize the number of students whom we might predict would have difficulty in medical school, we have decided to pay particular attention to the verbal SAT score in our admission process, as well as to scrutinize applicants' high school science and mathematics achievements with care.

The premise on which the H&M Program is based is that by eliminating the requirement for traditional premed requirements in college, students have more time to devote to their humanities majors and other pursuits and thus have time to broaden their backgrounds, which would be beneficial to their careers as physicians. These students bring to the medical school certain qualities and outlooks that positively impact the entire medical school community. They have been among the founders of various musical ensembles, theater groups, and art exhibitions, as well as members of the executive board positions of MSSM chapters of AMWA and SEOM. The first woman president of the Student Council was an

H&M student. There is no doubt that the MSSM community has been enriched by the diversity of interests brought to the campus by the H&M students.

The studies reported here should lead us to reconsider the need for the traditional science courses as a prerequisite for success in medical school. Numerous published reports<sup>5-7</sup> have questioned the emphasis on science knowledge in the selection of medical students and have suggested that studies in the humanities may enhance effective patient interaction and communication. By selecting highly qualified, intelligent students early in their college careers and allowing them to develop their curiosity in their chosen fields of interest, as well as involving themselves in community and extracurricular affairs, we have shown that such students successfully complete medical school and excel in clinical activities. We intend to track these students as they complete their residencies and establish their careers to be able to more fully evaluate their contributions.

Correspondence and requests for reprints: Mary R. Rifkin, PhD, Mount Sinai School of Medicine, Box 1475, New York, NY 10029.

---

#### References

1. Muller R. (chair). Physicians for the Twenty-first Century: Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine. Washington, DC: Association of American Medical Colleges, 1984.
2. Stimmel B, Smith K, Kase N. The Humanities and Medicine Program: the need for the traditional premedical requirements. *Acad Med.* 1995;70:438.
3. Ashikawa H, Hojat M, Zeleznik C, Gonnella JS. Reexamination of relationships between students' undergraduate majors, medical school performances, and career plans at Jefferson Medical College. *Acad Med.* 1991;66:458-64.
4. Zeleznik C, Hojat M, Veloski J. Baccalaureate preparation for medical school: does type of degree make a difference. *J Med Educ.* 1983;58:26-33.
5. Dickman RL, Samacki RE, Shimpfhauser FT, Katz LA. Medical students from natural science and nonscience undergraduate backgrounds. *JAMA.* 1980;243:2506-9.
6. Doblin B, Korenman S. The role of natural science in the premedical curriculum. *Acad Med.* 1992;68:539-41.
7. Neame RLB, Powis DA, Bristow T. Should medical students be selected only from recent school-leavers who have studied science? *Med Educ.* 1992;26:433-40.

## ● 1999 JACK MAATSCH MEMORIAL PRESENTATION

The Epistemology of Clinical Reasoning:  
Perspectives from Philosophy, Psychology, and Neuroscience

GEOFFREY R. NORMAN

Physicians' clinical reasoning has been an active area of research for about 30 years. The goal of the inquiry has been to reveal the processes whereby doctors arrive at diagnoses and management plans (although as Elstein correctly points out in his discussion of this paper,<sup>1</sup> the focus has been more on the former than on the latter) so that we could use this information to devise specific instructional strategies or support systems to make the acquisition and application of these skills more efficient and effective. Initially, these "clinical reasoning skills" were conceived as general, and content-independent, so that they could be observed in all clinicians working through any problems. That is, they were thought of as a general mental faculty, presumably rooted in the architecture of the mind, which would be brought to bear on solving clinical problems.

However, the research findings did not support this viewpoint. Elstein and Shulman<sup>2</sup> showed that whatever clinical reasoning was, it was definitely not skill-like, in that there was consistently poor generalization from one problem to another, a finding that ultimately sounded the death knell for evaluation methods such as patient management problems. The past 30 years have seen an accumulation of evidence, in medicine and many other disciplines,<sup>3</sup> about the nature of the process, and shown the importance and centrality of knowledge. The central issue of this revised research program is achieving an understanding of how knowledge is initially learned, how it is organized in memory, and how it is accessed later to solve problems.

A second research program in medical decision making also emerged from research of the early 1970s. As Elstein discusses in the companion paper, this program "views diagnosis making as opinion revision with imperfect information."<sup>1</sup> From the decision-analytic perspective, the best decisions arise from the application of a statistical decision rule to data; any other method is suboptimal. Thus, the research agenda is directed to identifying areas such as medicine where humans function in a suboptimal way, and attempting to understand the strategies, the heuristics and biases, they apply to arrive at these suboptimal decisions.

Elstein states that "it seems to me that decision theory is at least as promising as the study of categorization processes." He may well be correct. But the two schools highlight a fundamental epistemologic dilemma that the remainder of this paper addresses: Will we understand more about the nature of clinical diagnosis by focusing on the diagnostician and striving to understand the mental processes underlying diagnosis, or by focusing on the clinical environment and attempting to understand the statistical associations among features and diseases? To what extent is the world of clinical reasoning "out there" and comprehensible by understanding the relation between symptoms and diseases, and to what extent is it "inside" and understandable only by examining mental processes in detail?

Further dilemmas face us as we examine the research in clinical reasoning. "Organization of knowledge" is viewed as a critical determinant of expertise in medicine. But it is not really clear what is meant by organization of knowledge. Is knowledge organized hierarchically with general concepts at the top, more specific scripts in the middle, and specific instances at the bottom?<sup>4</sup> Is it organized

in networks with nodes and connections,<sup>5</sup> as a symptom-by-disease matrix,<sup>6</sup> as propositions with causal links,<sup>7</sup> as collections of semantic axes,<sup>8</sup> or as individual examples with no overarching concepts, as some of my earlier research claimed?<sup>9</sup>

A perusal of these various studies leaves the reader with only one overall impression—that the human mind is incredibly flexible and can organize and reorganize information at will and seemingly effortlessly to give the researcher exactly what he or she wants to hear. It is no coincidence that propositional networks are disturbingly idiosyncratic and not apparently reproducible.<sup>5</sup> My view is that all of these concept architectures are produced on the fly at retrieval, in order to satisfy the expectations of the researcher, and none can claim special status as *the* way knowledge is organized. Do you want the clinician to tell you the probability that myocardial infarction (MI) will present with referred pain to the back? Can do. The nature of the neural pathways linking the heart and the upper arm? Sure. The hair color of the last patient they saw with an MI? Red. Given this incredible diversity of knowledge from specific to general, it seems likely that any attempt to uncover a representation of knowledge consistent with a particular perspective from fairly directive probes will be successful; however, the ultimate form of this knowledge (if that is even an issue worth addressing) will remain elusive.

Still, if the clinician's mind is really that malleable, then this poses a serious challenge to the research tradition. Are there really any more "basic" or "primitive" forms of knowledge? How can we understand the nature of clinical reasoning if it appears to be this flexible? These were the questions that presented themselves as I reviewed the studies of clinical reasoning. As I thought about these issues, I began to explore other perspectives on the nature of knowledge and knowing from philosophy, psychology, and neuroscience, and started to identify common threads that, I think, can shed some light on these questions. As I did so, I found myself moving back and forth among three kinds of knowing, more or less from specific to general:

1. How does the clinician come to know about diseases? How might diseases be represented in his or her mind?
2. How do we as researchers come to understand domains of science, whether these are the diseases of clinical research or the workings of the clinician's mind?
3. What do we mean by knowing? What do we mean when we say we understand something?

In the remainder of this article I roam freely among these levels, since many of the writings I uncovered inform all levels. But I must begin with a disclaimer. My journeys in this field are as an amateur, and are recent. I have been heavily influenced in my interpretations by two books. The first is *Lessons from an Optical Illusion*, by Hundert,<sup>10</sup> who took the brave step of trying to find links among philosophy, psychology, and neuroscience. His goal was to place ethics in a context of these disciplines; mine is to turn these general truths to an understanding of clinical reasoning. A second major influence on my thinking is a book called *What is this Thing Called Science?*

by Chalmers<sup>11</sup>—a wonderful and readable review of classical philosophy and philosophy of science. I highly recommend both.

The starting point of my discourse is a critical examination of the concept of disease. My intention is to use the exploration of disease as a case study of how we come to know about things.

### What Is a Disease?

Through advances in biology, physiology, and molecular biology, we have come to a deep understanding of the mechanisms of many diseases. It seems almost nonsensical to now turn the clock back and ask what a disease is. But this small departure may serve us in good stead in understanding better what a concept is and how people identify concepts.

Let's take two examples:

- **Is syphilis a disease?** Absolutely. It fits the medical model to perfection. A bacterium invades the host, stimulating a diversity of processes that ultimately are manifested in clinical signs. Osler said "understand syphilis and you understand all of medicine." But there is a small historical glitch. Syphilis has been with mankind for millennia and the signs and symptoms were well established long before the bacterium was isolated.
- **Is heart disease a disease?** Yes. Put a label such as anterior myocardial infarction on it, and it looks even more like a disease. But likely we are all harboring the precursors of ischemic disease as cholesterol plaques slowly accrue in our arteries. So in a manner of speaking, the prevalence of heart disease approaches 100%. Can we then still speak of it as a disease? And by the way, although there are many risk factors for heart disease, there is no clear cause. The same is true for cancer. We can easily identify cancerous cells on pathology slides, and we can correlate the clinical course with the accumulation of malignant lesions, but we all have microscopic tumors in our thyroids, and a third of men who die of unrelated causes are found to have prostate cancer.

All of these things seem disease-like because we can "explain" them at some lower level—plaques, bacteria, malignant cells. But there are many other diseases listed in textbooks that have no clear causes, no microscopic correlates, no known mechanisms. And it is well to bear in mind that although anthropologists and historians have identified evidence of (for example) tuberculosis dating back several thousands of years, and although old writings in medicine clearly describe the symptoms and clinical course of tuberculosis, the cause, the tubercle bacillus, was identified, by Koch, only as recently as 1884, and effective therapy has been available only since the 1940s. So the existence of a causal mechanism is hardly sufficient to claim that something is a disease. More generally, it is likely that exceptions to any definition of disease will be common.

Campbell et al., in a classic article, "The Concept of Disease," reported presenting clinicians and lay people with a series of medical conditions and asking them whether or not they were diseases.<sup>12</sup> Perhaps not surprisingly, doctors were more prone than lay people to call things such as lead poisoning and tennis elbow diseases. But there was otherwise quite good concordance. Infectious diseases—malaria, tuberculosis, syphilis, polio—topped the list. Other common or serious medical problems—lung cancer, diabetes, multiple sclerosis, cirrhosis—came next. At the bottom were things such as hangover, senility, heatstroke, tennis elbow, and drowning, which had English, not Latin, labels. These authors concluded that the features that best predicted the labeling of a condition as a disease were that the condition (1) was associated with an abnormality of structure or function (i.e., it had a "cause") and (2) was likely to be treated by a doctor. The latter was the stronger determinant, but regrettably, this seems tautological. Since doctors are in the business of dealing with disease, describing a disease as some-

thing that doctors deal with does not, in my view, advance our understanding much.

Let us consider the first predictor for a moment. Arguably one simplistic but functional view is that if a condition simply represents a cluster of signs and symptoms (for example, carpal tunnel syndrome, low back pain) it is less disease-like. Presumably this reflects a concern that a condition's features and associations among the features may be an illusory correlation (which humans are particularly good at making)<sup>13</sup> and not "real." There is good reason for such a degree of skepticism. Historically, many syndromes that existed 100 years ago, such as self-pollution, have now disappeared, and there is every indication that many contemporary syndromes, such as chronic fatigue, sick-building syndrome, Gulf War syndrome, and the myriad health problems believed to be caused by breast implants may go the same way. Conversely, the ability to explain disease through some underlying mechanism lends authenticity to it. Angina becomes much more believable if we can find narrowing of the lumen of the coronary artery on angiography, even though the association with the clinical manifestations is weak.

### The Role of Basic Science

If we view the identification of the features of a disease as analogous to the findings of an experiment (in this case, an experiment conducted by a malicious deity) then one basis for distinguishing a disease from a non-disease is the extent to which the features can be explained by a scientific theory. Thus the infectious diseases are explained by a noncontroversial, and historically verified, theory of host and parasite. Chronic diseases such as atherosclerosis are a bit less disease-like since the theory underlying them is less secure. And as we move to syndromes such as chronic fatigue syndrome, we are less inclined to view them as diseases because no satisfactory scientific mechanism has yet been found to explain their features.

Turning to clinical reasoning, investigators such as Schmidt<sup>14</sup> and Patel,<sup>15</sup> in studying the role of basic science in clinical reasoning, have found repeatedly that clinicians rarely invoke mechanistic explanations. But as Schmidt has shown, the fact that they need not invoke mechanisms does not mean that they do not know them—the knowledge is available but is only rarely used. As he describes it, the knowledge is "encapsulated." While basic science may play only a minimal role in day-to-day-practice, it is arguably the only, or at least the major, route to understanding in this domain. Of course, basic science need not be restricted to biology. In the same way, the basic science of epidemiology was fundamental to understanding the transmission of AIDS, just as Snow in the 1880s understood the mechanism of cholera transmission (the London water supply) long before the bacillus was isolated.

I believe we can now posit an explanation for the paradoxical findings of Schmidt and Patel. In the normal course of events, clinicians making diagnoses deal at the syndrome level, where the nature of the causal mechanism is irrelevant. The history and physical exam are directed at revealing the syndrome-like manifestations, which then point to tests directed at the underlying processes, and therapy. The textbooks of clinical diagnosis for "old" diseases probably have not changed much since Osler's time. The signs and symptoms are pretty well what they have always been, although of course some historic scourges—smallpox, diphtheria, cholera—are now nearly unheard of in the West, and others, such as AIDS, have taken their place. But despite the changes in our understanding of disease, the clinician attempting to make a diagnosis is dealing almost exclusively at the syndrome level. Occasionally, some understanding of underlying processes may help to sort out some conundrum, but one suspects that clinicians appear rarely to use basic science simply because their investigations of history and physical are directed to labeling the syndrome. Clinical reasoning reverts to a historically earlier form of the disease, following the biologic dictum that ontogeny follows phylogeny—the fetus passes through all stages of evolution before birth.

Campbell<sup>13</sup> elaborated the notion of disease in philosophical terms, describing two basic positions: the “nominalist” perspective and the “essentialist” perspective. In the nominalist view, a disease is simply a collection of abnormalities that appear to arise together. Thus the historical diseases of dropsy, consumption, and plague were recognized long before any causal agent was detected, although etiologies (such as “bad humors”) were advanced. Conversely, the essentialist perspective presumes that the signs and symptoms arise from pathologic processes that can be identified and hopefully rectified. While it is tempting to place these two views in a historical order, the contemporary examples we have discussed indicate that the two perspectives represent extremes on a continuum, which, as we shall see, has parallels in both philosophy and psychology.

### What is a Concept? Lessons from Philosophy

We can make some general observations about the concept of disease. First, a disease, like any concept, does not exist entirely “out there” but rather, to some degree, is a mental construct. Second, the category or concept called “disease” is not an all-or-none proposition; rather, particular exemplars have different degrees of disease-ness. Finally, it is awfully difficult to devise an explicit rule to aid in distinguishing between diseases and non-diseases. A rule such as “diseases are what doctors deal with” works quite well but is singularly uninformative. And we sense, without proof, that any rule we may devise is not going to be coldly analytic, but must have sub-rules such as “the more Latinesque it is, the more disease-like it is.” So ironically, while it is relatively easy to devise rules to determine whether someone has a particular disease (although I will go on to show that the rules are not the whole story), it is a lot harder to devise rules for the overarching category called “disease.”

These issues are not at all specific to disease, but rather are part of a large body of knowledge extending in space across at least three disciplines—philosophy, psychology, and neuroscience—and in time as far back as Plato. To explore this further, I now venture (with considerable trepidation) into a more general inquiry into the nature of concepts. I begin by revisiting some philosophical views on the nature of concepts.

The origin of concepts has been, in some sense, a nature–nurture debate.<sup>9</sup> However, this argument has focused not on whether human traits are inherited or learned (the usual spin on nature versus nurture), but rather on whether categories or concepts such as beauty, disease, table, or tree exist “out there” to be learned by individuals as they develop and mature (which would suggest that an individual’s knowledge is formed from experience [nurture]) or are essentially a product of the mind (we impose order and category boundaries where none exists, as a result of the biological structure of the mind [nature]). A casual reading of any philosophy textbook reveals that this issue has been a central concern through the ages of the great minds—Plato, Aristotle, Descartes, Hume, Kant, etc. Let us briefly review the historical debate in mainstream philosophy, with a view to showing how thinking in philosophy can help to frame our perspective on clinical reasoning.

Modern philosophy began with Descartes, who emerges as the ultimate skeptic, and whose views have retained central status as the universal straw man for all his successors. His famous statement “cogito, ergo sum” (I think, therefore I am) has been a lodestone for philosophers and t-shirt makers for three centuries. Regrettably, this idea has been almost universally misunderstood. Most interpret it as a statement of the ultimate rational man; our humanity is defined in terms of our capacity for rational thought. Unfortunately, the statement had a much more humble meaning for Descartes. In continuing to question whether one could justify any external reality, to devise any conclusive argument for the existence of objects such as dogs and tables, Descartes was led to the desperate conclu-

sion that the only thing he could be really sure of was his own thoughts. I think, therefore I am.

The antithesis of this position was championed by the English empiricists Locke and Hume. Their view was that the mind was a *tabula rasa*, a clean slate on which one’s experience with the world was written. This interpretation seems perfectly acceptable for sensory experience, but is more difficult to sustain for higher concepts such as causation, temporality, or, for that matter, disease. Hume’s resolution was to suggest that these notions emerged as a result of experience.

Kant reframed the issue in a way that is central to our subsequent journey through psychology and neuroscience. He recognized that thoughts can occur only as products of interactions between the mind and the external reality of experience; we construct experience. He maintained a rigid boundary between those properties that our minds bring to experience (which are hardwired) and those that emerge from experience. He eventually created a list of 12 “primitives”—object, causation, temporality, and nine others—that he claimed the mind imposed on the world of experience.

Hegel went one step further and recognized that the external world can influence the categories and labels we apply. The categories themselves do not emerge from our minds, but are influenced by the objects of our perceptions. The mind is not simply a clean slate upon which all experience is written in coherent form (Hume); nor is it the case that there is no uniform order in the outside world and that all concepts are mental inventions (Descartes); nor finally does the mind impose fixed structure or constructs on sensory experience (Kant). Instead, the concepts and the content both grow and evolve (“become”) as a consequence of the interaction between the individual and the environment.

Finally, in this century, Wittgenstein extended these ideas further. He proposed that not only are concepts not fixed, they also are not definable by any set of logical rules. In pondering even commonplace concepts such as “dog,” he realized that any attempt to devise rules is doomed. A dog has four legs—but if one is amputated it’s still a dog. A dog barks—except an Egyptian Basenji. A concept—whether an abstract concept such as truth or a mundane concept such as dog, fork, or tree—emerges as a matter of “family resemblance.” Robins are more bird-like than penguins; malaria is more disease-like than alcoholism. Wittgenstein proposed that concepts or categories are derived from family resemblances, not from fixed sets of defining attributes.

Thus the philosophy of concepts evolved from a Cartesian view, which is entirely intra-psychic and questions any external reality, and an empiricist perspective that presumes that all order and concepts exist as natural categories to be discovered by the human observer, to a Kantian interaction, in which the mind provides the categories or concepts and the external reality provides the objects to fill the categories, to a Hegelian perspective, which is much more organic, and in which thoughts and concepts themselves evolve and change as a result of interactions with external reality. Ultimately, we reach the perspective of Wittgenstein, which places even fewer constraints on concepts, which are a matter of family resemblance and thus can be elaborated only through extensive experience with the world’s families.

Applying these notions to clinical reasoning, philosophy presents a larger framework in which to view our dilemma in defining a disease. To the extent that a disease is a concept, philosophy buttresses the middle ground between the notion that diseases exist entirely “out there” only to be discovered and learned and the notion that they are probably simply mental constructs. We can then think of the concept of disease as arising from an interaction between the thoughts of the perceiver and regular aspects and associations of the environment. Further, some diseases, such as syphilis, are more central members of the family; others, including the syndromes, are more peripheral.

As we shall see, this formulation finds remarkable support in research in both psychology and neuroscience, to which I now turn.

## What is a Concept? Lessons from Psychology

One division in psychology has been preoccupied with the same issue as the philosophers: how do people learn concepts such as table, dog, or truth? But instead of relying entirely on reason for understanding, psychology seeks evidence to understand how people create and learn concepts. Perhaps in the course of doing so, psychologists deliberately skirt some of the tough epistemologic issues that preoccupy philosophers. On the other hand, in my own reading, I was struck by how the one informs the other. A simple example:

The Müller Lyer illusion,<sup>16</sup> shown in Figure 1, is pretty well known to all. We see the one vertical element as being longer than the other. Even though we can measure them and show them to be the same, the illusion is inescapable—a fine example of how we impose order (sometimes biased order) on the external world. But psychologists have gone further with this illusion, and questioned precisely *why* it is an illusion. In the course of doing so, they provide a nice illustration of Hegel's interactive model of mind. One hypothesis is that it is an illusion because our minds are seeing it in three dimensions, so that the symbol on the left is seen as the outside corner of a wall nearest the viewer, and the one on the right is seen as the inside corner of a wall farthest away from the viewer. Although the two vertical lines are objectively the same size, since the one on the left is seen to be nearer than the one on the right, the right one is "actually" longer. Deregowsky<sup>17</sup> tested the illusion in Zulus, who spend their lives in round houses, and found that they did not see it as an illusion. So, it is not an illusion because our brains are "hardwired" to see it as such (unless Zulus have different hardwiring); it is an illusion because of the particular experiences we have had with the world. On the other hand, the illusion reminds us that our perceptions do not necessarily mirror reality, as they are also shaped by internal assumptions (in this case, about perspective and the inference of a third dimension from the two-dimensional representations on the retina) that sometimes lead us astray.

A second example from psychology leads us closer to our central concern with clinical reasoning. Most of us have, at one time or another, wondered whether the "red" we see is the same as the red seen by the person beside us. While the differences in perception are rarely likely to be as extreme as in the case of a childhood friend of mine whose color blindness was detected when he went to school and repeatedly drew green reindeer at Christmas, we have no real way of ever verifying the universality of "red." Is it just a linguistic device, or a cultural norm? After all, at some time we all had to learn, from our parents or friends, what red was. Perhaps it differs in different cultures. These questions, as they begin to cross the boundary between philosophy, psychology, and learning, are of more than passing interest.

Much of the fundamental work in concept formation has been done by Eleanor Rosch.<sup>18</sup> One area she studied was how colors are identified in different cultures. While, on the one hand, there appear to be small cultural differences in the boundaries between colors (e.g., the Navaho have only one word for blue and green (no wonder, with all that turquoise jewelry around),<sup>10</sup> Rosch showed that all cultures were unanimous in their choices of the best examples of red, yellow, or green. Even more interesting, Rosch discovered a primitive tribe, the Dani, who had words for "bright" and black only. She then taught them words for colors, using Dani words (e.g., tree) that were unrelated to color. One group learned the "primary" colors such as fire-engine red; the other learned Dani words for intermediate colors such as turquoise. The group learning red, yellow, and blue learned the associative words rapidly and effectively; the other group never did master the associations. Studies of this type provide support for the contemporary notion in philosophy that categories and concepts derive from our experience of the world; indeed there is surprising uniformity to these concepts

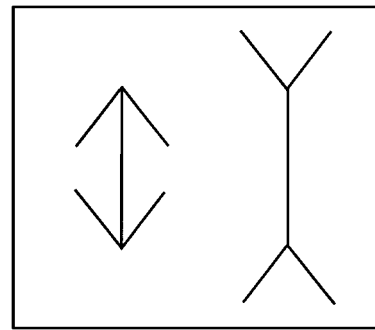


Figure 1. Müller Lyer optical illusion.

in precisely those areas where we might expect that experience (such as the experience of color) is also universal.

Prototype theory was perhaps the first theory of concepts to be seriously applied to clinical reasoning. Bordage and Zacks<sup>19</sup> used many of the methods of Rosch to demonstrate that the same kind of graded structure that distinguished the natural categories was present in disease categories. They found, for example, that diabetes was a much more prototypical endocrine disease than Hashimoto's disease or hyperthyroidism. It was volunteered more often by practitioners asked to name as endocrine disease, recognized more accurately and quickly, and so on.

These studies lead to two conclusions: first, there is evidence to substantiate our musings at the beginning of this talk that the concept of disease is a continuum, not a category. Second, the identification of conceptual prototypes such as diabetes, carrot, and robin, which transcend different cultures, argues for an external "nurture" basis for concepts—even high-level concepts such as disease.

Prototype theory, in its methods, seeks evidence for cultural or even transcultural norms for categories. In the extreme, prototype theory might be viewed as empirical evidence for a position that concepts and categories are derived entirely from universals in the environment, a position more extremely nurture-oriented than any we have considered except the positions of Locke and Hume.

Another psychological theory of concept formation, exemplar theory, while still holding to the implicit view that the concepts we learn reflect an external reality, is much more modest about the universality of such concepts. In this perspective, we are able to identify a member of a class or a concept, not because of any internal rules or because the sum of our experience has created prototypes of the class that are available for analysis and introspection, but because we have, in any category (dogs, chairs, diseases, sports cars), an innumerable number of instances of the category (my dog, Rover, Lassie, etc.). When we are faced with a categorization task, a first line of defense is a search through memory for similar examples of the class, and then, if we find an example that is sufficiently similar, we assume the new beast is also a dog. This description makes the process sound far more deliberate and available for introspection than the evidence suggests. Instead, if we inquire why a person decided that the new beast was a golden retriever, the new car was an Audi, or the skin lesion was actinic keratosis, the modal response would be "Because it looks like a golden retriever," or an Audi or actinic keratosis. Further justification may be forthcoming but it sounds suspiciously post hoc. This process is in fact unlikely to be available for conscious introspection.

I and some colleagues have done a series of studies in dermatology<sup>20</sup> and cardiology<sup>21</sup> in which we have found evidence for this mode of processing. As one example,<sup>20</sup> in a series of experiments we gave subjects (residents) practice with a set of dermatology slides covering 11 conditions, then subsequently tested them with a new set of slides. The slides were carefully chosen. Each was drawn from

a quartet of slides containing two typical slides that strongly resembled each other, and two atypical slides that resembled each other. Each subject was then tested with two other slides of the quartet. We balanced it all off, so that we could look at performances on typical-similar, atypical-similar, typical-different, and atypical-different slides. Thus we deliberately compared typicality (a property of the number of features and prototype theory) with similarity (a characteristic of exemplar-based reasoning). The results showed effects of both similarity and typicality. With immediate testing, similarity resulted in a gain of accuracy of about 50%, typicality a gain of about 12%. After ten days' delay, slides that were similar to those in the initial learning series were diagnosed about 25% more accurately, and typical slides were diagnosed about 25% more accurately.

We have continued to explore these phenomena. One concern is that it will work only with visually rich materials, where similarity is highly perceptual. Hatala<sup>21</sup> conducted a study with ECG interpretation, which, while still visual, is replete with quantitative rules. In this study, similarity to an ECG in the learning phase was based entirely on a one-line description (e.g., a "54-year-old accountant" and a "middle-aged banker" versus "an "80-year-old widow"). To demonstrate the effect, the match was to an ECG that was visually similar, but from an incorrect and confusable category (e.g., left bundle-branch block and anterior MI). When the description was matched, accuracy was 23%; when it was unmatched, it was 46%, and of course more residents who saw the matching description fell for the incorrect diagnosis. Further, it would seem that the process must have occurred without awareness. If they had known they were matching on the age and occupation, they would not have done it, since a moment's introspection reveals that this is irrational.

Both of these psychological theories—prototypes and instances—derive from a nurture view of concepts, namely that the concepts we learn are derived from our experiences. In fact, the exemplar models show precisely how specific experiences are available and used in subsequent judgments of category membership. However, as always, there is another side to the story. Psychology has been equally successful at deriving evidence to support the nature view, that what we see is influenced by our own minds. Admittedly, this is not a pure nature view, as we shall see, since the way our perceptions of the external world are biased derives itself from our experience with the world.

Cognitive psychology had its origins in an information-processing model based on the metaphor that the mind is like a computer. However, there was rapid accumulation of evidence showing just how un-computer-like humans are. One simple yet fundamental example is in information retrieval. The answers to questions such as "When did Columbus discover America?" and "What is the capital of Arkansas?" are available almost as soon as you hear the question inflection. Second, if asked about Albania, not Arkansas, you would know that you didn't know almost as rapidly. Contrast that with a search of the Web. Although the computer processes information at least a million times faster than does the mind, retrieval will inevitably take much longer. Further, it will take the computer longer still to decide that it doesn't know, since it will have to search every corner of its memory before it gives up. It is difficult to envision what kind of memory architecture humans must have to do this job, but it must be very different from the computer's RAM.

One model of memory that accommodates these observations is called human associative memory. The model emerged from studies of reading coupled with a phenomenon called the word-superiority effect,<sup>22</sup> which has relevance, surprisingly, to clinical reasoning as well as to many other domains. Imagine that I flash a four-letter word on the computer screen for a few milliseconds and ask you to identify the fourth letter. The phenomenon is this: when the fourth letter occurs in a real word such as "rink" or a pseudo-word such as "bink," the "k" is recognized faster and more accurately than

when it occurs in a non-word such as "nrik." While this seems perfectly plausible, it says some fundamental things about the nature of memory. That is, even at the perceptual level of recognizing individual letters, a process that must occur in milliseconds and without conscious introspection, identification is facilitated by memory of much higher-level concepts, the words themselves. This seems to illustrate beautifully the interactive nature of perception, showing that what we see can be influenced by what we expect to see.

The observations of the word-superiority effect were modelled by McLelland and Rumelhart<sup>23</sup> using a "connectionist" or parallel distributed processing (PDP) model, with multiple layers of nodes between input and output corresponding to letter elements, letters, and words, with links among nodes at all layers. Unlike expert systems or Bayesian models, these connectionist models had no pre-programmed rules: rather, they "learned" from experience, gradually building up strength among certain links connecting nodes.

Parallel distributed processing models have been continually refined (and renamed—they are now more commonly known as "neural networks"), and have found application in many settings, including clinical diagnosis, where they appear to be more effective diagnosis machines than the traditional expert systems. However, for present purposes, these applications are less important than the observation that the models have commonality with psychological views of concept formation, based on learning from examples. And as we shall see, the new name is not simply good public relations—neural networks bear a striking resemblance to models emerging from neuroscience.

I and my colleagues have taken the phenomenon that recognition depends, in part, on available concepts in memory into the clinical reasoning lab. In a series of studies in dermatology, radiology, and electrocardiography, we biased the subjects by providing a brief history suggestive of a particular diagnosis, then showed them a visual stimulus—an ECG, a slide of a skin lesion, or a head-and-shoulders picture. We have consistently found that the bias influences not only the differential diagnosis (which might be viewed as perfectly rational), but also the feature calls. Moreover, in a recent study using textbook examples of physical signs,<sup>24</sup> we showed that it was not simply a case that the history increased vigilance for that particular sign, and therefore the likelihood of detection. Rather, an incorrect history led students to misinterpret one sign as another—the inflamed parotid glands of mumps became the moon-shaped face of Cushing's disease, and the moon-shaped face of Cushing's became periorbital edema when linked with a history of nephrotic syndrome.

This phenomenon, that prior higher-level information either provided to the subject or available from memory can influence basic perceptual processes, has been demonstrated at all levels of expertise, from first-year students to cardiologists, so it is not simply a naive bias that can be erased with experience. LeBlanc's follow-up studies of strategies to "de-bias" subjects, under way in our lab, have shown that even fairly draconian measures are only partially successful; a finding that is not surprising since perceptual processes are not available to conscious introspection.

These findings, both in cognition of perception and in clinical reasoning, challenge a commonly held view that experts use "forward reasoning"; that is; they begin with the facts of the case and reason inductively to a logical conclusion, a view championed by Groen and Patel.<sup>25</sup> Their findings were derived from verbal introspections or written summaries, after the subjects had had time to read and reflect on the clinical case. It is my present view that the work on top-down processing, both in reading and reasoning, shows that deductive processes from hypothesized solutions are already occurring long before the case is in full view, and that the apparent induction of the expert simply reflects a coherent story told post hoc. One study done by Eva<sup>26</sup> substantiates this view. He had subjects read mystery stories, then recount their solutions. Half told their solutions "online" as they were reading; the other half,

as a summary after. On three measures, the latter group looked as if they were doing substantially more forward reasoning. However, the manipulation took place *after* the reasoning was over.

### What is a Concept? Lessons from Neuroscience

Finally, conspicuous in its absence from the discussion to date is the role of neuroscience in our understanding of concepts. I have described how cognitive psychology has provided examples of phenomena that help us to understand some aspects of clinical reasoning. Theories of concept formation and perception are a useful heuristic for testing apart aspects of clinical reasoning. But the skeptical reader could be forgiven for remarking that these theories seem more like useful demonstrations and analogies than real explanations, in a scientific sense.

Let me then venture into what is for me the largely uncharted territory of neuroscience. In doing so, I am moving closer to the more traditional interpretation of the nature–nurture debate than the way I originally framed it. That is, we now seek evidence from neuroscience that the brain and its structures (nature) are responsive to, and modified by, the environment (nurture). Further, just as basic science provides a framework for understanding disease, neuroscience may provide a framework for understanding the process of concept formation and clinical reasoning.

To advance the neuroscience argument, we need to discover evidence that categories “out there” can be localized to specific brain activities. Perhaps the most accessible argument about the impact of specific experiences on brain anatomy and brain development emerge from the phenomenon of plasticity—the discovery that there are critical periods in the development of the brain during which input from the environment is required in order for specific facilities to develop. The phenomenon is ubiquitous. Here are some examples:

- Children who have congenital cataracts must have them surgically removed before age 10, or they will be unable to recognize shape and pattern, although they will be able to learn colors. This was hypothesized to arise because of abnormal development in the visual cortex. Very recent research with newborns has extended this understanding further. Maurer studied children less than 9 months old who had had cataracts removed immediately following the surgery. Immediately after surgery, their vision was like a newborn’s—about 1/40 the acuity of an adult’s. But after only **one hour** of visual input, their acuity had improved to the level of a one-month infant. To quote the researcher: “It’s using the eyes and having the experience of seeing that’s driving the normal experience of vision after birth. . . . the brain was wired to be ready to receive visual images . . . but it’s got to have the input in order to do the learning.”<sup>27</sup>
- Animal experiments showed kittens raised in an environment that only allowed horizontal or vertical orientations never learned the other. Hubel and Wiesel<sup>28</sup> then showed that these selective deprivations are identifiable in the development of specific cells in the visual cortex. They went on to show that the brain development was incredibly specific, so that a single day of exposure at day 28 was sufficient to establish the orientation. Other researchers have gone on to establish that plasticity is associated with the presence of specific proteins.

The phenomenon of plasticity is direct evidence of an interaction between brain structures and the environment, and provides an explanation for the philosophical dilemma. Of course, such experiments do not provide direct evidence that higher-order concepts such as temperature, unemployment, love, or for that matter, tables, are associated with specific local changes. The next step is to move from the construction of perceptual maps of the environment to conceptual maps in different areas of the brain. This may not be as large a leap as it sounds; after all, the mechanism that

enables us to recognize Aunt Sally must involve links among the more primitive operators that isolate color, shape, and orientation. Thus, we move from brain mappings corresponding to perceptual inputs, which, as we have seen, develop and specialize as a consequence of interactions with the environment at highly specific developmental intervals, to mappings corresponding to the relations among these elements—a “mapping of types of maps,” according to Edelman.<sup>29</sup> This remains a theory thus far, the theory of “neuronal group selection.” I cannot pretend to be more than an intrigued observer, but it would seem that the evidence at hand regarding neural plasticity provides plausible mechanisms for such a neural correlate of concept formation. Indeed, as I discussed earlier, although neural networks were devised as a simulation device to test a model of concept learning involving parallel and distributed activation, there is a striking correspondence between the nodes and connections of neural networks and the proposed model of neuronal group selection.

### Conclusions

This review was intended to accomplish no more than to place the current debates around clinical reasoning in a larger context. There is, in all this, a Michigan State University (MSU) connection. The small research program focusing on clinical reasoning was begun by Elstein and Shulman at MSU in the early 1970s. The McMaster group joined the fray soon after, with me as their hired hand. But soon after this first cycle of studies was completed, there was a strong divergence in the field. Elstein moved his interest to normative approaches such as decision analysis, assuming that clinicians were suboptimal decision makers who could be made more optimal with training. Others who followed, including Patel and Groen, while disagreeing on the details, retained a strongly rationalist perspective. On the other side, Bordage pursued studies in prototype theory, and I began a research program around exemplar models. It is only recently, with the study leading to this review, that I began to appreciate the historical origins of this divergence.

The exciting conclusion from this review is that there appears to be a convergence among the three disciplines—philosophy, psychology and neuroscience—pointing to the reconciliation of these positions. While the constructs, the capacities for identifying regularities appear innate, these abilities are directly responsive to the environment, so that each individual’s concepts will be both communal and idiosyncratic. Moreover, this synthesis has some practical implications (believe it or not!). It appears to me that these thinkers are urging us to a reconciliation in our own field—expertise in clinical reasoning is neither mastery of analytical rules nor accumulation of experience, it is both. And the role of experience with individual examples in refining the concepts is critical. Moreover, the philosophical work and the demonstrations of optical illusions show us that the external environment is not delivered to the senses intact, but is filtered through the prisms of prior experience. These are important lessons for instruction in clinical reasoning.

The sum of these findings describes a model of clinical reasoning very different from the algorithmic processes used by the computer (except when, using neural networks, the computer mirrors the mind). An evident implication is that there is little to be gained in demonstrating that humans are suboptimal Bayesians or algorithm-appliers; they are suboptimal because they are using a substantially different basis for computation. While, on the one hand, this provides a strong rationale for computerized decision-support systems, the cautionary note that pervades this review is that the support system cannot intervene after the data are collected, since the data are themselves subject to interpretation in light of mental models.

The Jack Maatsch Memorial Presentation was sponsored by the Office of Medical Education Research and Development, Michigan State University, and presented at the annual AAMC–RIME meeting, October 27, 1999.

Correspondence: Dr. Geoffrey Norman, Department of Clinical Epidemiology and Biostatistics, Health Sciences Centre 2C14, McMaster University, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada; e-mail: {norman@mcmaster.ca}.

---

References

1. Elstein AS. Clinical problem solving and decision psychology: comment on "the epistemology of clinical reasoning." *Acad Med.* 2000;75(10 suppl):S134–S136.
2. Shulman LS, Elstein AS, Sprafka SA. *Medical Problem Solving: An Analysis of Clinical Reasoning.* Cambridge, MA: Harvard University Press, 1978.
3. Ericsson KA, Smith J. *Toward a General Theory of Expertise: Prospects and Limits.* Cambridge, U.K.: Cambridge University Press, 1991.
4. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. *Acad Med.* 1990;65:611–21.
5. McGaghie WC, Boerger RL, McCrimmon DR, Ravitch MM. Learning pulmonary physiology: comparison of student and faculty knowledge structures. *Acad Med.* 1996;71:S13–S15.
6. Papa FJ, Elieson B. Diagnostic accuracy as a function of case prototypicality. *Acad Med.* 1993;68(10 suppl):S58–S60.
7. Patel VL, Groen GJ, Frederiksen CH. Differences between medical students and doctors in memory for clinical cases. *Med Educ.* 1986;20:3–9.
8. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med.* 1994;69:883–5.
9. Norman GR, Brooks LR, Allen SW. Role of specific similarity in a medical diagnostic task. *J Exp Psychol Gen.* 1991;120:278–87.
10. Hundert EM. *Lessons from an Optical Illusion: On Nature and Nurture, Knowledge and Values.* Cambridge, MA: Harvard University Press, 1995.
11. Chalmers AE. *What is This Thing Called Science?* 2nd ed. St. Lucia, Australia: University of Queensland Press, 1999.
12. Campbell EJ, Scadding JG, Roberts RS. The concept of disease. *BMJ.* 1979;6193:757–62.
13. Chapman LJ, Chapman JP. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *J Abnorm Psychol.* 1969;74:271–80.
14. Boshuizen HPA, Schmidt HG. The role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cogn Sci.* 1992;16:153–84.
15. Patel VL, Groen GJ, Scott HM. Biomedical knowledge in explanations of clinical problems by medical students. *Med Educ.* 1988;22:398–406.
16. Gregory RL. *Eye and Brain: The Psychology of Seeing.* London, U.K.: Weidenfield and Nicolson, 1966.
17. Derogovsky JB. Illusion and culture. In: Gregory RL, Gombrich EH (eds). *Illusion in Nature and Art.* New York: Scribner, 1974.
18. Rosch E. Natural categories. *Cogn Psychol.* 1973;4:328–50.
19. Bordage G, Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. *Med Educ.* 1984;18:406–16.
20. Regehr G, Cline J, Norman GR, Brooks L. Effect of processing strategy on diagnostic skill in dermatology. *Acad Med.* 1994;69(10 suppl):S34–S36.
21. Hatala R, Norman GR, Brooks LR. Influence of a single example upon subsequent electrocardiogram interpretation. *Teach Learn Med.* 1999;11:110–7.
22. Reicher GM. Perceptual recognition as a function of meaningfulness of stimulus materials. *J Exp Psychol.* 1969;81:274–80.
23. McLelland JL, Rumelhart DE. An interactive activation model of context effect in letter perception. *Psychol Rev.* 1989;88:375–407.
24. Brooks LR, LeBlanc VR, Normal GR. On the difficulty of noticing obvious features in patient appearance. *Psychol Sci.* 2000;11:112–7.
25. Patel VL, Groen GJ. Knowledge based solution strategies in medical reasoning. *Cogn Sci.* 1986;10:91–116.
26. Eva KW, Norman GR. Is thinking aloud equivalent to post hoc explaining. Presented at the 1998 Research in Medical Education meeting, New Orleans, LA, 1998.
27. Maurer D, Lewis TL, Brent HP, Levin AV. Rapid improvement in the acuity of infants after visual input. *Science.* 1999;286:108–10.
28. Weisel TN. The postnatal development of the visual cortex and the influence of the environment. *Bioscience Reports.* 1982;2:351–77.
29. Edelman GM, Mountcastle VB. *The Mindful Brain.* Cambridge, MA: MIT Press, 1978.