REFERENCES

Damasio AR (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY, USA: Harcourt Brace

Dennett D (1991) *Consciousness Explained.* Boston, MA, USA: Little, Brown & Co.

Farah MJ, Illes J, Cook-Deegan R, Gardner H, Kandel E, King P, Parens E, Sahakian B, Wolpe PR (2004) Neurocognitive enhancement: what can we do and what should we do? *Nat Rev Neurosci* **5:** 421–425

Hall W (2004) Feeling 'better than well'. *EMBO Rep* **5:** 1105–1109

Koch C (2004) *The Quest for Consciousness: A Neurobiological Approach.* Denver, CO, USA: Roberts & Co.

Kureishi H (2002) *The Body: and Seven Stories.* London, UK: Faber & Faber

Metzinger T (2003) Der Begriff einer Bewusstseinskultur. In Kaiser G (ed), *Year Book 2002/2003 of the Wissenschaftszentrums Nordrhein-Westfalen* pp 150–171. Düsseldorf, Germany: Wissenschaftszentrum Nordrhein-Westfalen

Pinker S (1997) *How the Mind Works.* New York, NY, USA: Norton

Rose H (2004) Consciousness and the limits of neurobiology. In Rees D, Rose S (eds), *The New Brain Sciences: Perils and Prospects* pp 59–70. Cambridge, UK: Cambridge University Press

Rose S (2002) 'Smart drugs': Do they work? Are they ethical? Will they be legal? *Nat Rev Neurosci* **3:** 975–979

Searle JR (1995) The mystery of consciousness. *NY Rev Books* Nov 2. www.nybooks.com

Searle JR (2005) Consciousness: what we still don't know. *NY Rev Books* Jan 13 www.nybooks.com

Wolpe PR (2002) Treatment, enhancement, and the ethics of neurotherapeutics. *Brain Cogn* **50:** 387–395

Giovanni Frazzetto

doi:10.1038/sj.embor.7400384

The consequence of errors

From memory molecules to the criminal chromosome, erroneous conclusions continue to blight scientific research

Ancient Greek philosophers laid the groundwork for the scientific tradition of critical inquiry, but they nevertheless missed out on one aspect important to modern science. Many philosophers obtained their results through a tradition of contemplation and thought rather than experimental procedure, which, not surprisingly, led to errors. Aristotle's belief that the brain is a cooling organ for the blood was definitely not based on anything that scientists today would consider scientific evidence. He also thought that in humans, goats and pigs, males have more teeth than females, a notion easy enough to correct. His statement that flies have four legs was repeated in natural history texts for more than a thousand years despite the fact that a little counting would have proven otherwise.

Today, these errors are anecdotal, and science prides itself on having progressed from intuition-driven to solid, experiment-based reasoning. But modern science is not as infallible as it seems—it has erred in the recent past and still does today. To err is human. Given the increasing influence of science on nearly all aspects of daily life, the important question is how efficiently such errors are recognized and corrected.

The basis of every experiment is the acquisition of data. But even if this merely involves counting, it can be astonishingly difficult to obtain reliable data. In the 1950s and 1960s—centuries after the number of fly legs and male teeth had been corrected—James McConnell at Ann Arbor University (MI, USA) carried out experiments to condition planarians to associate a light stimulus with an electric shock so they would scrunch up their bodies in response to light. The educated worm was then ground up and fed to untrained littermates. Once they had cannibalized their brethren, these worms learned to contract in response to light twice as fast as compared with controls, according to McConnell. He concluded that the conditioned memory was stored in a molecule that could be transferred by ingestion (Rose, 1993). Today, we know that memory is not transferable in this way. But the implications of McConnell's experiment—that specific memories are stored in isolatable molecules—caused quite a stir at the time. "Eat your professor", the *New York Times* suggested (Zankl, 2004), and *TIME Magazine* discussed potential misuses, such as a police state or government brainwashing a whole population by lacing tap water.

…modern science is not as infallible as it seems—it has erred in the recent past and still does today

After several groups reproduced the experiments in worms and mammals and achieved the same results, the idea of memory transfer was well established. Questions finally arose from researchers who found it hard to train flatworms, let alone transfer trained behaviour through cannibalism. A fierce discussion followed. Potential confounding factors cited by 'successful' worm trainers included water temperature, time of day, phase of the moon, the direction the worm was heading when the electric shock was applied, or the slime trails it left behind. McConnell himself advised those who attempted to teach planarians to "keep changing the experimental situation until the right conditions are hit" (Travis, 1981). In the end, it was probably background variations—worms spontaneously contract in response to strong light—that facilitated self-deception to such an extent that wishful thinking prevailed.

Major errors also occurred when scientists, carried away by their discoveries, designed experiments to confirm their hypotheses and omitted the usual controls. In 1965, when the technology to visualize chromosomes had just been developed, a study was carried out to test for chromosome abnormalities among men in an institution for dangerous criminals. Among the 196 men studied, 8 showed a duplicated Y chromosome (Jacobs *et al*, 1965). This seemed to be extraordinarily high, so the researchers deduced from the small sample that XYY men tend to become criminals—although nobody had bothered to determine the

©2005 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

background prevalence of XYY karyotypes in the normal population. After this report, numerous screenings for XYY men were carried out in selected groups that were presumed to contain these karyotypes—in penal and mental-penal settings—and conclusions about 'XYY syndrome', not surprisingly, were further strengthened.

The image of the XYY criminal soon became highly popularized. Court proceedings for XYY defendants made legal history by introducing 'genetic predisposition to crime' as a defence strategy. The XYY myth appeared in popular culture, entered science textbooks and sparked a discussion on how criminal behaviour could be prevented through genetic screening. One reason for its success may be its apparent plausibility: aggressive tendencies were thought to be linked to the Y chromosome, thus explaining the perceived differences between men and women. When duplicated, the extra copy would carry these aggressive tendencies beyond usual bounds. The 'criminal chromosome' became a catchy slogan that summarized the idea (Fölsing, 1984; Zankl, 2004).

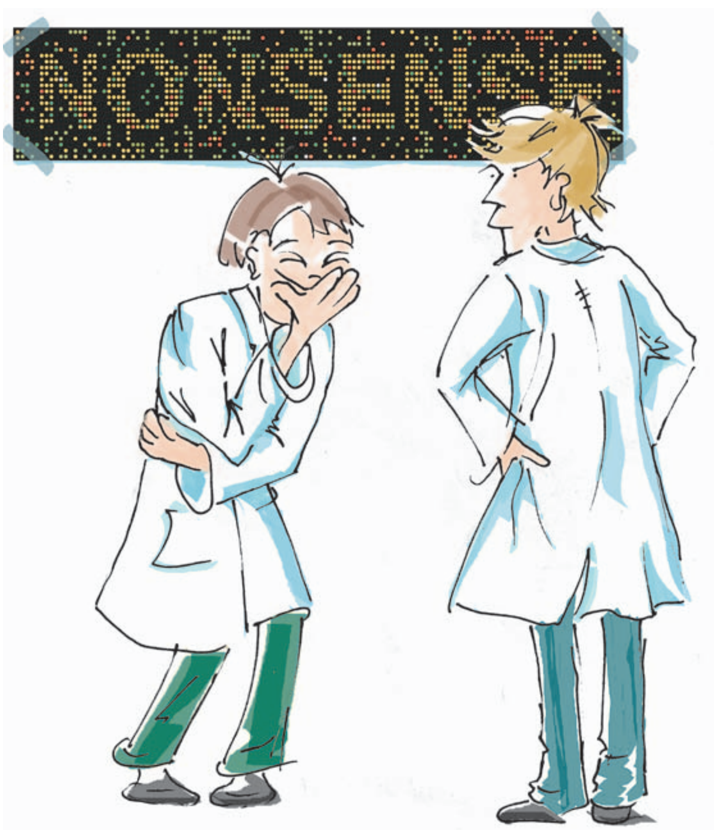
Today, a slightly disproportionate number of XYY men in criminal institutions is indeed expected, but the fact that most XYY men lead perfectly normal lives and in most cases do not even know about their karyotype, was not taken into account in the 1960s and 70s. The scientific reasoning that aggressive tendencies are attached to the Y chromosome has also been challenged. Several studies comparing XYY to XXY men, who also appear in criminal institutions with higher frequencies, found that "the two groups are more alike than different" (Theilgaard, 1984). Witkin and co-workers postulated that, although a wide spectrum of IQ is possible in both groups, their slightly lower average IQ may put XYY and XXY men at higher risk of being apprehended, thus reflecting a higher detection rate rather than a higher prevalence of criminal behaviour (Witkin *et al*, 1976).

Both memory molecules and the inheritance of criminal behaviour were hot topics in their time. In retrospect, it is easy to see how these ideas could spread so rapidly. When science is exciting, data may be over-interpreted in the heat of the moment. A biased researcher may interpret small background fluctuations as meaningful due to wishful thinking, as exemplified by the worm experiments. Critical voices that point to problems in reproducing the experiments arise only hesitantly and are dismissed by claims of incompetence. Bias also leads to experimental design or sample selections that will confirm a hypothesis rather than contradict it, as happened with the XYY myth. And finally, the more plausible a model is or the more it confirms common prejudice, the more open ears it finds, not only in the scientific community but also among non-scientists.

> When science is exciting, data may be over-interpreted in the heat of the moment

Molecular biology is particularly prone to fall into this trap. Its extraordinary successes in the past decades have raised expectations that genetics may answer almost any question about disease, behaviour or human nature. "We used to think our fate is in our stars. Today we know, in large measure, our fate is in our genes," James D. Watson said (Jaroff, 1989). This over-optimistic attitude is also reflected in terms such as 'blueprint' or 'language of God' when referring to the sequence of the human genome (Weigmann, 2004). One may ask whether such a scientific 'gold rush' creates more errors, when research is done with too much enthusiasm and anything that glitters is mistaken for gold.

In addition, novel technologies used in a newly established field may produce errors before the methods are properly established and tested for reliability. Whatever the reason, errors do occur frequently in molecular biology, and some fields are more susceptible than others. "There are fields that require the analysis of high-dimensional data sets, where the likelihood of getting a chance association might be high relative to the likelihood of finding a correct association," said Joel Hirschhorn, Assistant Professor of Genetics and Pediatrics at the Children's Hospital Boston and Harvard Medical School (Boston, MA, USA). And it is in these fields that new studies and methods



©2005 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

spring up like mushrooms, including gene expression studies, proteomics and studies determining the association between genes and disease. In a meta-analysis of gene association studies, Hirschhorn and co-workers found that less than half of the data are confirmed after further analysis, and that the majority are not reproducible (Lohmueller *et al*, 2003).

There is, of course, great potential in gene-association studies. But reproducible and valuable studies are swamped by others that may reflect wishful thinking rather than biological reality. Particularly when sample size is small and results are on the border of significance, scientists may give undue weight to random statistical fluctuations. "For association studies, the criteria of defining something as a positive result have never been agreed on," commented Hirschhorn.

It is not only in fields such as genetics that errors appear. Many studies that look for links between environmental factors and disease probably suffer from the same difficulties in calculating statistical significance. Prenatal exposure to influenza (Brown *et al*, 2004) or lead (Opler *et al*, 2004), parental age (Brown *et al*, 2002), maternal prepregnant body mass (Schaefer *et al*, 2000) and month of birth (Messias *et al*, 2004) have all been reported to affect schizophrenia risk. Similarly, caffeine (Maia & de Mendonca, 2002), cannabinoids (Ramírez *et al*, 2005) and nicotine (Dickerson & Janda, 2003) could potentially reduce the risk of Alzheimer's disease—good news for coffee, marijuana or smoking addicts, but information worth taking with a pinch of salt. The discovery that obesity risk may increase as a result of childhood dieting could have confused cause and correlation (Field *et al*, 2003). Several such relationships may turn out to be a lot weaker than originally thought, if they are reproducible at all.

Presumably, if such studies are not reproduced, they will fall into oblivion before causing major damage. But when models or dogmas are established on the basis of erroneous science, they are all the more difficult to eliminate. The first data to contradict a dogma often have a negligible impact, as the new information frequently does not make sense. "The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...'", as Isaac Asimov, physicist and science fiction novelist, once put it. Only once all these 'funny' data are integrated into a new model is there a chance for the old model to be replaced. "Old models are substituted only when the peers agree," commented Christian Pfeffer, Deputy Editor of the *Journal of Negative Results in Biomedicine*. "It may take a lot of data before a dogma is overthrown."

Findings that contradict a model or reports of the failure to reproduce an experiment may never reach the scientific community due to publication bias. 'Negative results' are rarely published in high-impact journals because they lack the 'news' factor, despite the fact that they may, in the long run, debunk scientific errors. It is only recently that journals are providing a forum for negative results. Even so, these journals still work hard to convince potential authors. "Scientists have a huge incentive to make the most of their data," said Hirschhorn. They tend to redefine the question and start rummaging about for positive results until they find some, rather than publishing negative data, he explained. "Some scientists may simply consider writing up negative results as a waste of time," Pfeffer said, "and others reason that they would only help their competitor if they publish negative data."

> ...the more plausible a model is or the more it confirms common prejudice, the more open ears it finds, not only in the scientific community but also among non-scientists

Even if false models are discarded by the scientific community, the legend can nevertheless persist in the public. Decades after being disproven, the myth of the criminal chromosome appeared in movies such as Alien3 (1992), in which a remote planet serves as a penal colony for XYY men. The comic figure Popeye convinced kids to eat spinach by attributing his superior strength to the leafy greens, despite its iron content having been miscalculated by a factor of 10 at the end of the nineteenth century. The error was detected in the 1930s, but spinach consumption was encouraged after World War II whenever possible, and still is today (Zankl, 2004). But scientific errors can have far more serious consequences than being coerced into eating spinach. A new medical treatment or diagnostic test based on erroneous data may raise false hopes in patients. Attaching unrealistic faith to preventive testing may lead to incorrect diagnoses and patients being treated for a disease they do not have (Gigerenzer, 2002). Genetic predispositions, distinctive features visible in brain scans or physical features such as slow heartbeat or large body size, have been suggested as reasonable indicators of a tendency towards violence in later life, with potentially damaging consequences for the unfortunate few who display these characteristics. In a press release to accompany the publication of a study linking brain size to violence (Raine *et al*, 2000), psychopathologist Adrian Raine from the University of Southern California (Los Angeles, USA) suggested that society should focus its resources on identifying the 5% of children "who will commit 50% of the crime and violence later in life" and channel them into activities that satisfy their stimulation-seeking and aggressive proclivities so they would "contribute to society as bomb-disposal experts, firefighters and test pilots" (Blakeslee, 2000). Most scientists would agree that this is questionable in many ways, but it illustrates how far incorrect estimation of predictability can go.

As the life sciences are having a growing impact on society—be it risk assessment, political counselling or medical treatment—the expectations of the accuracy and reliability of science also increases. Miscalculating the spread of an infectious disease or the success rate of a prospective therapy or diagnostic method can have disastrous consequences on everyday life. In many cases, errors are unavoidable, but premature enthusiasm may forego the need for critical evaluation. It is these cases in which, decades later, one wonders how science could have been so blind. Enthusiasm is an honourable quality in science, but only as long as experiments are carried out carefully and models are constantly challenged. "It is a good morning exercise for a research scientist to discard a pet hypothesis every day before breakfast. It keeps him young," said Konrad Lorenz, an Austrian ethnologist and Nobel laureate.

©2005 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

REFERENCES

Blakeslee M (2000) Size of brain linked to violence. University of Southern California, Los Angeles, CA, USA: Press release, Jan 31

Brown AS, Schaefer CA, Wyatt RJ, Begg MD, Goetz R, Bresnahan MA, Harkavy-Friedman J, Gorman JM, Malaspina D, Susser ES (2002) Paternal age and risk of schizophrenia in adult offspring. *Am J Psychiatry* **159:** 1528–1533

Brown AS, Begg MD, Gravenstein S, Schaefer CA, Wyatt RJ, Bresnahan M, Babulas VP, Susser ES (2004) Serologic evidence of prenatal influenza in the etiology of schizophrenia. *Arch Gen Psychiatry* **61:** 774–780

Dickerson TJ, Janda KD (2003) Glycation of the amyloid β-protein by a nicotine metabolite: a fortuitous chemical dynamic between smoking and Alzheimer's disease. *Proc Natl Acad Sci USA* **100:** 8182–8187

Field AE, Austin SB, Taylor CB, Malspeis S, Rosner B, Rockett HR, Gillman MW, Colditz GA (2003) Relation between dieting and weight change among preadolescents and adolescents. *Pediatrics* **112:** 900–906

Fölsing A (1984) *Der Mogelfaktor.* Hamburg, Germany: Rasch & Röhring

Gigerenzer G (2002) *Calculated Risks: How to Know When Numbers Deceive You.* New York, NY, USA: Simon & Schuster

Jacobs PA, Brunton M, Melville MM, Brittain RP, McClemont WF (1965) Aggressive behavior, mental sub-normality and the XYY male. *Nature* **208:** 1351–1352

Jaroff L (1989) The gene hunt. *TIME Magazine*, Mar 20. www.time.com

Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* **33:** 177–182

Maia L, de Mendonca A (2002) Does caffeine intake protect from Alzheimer's disease? *Eur J Neurol* **9:** 377–382

Messias E, Kirkpatrick B, Bromet E, Ross D, Buchanan RW, Carpenter Jr WT, Tek C, Kendler KS, Walsh D, Dollfus S (2004) Summer birth and deficit schizophrenia: a pooled analysis from 6 countries. *Arch Gen Psychiatry* **61:** 985–989

Opler MG, Brown AS, Graziano J, Desai M, Zheng W, Schaefer C, Factor-Litvak P, Susser ES (2004) Prenatal lead exposure, δ-aminolevulinic acid, and schizophrenia. *Environ Health Perspect* **112:** 548–552

Raine A, Lencz T, Bihrle S, LaCasse L, Colletti P (2000) Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Arch Gen Psychiatry* **57:** 119–127

Ramírez BG, Blázquez C, Gómez del Pulgar T, Guzmán M, de Ceballos ML (2005) Prevention of Alzheimer's disease pathology by cannabinoids: neuroprotection mediated by blockade of microglial activation. *J Neurosci* **25:** 1904–1913

Rose S (1993) *The Making of Memory—From Molecules to Mind.* New York, NY, USA: Anchor Books

Schaefer CA, Brown AS, Wyatt RJ, Kline J, Begg MD, Bresnahan MA, Susser ES (2000) Maternal prepregnant body mass and risk of schizophrenia in adult offspring. *Schizophr Bull* **26:** 275–286

Theilgaard A (1984) A psychological study of the personalities of XYY- and XXY-men. *Acta Psychiatr Scand Suppl* **315:** 1–133

Travis GDL (1981) Replicating replication? Aspects of the social construction of learning in planarian worms. *Soc Stud Sci* **11:** 11–32

Weigmannn K (2004) The code, the text and the language of God. *EMBO Rep* **5:** 116–118

Witkin HA *et al* (1976) Criminality in XYY and XXY men. *Science* **193:** 547–555

Zankl H (2004) *Der grosse Irrtum—Wo Wissenschaft sich täuschte.* Darmstadt, Germany: Primus

Katrin Weigmann

doi:10.1038/sj.embor.7400389