

REVIEW:

Scott E. Page

***The Difference: How The Power of Diversity Creates Better
Groups, Firms, Schools, and Societies***

Princeton, NJ: Princeton University Press, 2007. ISBN: 0-691-12838-3.

by

Yannis M. Ioannides

March 5, 2009

Department of Economics

Tufts University

Medford, MA 02155, USA

617-627-3294 (V) 617-627 3917 (F)

yannis.ioannides@tufts.edu <http://www.tufts.edu/~yioannid/>

Forthcoming, *Journal of Economic Literature* (article section)

Forthcoming, *Journal of Economic Literature* (article section)

Contents

1	Introduction	3
2	Does Diversity Generate Benefits?	4
2.1	The Empirical Evidence	5
2.2	Some Further Thoughts on the Interpretation of the Evidence	12
3	How We See Things	12
4	Summing up	21
5	References	23

1 Introduction

How can one tackle reviewing a book, for which entering "scott e page" the difference in google.com produces 24,000 entries? A *New York Times*, Science Section piece, titled in "Professor's Model, Diversity = Productivity" [Dreifus (2008)], a link to a video presentation at the College de France [College de France (2008)], and numerous reviews in blogs along with commentary are only a few of the hits. Max Bazerman's endorsement reads: "The book is brilliant. Page has a dazzling eclecticism." Kenneth J. Arrow's endorsement reads: "Scott Page has brought to our attention a practically important proposition: diversity of viewpoints is of the greatest importance in solving the problems that face us individually and collectively. Diversity among a group of problem solvers is more important than individual excellence. Page's exposition remarkably combines lightness and breadth of knowledge with rigor and evidence."

There is no obvious way to approach the review, especially since the book is not clearly aimed at a very specific audience. In fact, it is telling how diverse is the group of the "oh-so-many other people" [Page (2007), p. xiv] who commented on parts of earlier versions. I start with a brief summary and then present the book's thesis, with the evidence coming first. The presentation is sprinkled with my reactions throughout. I close with a summary evaluation.

I had known of, but had not really read carefully, Scott Page's work on modeling problem solving and its difficulties before I started reading *The Difference*. My anticipation was high. Full disclosure: I enjoy reading "pop science." Browsing the book raised some concerns. When will it get to the real science? I started reading it as an economist would, meaning with a desire to see what Page does with applying economic thinking in new territory. I did not connect with it completely at first. Still, one of the very early and most pleasant surprises was that the book makes an excellent case for expanding the scope of the language of economists. It helps sharpen one's understanding of how modeling of decision making can provide "toolkits" for studying richer aspects of everyday decision making. I will be providing examples to that effect. The book is fun to read, not because of linguistics-

tic embellishments, but more because of its reliance on “off-the-cuff” casual but endearing commentary — throughout, including footnotes¹ — without breaking the continuity of the text. I acknowledge that some might find the writing too idiosyncratic.

The book’s chapters are organized into Parts One through Five, titled respectively: “Unpacking the Tool Box”; “Diversity’s Benefits: Building from Tools”; “Diverse Values: a Conflict of Interests”; “The Pudding: Does Diversity Generate Benefits?” and “Going on the Offensive.”

In a nutshell, the book makes a good case for the beneficial effects of diversity, the *difference* perspective, for the functioning of social structures, at every scale. It argues that diverse membership of social groups, firms, schools and organizations improves their functioning in addition to being fair. The book’s underlying science is presented in the form of lively interpretations of statistical tools of decision making by means of enticing narratives. To use an analogy, it is reminiscent of how Edward Tufte² offers graphics as a vehicle for getting across cognitive notions. It made me go back and read some of Page’s scholarly papers, coauthored with Lu Hong, on which several of the book’s key ideas are based. I felt enriched after doing that. The more readers are so tempted, the better. The book is self-sufficient, of course, and may be read independently of Hong’s and Page’s papers.

2 Does Diversity Generate Benefits?

Societies are made up of individuals with diverse cognitive skills who organize themselves in diverse cities and productive enterprises. Such arrangements are sustained by a multitude of benefits that originate in the availability of diverse sets of cognitive tools. Differences in individuals’ identities are naturally associated with differences in preferences. Such differences may result, in turn, in ambiguous effects. Preference heterogeneity makes trade

¹Appropriately, this example is in a footnote; see, fn. 14, Chapter 8, Page (2007), p. 391. “I never would have believed that this [Page and Richardson (1992)] would sneak into this book, but it’s an example of how a perspective can move from one problem to another.”

²See, for example, Tufte (2001).

of private goods leave individuals at least as well off as autarky. But provision of public goods requires politically integrated mechanisms and institutions, which are less effective in matching provision with demand the greater is preference heterogeneity. Cognitive and preference diversity lead to higher variance of outcomes. These would be welcome, when diversity comes with sufficiently improved mean outcomes, but might not be so otherwise.

Page identifies three broad types of causes of cognitive diversity. These are training and experiences, which he treats as direct causes, and genes, which he considers as indirect and of minor consequence for his argument. A social scientist with a policy bent would want to know whether cognitive diversity would magnify itself as it is reproduced intergenerationally through training and experience. If there is an upper bound on productive cognitive diversity, then one wonders how would a society mediate its reproduction through its design of the relevant economically and politically integrated mechanisms?

2.1 The Empirical Evidence

References to evidence are dispersed throughout this book of 400 pages. But Chapter 13, “The Empirical Evidence, The Pudding,” *op. cit.* p. 313–335, would at once attract attention by any social scientist, especially by those with a serious interest in empirics. Chapter 13 comes with fifty six endnotes to more than seventy sources, which range from books and scholarly articles to unpublished works and individual suggestions. It is rich and a lot of fun to read. Let us sift through it, or as Page puts it, let’s start “eating.”³

Overall, Page recognizes that whereas cognitive diversity improves performance at problem solving and predicting tasks, the evidence for identity diversity is “far from unequivocal.” He also recognizes that since much of the evidence he discusses is not model-based, it has to be treated with caution. It suffices for him that “diverse groups merely *can* perform better” [*ibid.* p. 315]. Still, he needs evidence on circumstances where true group interactions are involved, not just activity by groups of individuals whose members work individually, and

³The attribution of the opening quotation of this chapter, “The proof of the pudding is in the eating” to Miguel de Cervantes’s *Don Quichote* is disputable. It may be even earlier than Cervantes. See Safire (2008).

where groups have learned how to exploit the benefits of diversity. Page approaches the evidence with a lot of optimism. In order to learn from the evidence one needs to assess what scope there might be for errors to obscure diversity benefits and for communication frictions to prevent attainment of benefits.

Over the ages, societies have learned from their own past experiences and from those of other societies. The story of modern civilization is a story of “nonzero [*social*] interactions” [*ibid.* p. 316] that continuously produce good and bad ideas. There is attrition of bad ideas and along with it of those societies that have adopted them, as well. New skills and techniques are being applied and combined, that is “heuristics” are utilized by “perspectives”.

Heuristics are techniques, sometimes very simple and rough, that we use for finding solutions. The “rule of 72” is a heuristic: it takes $72/(\% \text{ interest rate})$ years for an investment to double. Page defines perspectives as representations of the set of possible solutions. Newton’s conception of planetary motion and Mendeleev’s periodic table of elements are perspectives. Perspectives can take many forms and are intimately linked to scientific progress. For this to be possible, the cognitive tools developed have to be non-rival. For example, once drawing of blueprints was invented, everyone could use them. Although trial and error always play a role, it is the social element in adoption and large-scale experimentation that cements adoption of productivity-improving ideas. Perspectives can simplify problems. E.g., using polar coordinates instead of cartesian ones can vastly simplify certain though not all problems.

As Page says, robust evidence that diversity improves collective performance comes from prediction markets, which are very special types of markets. E.g., in election outcome-based markets, individuals trade precisely because they hold *different* beliefs about the election prospects of different candidates. In the process, the market clearing price for securities defined in terms of candidates’ electoral prospects aggregates the beliefs held by traders over the outcomes. Trading in the Iowa Electronic Markets on the eve of the 2008 US Presidential election produced a set of closing prices stunningly near the result: 0.535, for Barack Obama, and 0.464 for John McCain.⁴

⁴See http://iemweb.biz.uiowa.edu/pricehistory/PriceHistory_GetData.cfm. The outcome from trading on the electoral vote counts on <http://www.intrade.com> is equally impressive.

Page puts to some interesting uses his Diversity Prediction Theorem [*ibid.*, 208–209]. The theorem states that given a group (“crowd”) of predictive models, then the average squared error (collective error) is equal to the average individual error minus the variance of individual signals (prediction diversity). In other words, the larger is the prediction diversity and/or the smaller the variance of individual signals, the smaller the collective error. Thus, a diverse crowd always predicts more accurately than the average of the individuals. Page seeks through this theorem to explain the madness of crowds, by which he means phenomena ranging from random vandalism following sports events to stock market bubbles and crashes. A crowd can err egregiously only if its members lack both accuracy and diversity. If people can communicate with one another and thus share information and feedback, they can increase the accuracy of their models, thus reducing average individual error. They, however, can also become conformist and thus reduce the diversity of their own actions. As individuals often act in reaction to the opinion of others (peer pressure) as well as following their own impulses, it is easy to see that conformism implies greater magnification of individual variation. I have a concern here. Since the scope of this theorem is so wide, at least as Page sees it, it would be important to know how well it performs empirically in different instances of “stampedes” in the economic, social and political spheres and of varying seriousness.

Societies regularly take advantage of the superiority of collective performance in problem solving. Societies deliberately put together teams of diverse experts, “small-scale versions of Bletchley Park⁵” [*ibid.*, p. 322], to attack difficult problems. Organization theorists generally agree, Page argues, that teams of cognitively diverse agents improve rates of innovation over individuals in at least some tasks. Researchers would not engage in collaborative research, if doing so did not pay off. While science clearly went first for low-hanging fruit, interactions rather than collaborations among innovators (and scientists) have always been very important [Mokyr (2002)]. The transition to mass production came, as Page puts it, as “cognitive toolboxes diffused” [*ibid.*, p. 317]. Nowadays, rapidly growing scientific collaborations are

⁵Bletchley Park, also known as Station X, is an estate located in the town of Bletchley, Buckinghamshire, UK. During World War II, Bletchley Park famously hosted the United Kingdom’s main codebreaking establishment (code-named Ultra), whose successful endeavors included deciphering the German Enigma and Lorenz encoding machines.

facilitated by vastly decreased costs of collaboration [Goyal *et al.* (2006)].

I would add to Page's theoretical arguments that individuals representing diverse perspectives and interests can also bring continuity in committee decisions. Crémer (1986) shows that participation in organizations of infinite duration changes the incentives of agents with finite lives (a form of diversity *par excellence*), and will induce more cooperation than a static model would predict. This applies to the case, for example, of committees whose members have overlapping terms, like the Federal Open Markets Committee that sets US monetary policy.

Identity diversity is beneficial, Page argues, only if it is linked to cognitive diversity and provided that it is interpreted net of its costs. It is not surprising that the empirical evidence is mixed, but Page is prepared to be persuaded with only gross benefits being positive, instead of net ones. Page refers to Ellison *et al.* (2005) who report that when they measure labor force diversity of a particular large firm in terms of gender and tenure, they find the following: higher office-level gender diversity is associated with lower employee cooperation (and morale and satisfaction), but that tenure diversity has little or no effect. And, offices where the employees think the firm is accepting of diversity tend to be more cooperative, and have higher morale and satisfaction. That is, workers like the idea of diverse work-places and accordingly go along with it, but are actually more comfortable in homogeneous settings. Yet, more homogeneous units seem to be less productive overall, perhaps because they possess a less varied portfolio of talents on which to call.

Referring to a large number of studies conducted by others, Page concludes that the average net effect of diversity in workplace settings is modest, while there is evidence of costs, especially in the form of higher variance. Thus, indeed, gross benefits are likely to be positive. Racially diverse groups predict better, especially when race is a salient factor. Diversity in jury settings prompts whites to be more subtle and to base their prediction of guilt or innocence of a black defendant on finer interpretations. Again, it is in problem solving settings where diverse identities are associated with task-relevant tools that identity diversity pays off. Still, Page is quite optimistic, as a user of social science, when he states that the higher variance associated with identity diverse groups simply reflects the fact that

“we just haven’t quite figured out how to get along, and when we do, diversity improves performance” [*ibid.*, p. 328]. What if such variance is inherent in work-place diversity? I would also point to some encouraging evidence that diversity costs at the individual firm level may be offset, from the social viewpoint, at equilibrium through increased supply [Niederle *et al.* (2008)].

It is natural to wonder about how diversity of cities and of countries may be related to their economic performance. There is no doubt that there are benefits to clustering in cities, for firms and individuals. Or else, we would observe economic activity spread out evenly over space. As Alfred Marshall argued long ago, dense concentration of economic activity improves productivity for several reasons. One is known as human capital or knowledge spillovers, and is like sharing of cognitive tools, as Page refers to it. Workers casually exchange knowledge about technology and production conditions at their places of employment. Such transfers happen fortuitously, but are sometimes sought out deliberately by firms, as anecdotal evidence about life in California’s Silicon Valley testifies.

A second reason is labor market pooling. At any point in time, firms are subject to idiosyncratic shocks (e.g. as a result of changing demand for their particular products) that makes them want to hire or fire workers. If these shocks are idiosyncratic, when one firm is firing, another firm may well be hiring. If firms locate in close proximity to one another, it is easy for workers to move from firms experiencing bad times to those experiencing good times. In Silicon Valley “[t]he joke is that you can change jobs and not change parking lots” [Saxenian (1994), p. xi]. Agglomeration in effect helps workers insure against idiosyncratic shocks. It is also attractive for firms because it weakens the impact on wages of their own idiosyncratic shocks. Even though employment in bad times is more costly than it would otherwise have been, this is more than compensated for by the ease of expansion in good times [Krugman (1991), 38–49, 123–127; Overman and Puga (2009)].

A third force for agglomeration comes from the greater variety of intermediate products and richer mix of labor skills and expertise that are available in larger urban areas. Greater variety of goods and services lowers prices and wages and also enhances firms’ options in choosing technologies for production and distribution of their products [Krugman, *op. cit.*;

Rosenthal and Strange (2004)]. Clearly, variety is an important force of benefits of urban areas and increases with size. Within spillovers, generally, researchers have sought to identify whether individuals benefit from the presence of other individuals with similar, higher or simply diverse skills. Similarly, firms may benefit from concentration of other firms in their own industry, known as localization or Marshall–Arrow–Romer externalities, or from greater industrial concentration (Jacobs externalities). There undoubtedly exist both these types of externalities, but how important are they, quantitatively?

I would think that the range of estimates of 6 to 20 percent, reported by Page, for the effect of doubling city size on the productivity per worker is somewhat on the upper side. The estimation of such effects is notoriously difficult in part because of definitional problems and in part because of self-selection. What is a city? If a city doubles, do we reproduce its composition in terms of the socioeconomic profile of the population? Do we have an American city in mind, more precisely the concept most often used namely a US metropolitan area, or do cities in less developed countries also qualify? If smarter (or better educated) people are attracted by the amenities of large city life, and then tend to work harder, not everything is held constant when we compare productivity across urban and rural populations.

Beyond such compositional issues, two types of estimates must be mentioned. One is due to Ciccone and Hall (1996), who use state-level data and estimate the effect at 6%. This was recently updated by Chen and Weisberger (2007), who use metro-level as well as state-level industry data. These researchers seek to distinguish between agglomeration effects, for which the evidence is robust, and the balance of agglomeration and congestion effects, which the literature refers to as “net increasing returns to scale,” for which the evidence is much weaker. When they use time- and state-fixed effects, they estimate the net returns from agglomeration to be negative, at around -1%. When they include cyclical controls, their estimate falls to -7%.⁶ Another approach, due to Lucas (2001), acknowledges that the intensity of effects attenuates with distance between firms, which is in fact confirmed

⁶From these two point estimates they are tempted to conclude that on net, there are no increasing returns to scale. However, when they instrument for density using either lagged density or lagged birthrates, their estimates of the net effect are again positive rising up to 5%, for gross state product, but still varying quite a lot for different one-digit industries.

empirically by Rosenthal and Strange (2003). It leads to estimates that are close to their upper bound, which is the share of land in urban production, 0.05 in Lucas's case. But, there are dissenting views; see Ciccone and Peri (2006), in particular, who use US data and find net agglomeration effects of 0%, at the city level, and around 2%, at the state level. All these studies point to lower estimates for the benefits of agglomeration than Page seems comfortable with.

That urban size is a blessing due to coexistence of many activities is a strong argument in favor of Page's interpretation. In a fundamental sense, human societies engage in continuous problem solving, that goes on in schools, factories and firms' headquarters, and is facilitated by the serendipity of urban encounters that are conducive to creativity. Cultural diversity is a growth factor, but racial diversity might not be.

Ethnic diversity at a national scale has, as Page thoughtfully argues, a complex effect on a country's economic performance. Federal countries can be wealthy in spite of their ethnic diversity, like Canada, Switzerland and the US. These countries have developed effective political institutions that are flexible enough to withstand upheavals and to accommodate the need to balance majority-based decision making with adequate representation of minority and constituent issues. Ethnic diversity that is accommodated within federated structures can mitigate inter-ethnic conflict associated with day-to-day frictions, while being consistent with the notion that broader resource allocation questions may be resolved only at the level where sovereignty is recognized. Arguing along the lines of the arguments in Alesina, Spoloare and Wacziarg (2000), one must recognize that this cannot be seen independently of the expansion of international trade and the emergence of international multilateral institutions. For example, increasing reliance on free trade worldwide favors breakup of what otherwise would have been multiethnic political entities. Therefore, it is hard to draw general conclusions about the economic consequences of ethnic diversity. So, Page's conclusion that evidence that identity diversity in countries, cities or groups, is conducive to improved economic performance is not fully persuasive.

Overall, Page is honest in not overstating the case for empirical support of his claims. Since that is not his emphasis either, I would not insist further. But, he should be more

forceful in encouraging such research. In addition, the book does rely extensively on results from simulations, which are summarized in tables and referred extensively in the book. Since he has carried out the simulations himself, and is clearly very talented in that area, why not provide a web site where interested people can try out things on their own, like tinkering with parameter values and the like? We know that doing so sometimes leads to further breakthroughs in theorizing, especially in areas where closed-form solutions were originally thought as impossible to obtain.

2.2 Some Further Thoughts on the Interpretation of the Evidence

But, one wonders, have groups, schools, firms, and societies sought to organize themselves so as to best avail of the power of diversity, let alone succeed in doing so? Is the evidence sampled randomly from instances of successes and failures? Or, should we worry that there is sample selection and we only get to see the success stories? The objects of study that would bear upon the evidence are very heterogeneous, and in particular, in terms of size. Is there anything that can be learned from the heterogeneity in the evidence itself?

3 How We See Things

We differ in how we see things. Some of us occasionally focus hard on the visible (to all) and invisible (to some) parts of our world and come up with organizing principles. Scientists working with models are at ease with these notions. The lay public might mistake the organizing principles for purported truth. For example, entropy is a theoretical concept, not the truth. Mendeleev's periodic table of elements codified knowledge as he saw it, but also encouraged speculation that vacant spots on the table could be filled once previously unknown existing elements have been discovered. The notion of discovery must be broadly understood, in the sense that new elements being discovered might be very short-lived. Discovery may not be codifiable in the absence of representation. Page prefers the notion of a perspective, which as defined above encompasses representations that encode objects, events,

or situations. Mendeleev’s discovery indicates how the choice of perspective facilitates discovery. Perspectives must signal functional relationships between underlying objects within the context of internal languages.

Page directs attention to the fact that different scientists work with different visual representations, such as chemical bonds for chemists or supply and demand diagrams for economists. It is clearly essential for scientific progress for a science to avail itself of abstract forms of representation. Whether they are strictly mathematical or not, they should uniquely depict the objects of interest. Building novel perspectives in terms of other perspectives helps organize knowledge, and the more obvious they may seem *ex post* the more effective they have been. Of course, not all perspectives are obvious, like the notion that we humans are made of “tiny vibrating strings” — clearly, that is not so obvious. (I like this mnemonic).

Perspectives simplify problems, but they do not constitute unique representations of objects. Hong and Page (2001) see this in a more abstract setting, which also facilitates a number of other claims. Hong and Page prove that collections of diverse agents can locate optimal solutions to difficult problems, even if agents’ abilities are bounded. The trick lies in the fact that however unsophisticated a problem solver’s approach is, she may reach an optimal solution to a difficult problem, provided that she has the right perspective, the right representation. The Hong–Page theory of problem solving does not impose very specific structure on problem solving production functions, even when problem solvers have identical ability. The way problem solvers approach a problem matters for the ease with which it may be solved. Problem solving, as understood by Hong and Page, involves representing problems by means of bits and agents’ actions as the flipping of bits. Hong and Page argues persuasively that binary strings are the most concise and convenient way of addressing problem solving. Still, from the perspective of the book, it is quite fascinating that Page gets so much conceptual mileage out of abstract ideas that have very restricted symbolic representations. I am wondering, however, how well these problem solvers would do with respect to computationally complex tasks. As modeled by computer scientists, computational complexity offers formal assessments of the computational difficulty of problems [Papadimitriou

(1994)]. How do Page's and Hong's problem solvers stand up to such scrutiny?

Heuristics are rules, which when applied with perspectives clarify problems. But there exists a large number of heuristics, and within professional and scientific disciplines, new heuristics are created regularly. Different types of heuristics are helpful in different disciplines. A management problem may be successfully attacked with a heuristic that says "go after that fifth of matters to be decided that account for eighty percent of the benefits." Of course, that would be the most important "fifth" of decisions, provided that the benefits are quantifiable. Other problems, like personnel problems, may involve matters of ethics and notions of fairness, typically to be understood within the norms of the particular enterprises, that are not inherently quantifiable. One wonders, does the Hong and Page approach apply to such settings?

We learn and innovate by creating finer partitions that restrict the range of the interpretations that we can draw, based on events we observe. Interpretations may allow us to learn from events that we had not anticipated because they were previously ignored. We form interpretations by ignoring certain aspects and highlighting others in order to draw causal inferences. Page defines an interpretation as a category of objects, a map from different kinds of objects into words. A word in an interpretation may not represent an object uniquely and yet must involve some regularity. For example, an interpretation, like that of a "Soccer Mom," performs an informational role only if married women of certain age, with soccer-playing children and working spouses behave in a manner that is distinct from that of other members of a suburban community. Such categorizations may be helpful in some instances of discourse, but are not helpful in others, like in predicting electoral outcomes. For example, there is no reason to expect that soccer moms would vote similarly. Speech is made more concise when interpretations conjure up images of certain types of behavior, even if this is so due to correlation and not necessarily causation. Making such distinctions finer may be critical for the usefulness of categories. Since reality consists of numerous attributes, most of us consider only some and are prepared to lump together many into broader categories.

All this applies to perception of space, especially in urban areas where people live with mental maps and associated heuristics. Page recalls a powerful example about how Chicago's

urban transit authority redefined transport links as hubs and thus added new options to commuters at lower costs. Rethinking of familiar objects and creating specialized objects that may be used in modular fashions typically improves productivity and user satisfaction. For example, Mary Phelps Jacob, a New York socialite, received a patent in 1913 for her invention of the first modern brassiere, a substitute for corset-like garments that women used to wear. It completely revolutionized the women’s apparel industry.

Chapter 4 deals with predictions, broadly construed — the book’s niche. A predictive model is an interpretation together with a prediction for each set or category created by the interpretation. Page thinks of a prediction as something broader than most people think.

Let’s say the world may be in a state x , one of the elements of set X . An *outcome function* maps each possible state to an outcome, $F : X \rightarrow \Theta$. The outcome may be binary, no/yes, ($\Theta = \{0, 1\}$), in the case of a classification problem, or a parameter value in $\Theta = [0, \infty)$, in the case of an estimation problem. An interpretation partitions the states of the world into distinct categories. I classify politicians as either conservatives or liberals. My neighbor Joan does so in terms of such “identity” categories as age, race and gender. An individual i ’s interpretation is the set of categories that she uses to partition the state of the world: $\Phi_i = \{\phi_{i1}, \dots, \phi_{im}\}$. $\Phi_i(x)$ denotes the category in i ’s interpretation that state of the world x belongs to. For example, Anna is more sophisticated than me, if every category in her interpretation is contained in a category of mine, in other words, her categories are finer than mine: $\Phi_A(x) \subseteq \Phi_Y, \forall x \in X$, with strict inclusion for at least one state of the world. A predictive model $M_i(x)$ is a signal from each state of the world, that is a mapping from states of the world into an outcome, such that if two different states are in the same category then they map into the same outcome.

It is important in Page’s view that such signals, *interpreted* signals, be distinguished from *generated* signals. The latter are simply random draws from a distribution. Interpreted signals are filtered, they come from a controlled sampling, not a random sampling process, as the above description makes clear. Filtering them through partitions and predictive models renders them no longer independent draws from a distribution, unlike generated signals, which have the appropriate statistical properties.

Chapter 5 provides a nice summary of how different people may excel in different things, in the use of different toolboxes. There are diminishing returns in any one of these toolboxes, but by adding toolboxes, we also add additional capabilities to our portfolios. So, people differ in terms of such portfolios, which complement those of others when we consider collective problem solving and prediction by diverse people.

Chapter 6, arguably the heart of the book, takes up how diversity pays off. It sets out with tight analytical foundations, which are borrowed from Hong and Page (2001) and model problem solving as an activity by firms and groups of individuals, and Theorem 1, Hong and Page (2004). Absorbing it demands a bit of a commitment, so let me interpret it. First, it is important for the main idea to think of heterogeneity in how individuals represent problems internally, their perspectives, and in the algorithms they use to generate solutions, their heuristics. Problem solving as an economic activity does not have to satisfy familiar properties. That is, while a collection of bounded but diverse agents can locate optimal solutions to difficult problems, problem solving firms can exhibit arbitrary marginal returns to problem solvers. The order in which problem solvers are applied to a problem can matter, so that the standard story of decreasing returns to scale may not apply to problem solving firms.

Second, as a group of problem solvers becomes large, the group of individually best performing among them become more and more alike and likely to search in similar areas of the solution space. Therefore, the gain to the problem solving ability of a society is offset by individuals' diminishing diversity in their functioning as problem solvers, and regardless of how their joint activity is organized. Individuals who are randomly chosen, on the other hand, are very unlikely to have common locally acceptable solutions. This is subject to conditions, of course: the problem must be *hard*, in the sense that no individual problem solver can always locate the global optimum; all problem solvers can always find a local optimum for the problem; the global optimum does not coincide with the local optimum of a non-trivial group of problem solvers; and the numbers of problem solvers is large. The result is powerful: the best groups of problem solvers need not be groups of the best!

The central message of Chapter 6 is subtle, requires qualifications and relies on diversity, a

property of groups of individuals. An individual's value to a problem solving process depends on her ability to help improve the collective decision, a point forcefully made by Hong and Page (2001). So, it is contextual in that it depends on the perspectives and heuristics of all others who work on the problem. Whereas an individual's own ability, as measured by IQ or SAT scores, may be a good predictor of her own problem solving ability, it is how differently she thinks that matters critically for collective decision making. Since this may well vary with the particular setting, it is appropriate to think of a person's ability as the sum total of perspectives, heuristics and mental models.

Even if solvers have identical abilities, their contributions to group problem solving can be arbitrary, making hard to attribute solutions to specific individuals and requiring that compensation arrangements not be fully linked to individual contributions. And, of course, with individuals' being complex creatures themselves, the benefits of diversity "also apply within individuals." That suggests an area where real empirical support would be most welcome. Is the story of more effective people one that rests on "internal diversity," whereby such people may be able to deal with diverse circumstances with radically different ways? Or, is it that more talented people can be good in many things?

Chapter 7 is about information aggregation. An influence by Surowiecki's *Wisdom of Crowds* [Surowiecki (2004)] is evident. Surowiecki's necessary conditions are critical: the members of the crowd must use diverse predictive models, they should predict and not influence one another. In this fashion, a collection of people, each of whom knows well a part, "can know the whole" [*ibid.* p. 181], provided each one is somewhat accurate and all of them somewhat diverse, so as to offset each other's biases. However, there are many instances where interdependencies among members of a group are important for its cohesiveness and survival.

Chapter 8 takes this theme further. For a crowd to be wise its members must be individually smart or collectively diverse; it is better yet, if they are both. Polling a crowd is inferior to an information market, where people can place bets and stand to benefit (or lose) from their bets. Less accurate predictions are driven out, and more accurate ones survive, provided that people have confidence in their predictions. This is exactly what a market

setting elicits: people put their money where their prediction is. A crowd of models characterized by spontaneous diversity is necessary for good predictions.⁷ Large crowds are more likely to satisfy these conditions, but smaller crowds, like committees and juries, must be encouraged to employ diverse models. Empirical evidence here may be anecdotal, as it is very hard to quantify the logically coherent but diverse mental approaches which successful investors like Warren Buffett use instinctively.

Chapter 9 is about diverse preferences, where a distinction is drawn between *fundamental* preferences and *instrumental* preferences. The former is valuing what we want, and the latter is about how we get what we want. The latter involves internalizing how what we want may be obtained in a particular social setting. Fundamental preference for a large house is valuing it in its own right. When one believes that the attention of certain members of the opposite sex may be easier to secure if he has a larger house than his neighbors, we have an instance of instrumental preferences for a large house. All of us want everyone to have access to good schools, less inequality, crime-free neighborhoods, etc. There are many politically integrated ways to achieve such objectives. People interested in different things look very differently at ways of acquiring them. When fundamental preferences are diverse, underprovision of common resources is likely. Differences in fundamental versus instrumental preferences diversity are critical for the composition of decision-making bodies. Fundamental preference diversity calls for diverse committees.

Chapter 10 deals with aggregation of preferences, a task that is necessary for groups, teams, communities and societies to be able to function. The chapter reviews the massive

⁷For a collection of individuals, the collective prediction is the average (in the set-theoretic sense) of individual predictions: $\bar{M}(x) = \sum_i M_i(x)$. The collective prediction is more accurate the more individually sophisticated the individual prediction models are. That is, by partitioning the states of the world into many, fine, categories, collectively diverse prediction models create different partitions. The interpreted signals of two different individuals 1 and 2 are based on independent interpretations, if and only if the probability of the intersection of any two categories, with one being used by 1 and the other by 2, is equal to the product of the respective probabilities. If two individuals use independent interpretations, then they look at different dimensions given the same representation. Hong and Page (2008) show (under some restrictions on the individuals' predictive models) that for problems with binary outcomes (classification problems), independent interpreted signals must imply negatively correlated predictions.

literature in how societies and formal decision making bodies aggregate preferences, taking individual preferences as given. It dwells at length and elegantly on Arrow's Impossibility Theorem. Recall the theorem's powerful simplicity. It is impossible to arrive at an aggregation of (arbitrary) preferences over more than two alternatives, that is based on individual preference orderings (which are individually rational, that is complete and transitive) that satisfies unanimity, independence of irrelevant alternatives and non-dictatorship. Arrow's theorem predicts unpleasant outcomes, not only that it may be impossible to aggregate preferences, but also that shrewd committee people can manipulate voting outcomes by suitable sequencing of decisions, and that the tragedy of the commons may be inescapable. A number of newer, noteworthy but less well known, theorems provide additional rigor, and Page does a superb job in reviewing them. This is one of the best parts of this fascinating book.

Specifically, Charles Plott has proven that when preferences involve more than one dimension, there will be no alternative that dominates every other alternative in a pairwise vote. However, for what the theorem really implies, it matters if preferences are instrumental or fundamental. The theorem is less problematic in the latter case, because the same objective may be attained in different ways, which might create fewer tensions, instead of the need to deviate from fundamental preferences. McKelvey's cycling theorem and its generalization by Schofield assert that when preferences are multidimensional, starting from any alternative, appropriate sequences of majority votes can get a society to any other alternative. Page cites empirical evidence from districts in Los Angeles County that confirms that the representativeness of choices made by voters, as indicated by the preference of the median voter, is greater the more homogeneous they are. Given a vast variety of voting settings in a country like the US, one surely hopes that a lot more evidence may be brought to bear on this issue. Social choice impossibility theorems are in practice about worst-case scenarios. It is important to know the practical consequences of logical limitations on social decision making.

And then there is Gibbard–Satterthwaite: what happens to social decision making if people misrepresent their preferences in order to achieve palatable outcomes, an instance of what is commonly known as strategic voting? This theorem states that with more than two

outcomes, any nondictatorial decision rule for aggregating diverse preferences is manipulable. That is, knowing how others would vote, an individual can manipulate the outcome by deviating from her true preferences in order to get an outcome that has a chance to win, instead of an ideal outcome. Again, the consequences are more serious for diverse fundamental than instrumental preferences. People have less incentives to misrepresent their preferences in the latter case. Preference diversity has consequences in how much people trust one another. Perhaps it would be here than a richer set of empirics would be most helpful, which seems to be possible. But, the conceptual discussion is enriching, although some of the metaphors used are a bit too homey and might turn off people unnecessarily.

Chapter 14 is a compendium of how the book's ideas can be put to work. A central message of the book is that societies need diverse perspectives — multitudes, dissonance — and not a single perspective, interpretation, heuristic or predictive model. Page discusses the following experiment: how can we predict how much leather would be produced by a cow? A crowd of diverse models would do best. One model would be along the lines of typical engineering problem solving: start from a sphere, stretch it, and keep deforming it, while keeping track of the surface areas of the resulting objects. Alternatively, consider Page's suggestion that "we might even tape together a few Gateway Computer boxes⁸ until we've reached cowlike proportions." [*ibid.* p. 342]. This is an instance where Page's knack for mnemonics comes in handy. A crowd of models is supportive of diversity in hiring and university admissions.

Different types of diversity will produce further diversity. For that, it might be necessary to destroy walls erected and maintained by stereotyping. Most interesting in this connection is Page's argument that by making stereotypical inferences about people with different identities, "we reduce their incentives to accumulate tools outside these stereotypes. We limit opportunity. To use Glenn Loury's phrasing, we stigmatize [*ibid.*, p. 366]." Stereotyping is in an algorithm that ignores individual differences and leads to inaccurate predictions.

A second message is that diverse perspectives on how organizations solve problems come

⁸Readers should know that Gateway, Inc., a computer manufacturer, uses as its logo a cow-like pattern, black on white, which features prominently on its shipping material.

from the use of complete outsiders in certain roles, provided that doing so lets them stay as outsiders, or else they would no longer think differently. This justifies the practice of external reviews of organizations, like academic departments in universities, by groups of outsiders, or the hiring of consultants. External review boards do not have to be made up of individuals with comparable abilities to those being reviewed. External reviewers might not necessarily be experts themselves, but they are outsiders who are familiar with the business conducted in universities. This notion also underscores the value of interdisciplinary perspectives.

A third message is that the business of society may be equally well served by a diverse citizenry who possess diverse predictive models. This increases the likelihood of providing more checks on bad ideas. To Page's speculation that the lack of cognitive diversity "might explain the lack of stability in developing countries," I would add that such countries need diversity of interpretations (in the sense of Page) in order to be able to cope in an complex international arena. Social decision making depends critically on the nature of preferences. If individuals' preferences are diverse and fundamental, then we are more likely to see misrepresentations of preferences and manipulations of agendas. If they are diverse and instrumental, they simply involve disagreements over the means but not the ends. E.g., two officemates who both enjoy fresh air may disagree over keeping their office window open, because one of them dislikes bugs getting in. This is an instance where a disagreement may be reconciled once the nature of the object has been clarified: "... purchase a screen for the window ..." [*ibid.* p. 350]. This is a simple remedy, for sure.

4 Summing up

Page's rhetoric moves us way beyond accepting diversity as a matter of taste and on to realms where diversity affects the "production function," too, that is the social opportunity set defined in terms of social outcomes. He says at the closing of the text: "when we meet people who think differently than we do, ..., we should see opportunity and possibility. We should recognize that a talented 'I' and a talented 'they' can become even more talented 'we.'" That happy vision rests not on blind optimism, or catchy mantras.

It rest on logic. A logic of diversity.” [*ibid.*, p. 375].

This is most emphatically an eloquent manifesto in defence of *diversity*, of *The Difference*. Page is, at the same time, offering a different way to “market” economic science, and his goals are lofty indeed. Page’s fascinating work is provoking economics in a new ways, especially in applying it to problem solving. Still, it would take our “battle-hardened” empirical toolkits to make the point of *The Difference* fully compelling and thus to help establish a “science of difference.” In the treacherous world of empirics, we cannot always control for everything, identities can be deceiving [Loury (2008)], and causality can be nebulous, because computation, too, can be emergent [Crutchfield and Mitchell (1995); Ioannides (2008)]. There also remains the issue of how to implement policies that enhance diversity by improving the representation of disadvantaged groups. In democratic societies such policies may involve discriminatory practices and lead to dilemmas, especially when identities are not obvious nor visible [Fryer and Loury (2007)]. Still, there is enormous public interest in overcoming the accumulated effects from past disadvantaged access to resources and skills.

We should not expect that all these important issues would be addressed by a single book. This book does make an excellent case for the value of further research and thinking about implementation of diversity-enhancing policies. Within such an agenda, sharply focused empirical research deserves special attention. It is therefore even more well-worth reading!

5 References

- Alesina, Alberto, Enrico Spolaore, and Romain Wacziarg. 2000. “Economic Integration and Political Disintegration.” *American Economic Review* 90(5):1276–1296.
- Chen, Yu-chin, and Noah Weisberger. 2007. “Local Labor Markets and Increasing Returns to Scale: How Strong Is The Evidence?” working paper, Department of Economics, University of Washington.
- Ciccone, A., and R. E. Hall, 1996, “Productivity and the Density of Economic Activity,” *American Economic Review*, 86: 54-70.
- Ciccone, Antonio, and Giovanni Peri. 2006. “Identifying Human-Capital Externalities: Theory with Applications.” *Review of Economic Studies* 73: 381 – 412.
- College de France. 2008. “La sagesse collective : principes et mécanismes.” (Collective Wisdom : Principles and Mechanisms) http://cerimes.education.fr/sagesse-collective-principes-mecanismes-catalogue-college-france-philosophie-fiches-a223154s100001017t88_.html
- Crémer, Jacques. 1986 “Cooperation in Ongoing Organizations.” *The Quarterly Journal of Economics* 101(1):33–49.
- Crutchfield, James P., and Melanie Mitchell. 1995. “The Evolution of Emergent Computation.” *Proceedings of the National Academy of Sciences USA*. 92: 10742–10746.
- Dreifus, Claudia. 2008. “Professor’s Model, Diversity = Productivity.” *New York Times* January 8; accessed, December 1, 2008, at:
<http://www.nytimes.com/2008/01/08/science/08conv.html>
- Fryer, Roland G., and Glenn C. Loury. 2007. “Valuing Identity: The Simple Economics of Affirmative Action Policies.” working paper, Department of Economics, Harvard University.
- Goyal, Sanjeev, Marco J. van der Leij, and José Luis Moraga-González. 2006. “Economics: An Emerging Small World.” *Journal of Political Economy*. 114(2):403-412.

- Hong, Lu, and Scott E. Page. 2001. "Problem Solving by Heterogeneous Agents." *Journal of Economic Theory* 97:123–163.
- Hong, Lu, and Scott E. Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences USA*. 101(46): 16385–16389.
- Hong, Lu, and Scott E. Page. 2008. "Interpreted and Generated Signals." working paper, Departments of Political Science and Economics, University of Michigan, August.
- Ioannides, Yannis M. 2008. "Emergence." Durlauf, Steven N., and Laurence E. Blume, eds., *The Palgrave Dictionary of Economics*, 2nd ed., Basingstoke, Hampshire: PalgraveMcMillan.
- Krugman, Paul. 1991. *Geography and Trade*. Cambridge, MA: MIT Press.
- Loury, Glenn C. 2008. "Is He One of Us? Reflections on Identity and Authenticity." The 2008 Opening Convocation Address Brown University, September.
<http://today.brown.edu/articles/2008/09/loury>
- Lucas, Jr. Robert E. 2001. "Externalities and Cities." *Review of Economic Dynamics* 4(2):245–274.
- Mokyr, Joel. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton, NJ: Princeton University Press.
- Niederle, Muriel, Carmit Segal, and Lise Vesterlund. 2008. "How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." NBER working paper No. 13923.
- Overman, Henry G., and Diego Puga. 2009. "Labor Pooling as a Source of Agglomeration: An Empirical Investigation." presented at the Conference on the Economics of Agglomeration, NBER Cambridge, MA, December 2007.

- Papadimitriou, Christos H. 1994. *Computational Complexity*. Reading, MA: Addison Wesley.
- Rosenthal, Stuart A., and William C. Strange. 2003. "Geography, Industrial Organization and Agglomeration." *Review of Economics and Statistics* 85(2):377–393.
- Rosenthal, Stuart A., and William C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." in Henderson, J. Vernon, and Jacques-Francois Thisse, eds., *Handbook of Regional and Urban Economics: Vol. 4, Cities and Geography*, North-Holland.
- Safire, William. 2008. "The Proof Isn't really in the Pudding." *International Herald Tribune* October 19; accessed on December 1, 2008 at:
<http://www.iht.com/bin/printfriendly.php?id=17077463>
- Saxenian, AnnaLee. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
- Surowiecki, James. 2004. *The Wisdom of Crowds: Why The Many Are Smarter than The Few and how Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York, NY: Doubleday.
- Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.