The Two Cultures of Mathematics.

W. T. Gowers

In his famous Rede lecture of 1959, entitled "The Two Cultures", C. P. Snow argued that the lack of communication between the humanities and the sciences was very harmful, and he particularly criticized those working in the humanities for their lack of understanding of science. One of the most memorable passages draws attention to a lack of symmetry which still exists, in a milder form, forty years later:

> A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold: it was also negative. Yet I was asking something which is about the scientific equivalent of: *Have you read a work of Shakespeare's?*

I would like to argue that a similar sociological phenomenon can be observed within pure mathematics, and that this is not an entirely healthy state of affairs.

The "two cultures" I wish to discuss will be familiar to all professional mathematicians. Loosely speaking, I mean the distinction between mathematicians who regard their central aim as being to solve problems, and those who are more concerned with building and understanding theories. This difference of attitude has been remarked on by many people, and I do not claim any credit for noticing it. As with most categorizations, it involves a certain oversimplification, but not so much as to make it useless. If you are unsure to which class you belong, then consider the following two statements.

(i) The point of solving problems is to understand mathematics better.

(ii) The point of understanding mathematics is to become better able to solve problems.

Most mathematicians would say that there is truth in both (i) and (ii). Not all problems are equally interesting, and one way of distinguishing the more interesting ones is to demonstrate that they improve our understanding of mathematics as a whole. Equally, if somebody spends many years struggling to understand a difficult area of mathematics, but does not actually do anything with this understanding, then why should anybody else care? However, many, and perhaps most, mathematicians will not agree equally strongly with the two statements. Certainly, Sir Michael Atiyah would not, as he showed in an interview in 1984 [A1].

> MINIO: How do you select a problem to study?
>
> ATIYAH: I think that presupposes an answer. I don't think that's the way I work at all. Some people may sit back and say, "I want to solve this problem" and they sit down and say, "How do I solve this problem?" I don't. I just move around in the mathematical waters, thinking about things, being curious, interested, talking to people, stirring up ideas; things emerge and I follow them up. Or I see something which connects up with something else I know about, and I try to put them together and things develop. I have practically never started off with any idea of what I'm going to be doing or where it's going to go. I'm interested in mathematics; I talk, I learn, I discuss and then interesting questions simply emerge. I have never started off with a particular goal, except the goal of understanding mathematics.

This interview first appeared in the Mathematical Intelligencer, but was reprinted in the general section of Atiyah's collected works. I strongly recommend his general essays and lectures on mathematics to anybody who wishes to sort out their own ideas about the significance of mathematics, and shall refer to them several times in this article.

Another person who would not have attached equal weight to the two statements was Paul Erdős, who bequeathed to the world an enormous number of fascinating problems, as well as solutions to many others, but is not associated to the same extent with the development of theory. This is not to deny that Erdős was trying to understand mathematics: many people who have solved an Erdős problem (alas, I am not one of them) will testify that, as they have thought harder and harder about it, they have been led in unexpectedly fruitful directions and come to realize that the problem was more than the amusing curiosity that it might at first have seemed. So when I say that mathematicians can be classified into theory-builders and problem-solvers, I am talking about their *priorities*, rather than making the ridiculous claim that they are exclusively devoted to only one sort of mathematical activity.

It is obvious that mathematics needs both sorts of mathematicians (as Atiyah himself says at the end of [A2]). It is equally obvious that different branches of mathematics require different aptitudes. In some, such as algebraic number theory, or the cluster of subjects now known simply as Geometry, it seems (to an outsider at least - I have no authority for what I am saying) to be important for many reasons to build up a considerable expertise and knowledge of the work of other mathematicians are doing, as progress is often the result of clever combinations of a wide range of existing results. Moreover, if one selects a problem, works on it in isolation for a few years and finally solves it, there is a danger, unless the problem is very famous, that it will no longer be regarded as all that significant.

At the other end of the spectrum is, for example, graph theory, where the basic object, a graph, can be immediately comprehended. One will not get anywhere in graph theory by sitting in an armchair and trying to *understand graphs* better. Neither is it particularly necessary to read much of the literature before tackling a problem: it is of course helpful to be aware of some of the most important techniques, but the interesting problems tend to be open precisely because the established techniques cannot easily be applied.

Let me now briefly mention an asymmetry similar to the one pointed out so forcefully by C. P. Snow. It is that the subjects that appeal to theory-builders are, at the moment, much more fashionable than the ones that appeal to problem-solvers. Moreover, mathematicians in the theory-building areas often regard what they are doing as the central core (Atiyah uses this exact phrase) of mathematics, with subjects such as combinatorics thought of as peripheral and not particularly relevant to the main aims of mathematics. One can almost imagine a gathering of highly educated mathematicians expressing their incredulity at the ignorance of combinatorialists, most of whom could say nothing intelligent about quantum groups, mirror symmetry, Calabi-Yau manifolds, the Yang-Mills equation, solitons or even cohomology. If a combinatorialist were to interrupt such a gathering and ask roughly how many subsets of $\{1, 2, \ldots, n\}$ can be found such that the symmetric difference of any two of them has size at least $n/3$, the response might very well be a little frosty. (This problem is very easy if and only if one knows the appropriate technique, which is to choose sets randomly and show that the chances of any given pair of them having a symmetric difference of size less than $n/3$ are exponentially small. So the answer is e^{cn} for some constant $c > 0$.)

My main purpose here is to defend some of the less fashionable subjects against criticisms commonly made of them. I shall devote most of my attention to combinatorics, since this is the area I know best. However, what I say applies to other areas as well. I often use the word "combinatorics" not quite in its conventional sense, but as a general term to refer to problems that it is reasonable to attack more or less from first principles. (This is really a matter of degree rather than an absolute distinction.) Such problems need not be discrete in character or have much to do with counting. Nevertheless, there is a considerable overlap between this sort of mathematics and combinatorics as it is conventionally understood.

Why should problem-solving subjects be less highly regarded than theoretical ones? To answer this question we must consider a more fundamental one: what makes one piece of mathematics more interesting than another? Once again, Atiyah writes very clearly and sensibly on this matter (while acknowledging his debt to earlier great mathematicians such as Poincaré and Weyl). He makes the point (see for example [A2]) that so much mathematics is produced that it is not possible for all of it to be remembered. The processes of abstraction and generalization are therefore very important as a means of making sense of the huge mass of raw data (that is, proofs of individual theorems) and enabling at least some of it to be passed on. The results that will last are the ones that can be organized coherently and explained economically to future generations of mathematicians. Of course, some results will be remembered because they solve very famous problems, but even these, if they do not fit into an organizing framework, are unlikely to be studied in detail by more than a handful of mathematicians.

Thus, it is useful to think not so much about the intrinsic interest of a mathematical result as about how effectively that result can be communicated to other mathematicians, both present and future. Combinatorics appears to many to consist of a large number of isolated problems and results, and therefore to be at a disadvantage in this respect. Each result individually may well require enormous ingenuity, but ingenious people exist, especially in Hungary, and future generations of combinatorialists will not have the time or inclination to read and admire more than a tiny fraction of their output.

Let me attempt to answer this criticism. It is certainly rare in combinatorics for somebody to find a very general statement which suddenly places a large number of existing results in their proper context. It is also true that many of the results proved by combinatorialists are somewhat isolated and will be completely forgotten (but this does not distinguish combinatorics from any other branch of mathematics). However, it is not true that there is no structure at all to the subject. The reason it appears to many mathematicians as though combinatorics is just a miscellaneous collection of individual problems and results is that the organizing principles are less explicit.

If the processes of abstraction and generalization, which are so important in mathematics, are of only limited use in combinatorics, then how can the subject be transmitted to future generations? One way of thinking about this question is to ask what the requirements of tomorrow's combinatorialists are likely to be. As I have said already, their priority is likely to be solving problems, so their interest in one of today's results will be closely related to whether, by understanding it, they will improve their own problem-solving ability. And this brings us straight to the heart of the matter. The important ideas of combinatorics do not usually appear in the form of precisely stated theorems, but more often as general principles of wide applicability.

An example will help to make this point more clearly. One form of Ramsey's theorem is the following statement.

Theorem. *For every positive integer k there is a positive integer N, such that if the edges of the complete graph on N vertices are all coloured either red or blue, then there must be k vertices such that all edges joining them have the same colour.*

The least integer N that works is known as $R(k)$.

The following argument shows that $R(k) \leq 2^{2k}$. Let G be a graph with 2^{2k} vertices, and let us suppose for convenience that the vertices are totally ordered. Let x_1 be the first vertex. Then by the pigeonhole principle there is a set of vertices A_1 of size at least 2^{2k-1} such that every edge from x_1 to A_1 has the same colour. Now let x_2 be the least vertex of A_1. By the pigeonhole principle again there is a set $A_2 \subset A_1$ of size at least 2^{2k-2} such that every edge from x_2 to A_2 has the same colour. Continuing this process, we obtain a sequence $x_1, \ldots, x_{2k}$ of vertices and a sequence $A_1 \supset \ldots \supset A_{2k}$ of sets such that $x_i \in A_{i-1}$ for every i, and every edge from x_i to A_i has the same colour. It follows that the colour of the edge joining x_i to x_j depends only on $\min\{i, j\}$. By the pigeonhole principle again, we can therefore find a subset H of $\{x_1, \ldots, x_{2k}\}$ of size at least k such that this colour is always the same, so that all edges joining vertices in H have the same colour.

A slight variant of the above argument improves the bound to $\binom{2k}{k}$. However, I am more interested here in *lower* bounds for $R(k)$. One of Erdős's most famous results is that $R(k) \geq 2^{k/2}$, and it is proved as follows. Instead of trying to construct a clever colouring, one simply colours the edges randomly in the most obvious way. That is, each edge is red with probablity $1/2$, and blue with probability $1/2$, and all the choices are independent. Let the number of vertices be N, and let $\{x_1, \ldots, x_k\}$ be a set of k of them. The probability that every x_i is joined to every x_j with a red edge is $2^{-\binom{k}{2}}$, and so is the probability that they are all joined by blue edges. Therefore, the expected number of sets of k vertices all joined by edges of the same colour is $2^{1-\binom{k}{2}}\binom{N}{k}$. If this is less than one, then it must be possible for there to be no such sets of k vertices. A small calculation shows that it is less than one when $N = 2^{k/2}$.

This result of Erdős [E] is famous not because it has large numbers of applications, nor because it is difficult, nor because it solved a long-standing open problem. Its fame rests on the fact that it opened the floodgates to probabilistic arguments in combinatorics. If you understand Erdős's simple argument (or one of many other similar arguments) then, lodged in your mind will be a general principle along the following lines:

> if one is trying to maximize the size of some structure under certain constraints, and if the constraints seem to force the extremal examples to be spread about in a uniform sort of way, then choosing an example randomly is likely to give a good answer.

Once you become aware of this principle, your mathematical power immediately increases. Problems such as the one I mentioned earlier of finding a large number of sets with large symmetric differences between any two of them suddenly switch from being impossible to being almost trivial.

Of course, there is much more to probabilistic combinatorics than keeping an eye out for simple applications of the above principle. For example, one often decides to use probabilistic methods and then finds that estimating the relevant probabilities is not easy. However, a great deal of work has now been done in this area, and many clever techniques invented. Some of these can again be encapsulated in the form of useful principles. One of them is the following, known to its friends as Concentration of Measure:

> if a function f depends in a reasonably continuous way on a large number of small variables, then $f(x)$ is almost always close to its expected value.

The full significance of measure concentration was first realized by Vitali Milman in his revolutionary proof [M] of the following theorem of Dvoretzky [D].

Theorem. *For every positive integer k and every $\epsilon > 0$ there is an n such that every n-dimensional normed space X has a k-dimensional subspace with Banach-Mazur distance* *at most $1+\epsilon$ from ℓ_2^k.*

This is equivalent to saying that we can find vectors $x_1,\dots,x_k \in X$ such that

$$\Big(\sum_{i=1}^k |a_i|^2\Big)^{1/2} \le \Big\|\sum_{i=1}^k a_i x_i\Big\| \le (1+\epsilon)\Big(\sum_{i=1}^k |a_i|^2\Big)^{1/2}$$

for every sequence $a_1,\dots,a_k$ of scalars. A more geometrical (and even more counterintuitive) way to reformulate the theorem is to say that every n-dimensional symmetric convex body has a k-dimensional central cross-section which contains a k-dimensional ellipsoid B and is contained in $(1+\epsilon)B$, and is therefore almost ellipsoidal itself.

Milman's approach to the theorem was to choose a k-dimensional subspace of X randomly. Before doing this, one must choose a sensible probability measure, which can be done using a theorem of Fritz John. This states that there is a basis $x_1,\dots,x_n$ of X which is not hopelessly far from orthonormal, in the sense that

$$\Big(\sum_{i=1}^n |a_i|^2\Big)^{1/2} \le \Big\|\sum_{i=1}^n a_i x_i\Big\| \le \sqrt{n}\Big(\sum_{i=1}^n |a_i|^2\Big)^{1/2}$$

for every sequence $a_1,\dots,a_k$ of scalars. One then takes the natural measure on the Grassmannian $G_{n,k}$, with respect to this basis. (Actually, this is an oversimplification, but the precise details need not concern us here.)

In the case of the two results I mentioned earlier, it was easy to see, with hindsight, that random methods were sensible. By contrast, in this example there seems to be no reason to expect them to work. After all, one might think that a typical cross-section of an irregular convex body would itself be irregular. However, once one is familiar with the idea of concentration of measure, one notices that the norm of a vector $\sum_{i=1}^n a_i x_i$ depends on several variables a_i. Moreover, since we are interested only in the ratio of this norm to the ℓ_2-norm $\Big(\sum_{i=1}^n |a_i|^2\Big)^{1/2}$, we can assume that none of the a_i vary by all that much. Therefore, by concentration of measure, we might expect the ratio of the X-norm to the ℓ_2-norm to be very close to its expected value nearly all the time. And now the idea of almost ellipsoidal cross-sections has ceased to be counterintuitive.

The main result needed to make the above argument precise is the following consequence of Lévy's isoperimetric inequality on the sphere. Let $f : S_n \to \mathbb{R}$ be a function with median M. Let $A \subset S_n$ be the set of all points x such that $f(x) \le M$. Then the probability that a random point in S_n has distance more than ϵ from A is at most $\sqrt{\pi/2}\exp(-\epsilon^2 n/2)$. Replacing f by $-f$ shows also that almost every y is close to a point x with $f(x) \geq M$. Now let $f(a_1, \ldots, a_n) = \left\| \sum_{i=1}^n a_i x_i \right\|$. Since f is reasonably continuous, and since most points y are close to a points $x \in A$, it follows that $f(y)$ is not much larger than M. Similarly, for most y, $f(y)$ is not much smaller than M.

Dvoretzky's theorem, especially as proved by Milman, is a milestone in the local (that is, finite-dimensional) theory of Banach spaces. While I feel sorry for a mathematician who cannot see its intrinsic appeal, this appeal on its own does not explain the enormous influence that the proof has had, well beyond Banach space theory, as a result of planting the idea of measure concentration in the minds of many mathematicians. Huge numbers of papers have now been published exploiting this idea or giving new techniques for showing that it holds.

I could mention many more general principles such as these. Here are just a few. They are not of equal importance, and, as I have said, they are not precise, but they have all been exploited a great deal.

(i) Obviously if events $E_1, \ldots, E_n$ are independent and have non-zero probability, then with non-zero probability they all happen at once. In fact, this can be usefully true even if there is a very limited dependence. [EL,J]

(ii) All graphs are basically made out of a few random-like pieces, and we know how those behave. [Sze]

(iii) If one is counting solutions, inside a given set, to a linear equation, then it is enough, and usually easier, to estimate Fourier coefficients of the characteristic function of the set.

(iv) Many of the properties associated with random graphs are equivalent, and can therefore be taken as sensible definitions of pseudo-random graphs. [CGW,T]

(v) Sometimes, the set of all eventually zero sequences of zeros and ones is a good model for separable Banach spaces, or at least allows one to generate interesting hypotheses.

My main point about such principles is not so much that they are useful, which is not particularly surprising, but that they play the organizing role in combinatorics that deep theorems of great generality play in more theoretical subjects. When one is trying to understand a result, it saves a lot of time if one can reduce it to two or three main ideas.

Having done this, one may feel no need to follow the details closely. Being familiar with general principles and a few examples of their application, one knows that one could work out the details oneself if necessary. For example, I memorize Erdős's proof of the lower bound for $R(k)$ as follows: choose the colouring randomly and do the obvious calculations. Another example is the following remarkable theorem of Kashin [K].

Theorem. *For every positive integer n there is an orthogonal decomposition of $\mathbb{R}^{2n}$ (with respect to the usual inner product) into two n-dimensional subspaces X and Y such that for every $x \in X \cup Y$ the ratio of the ℓ_1-norm of x to the ℓ_2-norm of x lies between $c\sqrt{2n}$ and $\sqrt{2n}$, for some absolute constant $c > 0$.*

A proof by Szarek [Sza] of this theorem goes like this (if you know the right vague ideas): a simple volume argument shows that almost none of the unit ball of ℓ_1^n is contained in the corners (a rough word meaning parts where the ratio of the ℓ_1-norm to the ℓ_2-norm is small), so it is easy to show that a random decomposition works.

A third example is the following theorem of Roth [R].

Theorem. *For every $\delta > 0$ there exists N such that every subset of $\{1, 2, \ldots, N\}$ of size at least δN (that is, density at least δ) contains an arithmetic progression of length three.*

The condensed proof of this is as follows. Look at a subset $A \subset \mathbb{Z}/N\mathbb{Z}$ instead and consider the Fourier coefficients of the characteristic function of A. If these are small (apart from $\hat{A}(0)$) then A is essentially random, so it contains plenty of progressions of length three. Otherwise, there is a large coefficient which allows us to pass to a subprogression of $\{1, 2, \ldots, N\}$ where A has larger density.

Let me summarize what I have said so far. I have been trying to counter the suggestion that the subject of combinatorics has very little structure and consists of nothing but a large number of problems. While the structure is less obvious than it is in many other subjects, it is there in the form of somewhat vague general statements that allow proofs to be condensed in the mind, and therefore more easily memorized and more easily transmitted to others.

There are, however, many other criticisms of combinatorics. One I have heard is that the subject lacks direction, or goals of a general kind. Another is that the subject is not particularly deep. A third is that it does not have interesting connections to other parts of mathematics (meaning the "central core" of mathematics). A fourth is that many of the results do not have applications.

These criticisms can be answered in a similar way. Consider first the notion that there are not general goals in combinatorics. I quote again from the interview with Atiyah [A1]:

> I was thinking more of the tendency today for people to develop whole areas of mathematics on their own, in a rather abstract fashion. They just go on beavering away. If you ask what is it all for, what is its significance, what does it connect with, you find that they don't know.

Atiyah was not particularly referring to combinatorics, but he makes a powerful point, and it is as important for combinatorialists as it is for anyone else to show that they are doing more than merely beavering away.

Some branches of mathematics are dominated by a small number of problems of universally acknowledged importance. One can justify many results by saying that, in however small a way, they shed light on the Riemann hypothesis, the Birch-Swinnerton-Dyer conjecture, Thurston's geometrization conjecture, the Novikov conjecture or something of the kind. Other branches of mathematics derive their appeal from an abundance of mysterious phenomena that demand explanation. These might be striking numerical coincidences suggesting a deep relationship between areas that appear on the surface to have nothing to do with each other, arguments which prove interesting results by brute force and therefore do not satisfactorily explain them, proofs that apparently depend on a series of happy accidents or heuristic arguments that work well but are hard to make rigorous.

It would be difficult to demonstrate that combinatorics had many general goals of the sort just mentioned (with the one very definite exception of the P=NP problem). However, just as the true significance of a result in combinatorics is very often not the result itself, but something less explicit that one learns from the proof, so the general goals of combinatorics are not always explicitly stated. To illustrate this, let me return to Ramsey's theorem and the bounds for the function $R(k)$. Despite the simplicity of the arguments I gave earlier, they give almost the best known estimates. To be precise, the following problem is open.

Problem. *(i) Does there exist a constant $a > \sqrt{2}$ such that $R(k) \geq a^k$ for all sufficiently large k. (ii) Does there exist a constant $b < 4$ such that $R(k) \leq b^k$ for all sufficiently large k?*

I consider this to be one of the major problems in combinatorics and have devoted many months of my life unsuccessfully trying to solve it. And yet I feel almost embarrassed to write this, conscious as I am that many mathematicians would regard the question as more of a *puzzle* than a serious mathematical problem.

The reason I think so highly of this problem is that it has become clear to me (as it has to many others who have tried it) that it is very unlikely to be solved by a clever *ad hoc* argument, tailored to this problem only. (Actually, I am referring mainly to the part of the problem concerning the upper bound for $R(k)$.) To be a little vague about it, the proof that $R(k) \leq 4^k$ has a local character, in the sense that one throws away most of the graph and concentrates on small neighbourhoods of a few vertices. A better bound seems to demand a more global argument, involving the whole graph, and there is simply no adequate model for such an argument in graph theory. Therefore, a solution to this problem is almost bound to introduce a major new technique.

One of the frustrations of the problem is that, while random colourings do not give a lower bound significantly better than $2^{k/2}$, it looks as though no departure from randomness (such as partitioning the vertices into five classes and making the probability of an edge being red vary according to whether it lies inside a class or joins two classes) improves the lower bound. However, if one tries to pursue this line of thought, one gets bogged down with separate analyses of all sorts of different graphs. None of them is individually threatening, but they are hard to pursue systematically. This has led me to fantasize that there might be a sort of classification of red-blue colourings (or, equivalently, graphs) which would enable one to solve the problem by checking for each class that the bound had to be less than $(3.99)^k$, say.

The idea of classifying graphs sounds odd at first, so let me state a result which has been applied over and over again. First, we need to define a notion of a pseudorandomness. Let G be a graph and let A and B be disjoint sets of vertices in G. The pair (A, B) is said to be ϵ-*uniform* if there exists a real number $\alpha > 0$ such that, whenever $A' \subset A$ and $B' \subset B$ are sets of cardinality at least $\epsilon|A|$ and $\epsilon|B|$ respectively, the number of edges joining A' to B' differs from $\alpha|A'||B'|$ by at most $\epsilon|A'||B'|$. If ϵ is small, this says that the bipartite graph consisting of the edges of G between A and B is rather like a random graph with edge-probability α. The next result is due to Szemerédi [Sze], and is known as his uniformity lemma, or regularity lemma.

Theorem. *Let G be a graph, let $\epsilon > 0$ and let k be a positive integer. Then there exists a constant K, depending on k and ϵ only, such that the vertices of G can be partitioned into m sets $A_1, \ldots, A_m$, with $k \leq m \leq K$, such that at least $(1-\epsilon)\binom{m}{2}$ of the pairs (A_i, A_j) (with $i < j$) are ϵ-uniform.*

This, as the alert reader will have noticed, is a precise formulation of the general principle (ii) mentioned earlier (but I stress that not all such principles can be made precise).

Unfortunately, although Szemerédi's uniformity lemma is an ideal tool for many problems, there are many others, such as finding better bounds for Ramsey's theorem, about which it tells us nothing. Therefore, one general goal of graph theory is to find more detailed and more refined classifications. Another, somewhat opposed to it, is to find ways of doing without Szemerédi's uniformity lemma. Significant progress in either of these projects would have a correspondingly significant impact on graph theory. (This is almost a tautology, since a good measure of the progress is whether it allows one to solve interesting problems.)

In general, as one gains experience at solving problems in an area such as combinatorics, one finds that certain difficulties recur. It may not be possible to express these difficulties in the form of a precisely stated conjecture, so instead one often focuses on a particular problem which involves those difficulties. The problem then takes on an importance which goes beyond merely finding out whether the answer to it is yes or no. This explains why it was possible for so many of Erdős's problems to have hidden depths.

What of the criticism that combinatorics is a shallow subject? One of the great satisfactions of mathematics is that, by standing on giants' shoulders, as the saying goes, we can reach heights undreamt of by earlier generations. However, most papers in combinatorics are self-contained, or demand at most a small amount of background knowledge on the part of the reader. Contrast that with a theorem in algebraic number theory, which might take years to understand if one begins with the knowledge of a typical undergraduate syllabus.

This criticism reflects the different priorities of theory-builders and problem-solvers. A theory-builder will tend to say that Theorem A is deep because it uses Theorem B which uses Theorem C etc., all of which were, individually, significant results. A problem-solver may well not have a long chain of logical dependences of this kind. However, if we consider a more appropriate kind of dependence, based on general principles again, then the picture changes. It will often be the case that, while there is no formal dependence between two results, there would have been no hope of proving one of them if one was unaware of the general principles introduced in the proof of the other. Chains of this kind of dependence can be quite long, so combinatorialists too can have the satisfaction of solving problems that would have been well out of reach a generation ago. In this way, one feels that the subject as a whole is progressing.

Now all my arguments so far have been internal to combinatorics. I have tried to show that the subject has a coherence and sense of direction that are not obvious to an outsider but are nonetheless important. However, I have said nothing about how mathematics as a whole can benefit from progress in combinatorics. Let us once again consider what Atiyah has to say about matters of this sort [A3, §6].

> ... the ultimate justification for doing mathematics is intimately related with its overall unity. If we grant that, on purely utilitarian grounds, mathematics justifies itself by some of its applications, then the whole of mathematics acquires a rationale provided it remains a connected whole. Any part that drifts away from the main body of the field has then to justify itself in a more direct fashion.

How does combinatorics fare if we accept the need for a justification of this kind? An obvious remark is that one can in fact justify combinatorics in a direct fashion, because of its intimate connection with computer science, the utility of which is obvious. (Strangely, when the programme committee of the 1998 International Congress of Mathematicians listed the connections between the various sections, they did not recognise this one.) As for connections with other subjects, there are applications of combinatorics to probability, set theory, cryptography, communication theory, the geometry of Banach spaces, harmonic analysis, number theory ... the list goes on and on. However, I am aware as I write this that many of these applications would fail to impress a differential geometer, for example, who might regard all of them as belonging somehow to that rather foreign part of mathematics that can be safely disregarded. Even the applications to number theory are to the "wrong sort" of number theory.

It is perhaps helpful to consider various different ways that one branch of mathematics can be of benefit to another. Here is a list, in order of directness (but not exhaustive), of how area A can help area B.

(i) A theorem of A has an immediate and useful consequence in B.

(ii) A theorem of A has a consequence in B, but it takes some work to prove it.

(iii) A theorem of A resembles a question in B sufficiently closely to enable one to imitate or adapt the A-proof and answer the B-question.

(iv) In order to solve a problem in A, one is led to develop tools in B which are of independent interest.

(v) Area B contains definitions which resemble those of area A. (To give just one example, sometimes one wishes to define a notion of independence which behaves in some but not all ways like linear independence in vector spaces.) Area A then suggests fruitful ways of organizing and tackling the results and problems in B.

(vi) If one gains an expertise in area A, then one picks up certain habits of thought which enable one to make a significant contribution to area B.

(vii) Area A is sufficiently close in spirit to area B, that anybody who is good at area A is likely to be good at area B. Moreover, many mathematicians make contributions to both areas.

The less direct relationships such as (iv) to (vii) are correspondingly less visible. Nevertheless, their contribution to the interconnectedness of mathematics should not be underestimated. I feel this particularly acutely after working for several years in the geometry of Banach spaces. My initial reason for working in the area was simply that I found results such as Dvoretzky's theorem, and several very natural open problems, fascinating. Later I learnt that my chosen branch of mathematics was strongly criticized for having broken free from its classical roots in differential equations and thereby lost its purpose. I would concede that there are not many connections of type (i) and (ii) from pure Banach space theory to other areas, although there are some deep ideas which many feel are underexploited. However, as soon as one considers the looser connections, they are there in abundance. The exploitation of concentration of measure provides a good example of (iii) (or (iv) - the classification is somewhat artificial). As for (vi) and (vii), I can speak from personal experience, having worked recently on several problems outside Banach space theory. Although I have not applied Banach-space results, my experience in that area has enabled me to think about problems in ways that would not otherwise have occurred to me. And I am far from alone in this, as many mathematicians who have worked on Banach spaces have worked successfully in other areas as well, such as harmonic analysis, partial differential equations, C^*-algebras, probability and combinatorics.

The only way of dismissing these connections, it seems to me, is if one believes that they are all connections between one uninteresting, unimportant area and another. This is such a radical view, and dismisses so much of modern mathematics, including many results with direct practical applications, that it is hard to believe that anybody would take it seriously. Yet it seems that some do, which brings me back to the idea that there are two cultures within pure mathematics. It is probably true that there are more connections *within* the problem-solving and theory-building parts of mathematics than there are *between* them, which is why the "two cultures" tag is appropriate. (I should say once again that these labels are useful oversimplifications, and that I fully recognise that many people are attracted to what I have called theory-building subjects because they would ultimately like to contribute to the solution of certain problems.)

If it is true that pure mathematics can be divided into two broad cultures with not a great deal of communication between them, one can still ask whether this matters. In my opinion it does. One reason is that this situation has many undesirable practical consequences. For example, mathematicians from one culture may well find themselves making decisions that affect the careers of mathematicians from the other. If there is little mutual understanding between the two cultures, then making such decisions fairly, which is difficult under the best of circumstances, becomes even harder. (I myself have absolutely no complaint, but I have had a very lucky career.) A second effect is that potential research students who are naturally suited to one culture can find themselves under pressure to work in an area from the other, and end up wasting their talent. This is a particular danger in a department which becomes over-dominated by a small number of subjects.

These effects are perhaps inevitable by-products of academic life, and are not my main reason for urging better communication across the great divide. The most important disadvantage of the current lack of communication is that it represents a great missed opportunity. I have occasionally heard mathematicians on the theory side complain of a problem that it has been attacked with all the known tools, but that a stubborn core remains which is "essentially combinatorics". This is a way of saying that the problem is very difficult, but, more to the point, difficult in exactly the way that attracts mathematicians of a problem-solving mentality. Unfortunately, it is also difficult to reach a level of understanding where one can appreciate the essentially combinatorial nature of the underlying problem. Such a situation is tailor-made for cross-cultural collaboration, but collaboration of this kind would require greater efforts on the part of problem-solvers to learn a bit of theory, and greater sympathy on the part of theoreticians towards mathematicians who do not know what cohomology is. Such efforts cannot fail to enrich both mathematical cultures.

References.

[A1] M. F. Atiyah, *An interview with Michael Atiyah*, Math. Intelligencer **6** (1984), 9-19.

[A2] M. F. Atiyah, *How research is carried out*, Bull. I.M.A. **10** (1974), 232-4.

[A3] M. F. Atiyah, *Identifying progress in mathematics*, ESF conference in Colmar, C.U.P. (1985), 24-41.

[CGW] F. Chung, R. L. Graham and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1990), 345-362.

[D] A. Dvoretzky, *Some results on convex bodies and Banach spaces*, Proc. Symp. on Linear Spaces, Jerusalem 1961, 123-160.

[E] P. Erdős, *Some remarks on the theory of graphs*, Bull. A.M.S. **53**, 292-4.

[EL] P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, in A. Hajnal et al. eds, Infinite and finite sets, North-Holland, Amsterdam, 609-628.

[J] S. Janson, *Poisson approximation for large deviations*, Random Structures and Algorithms **1**, 221-230.

[K] B. S. Kashin, *Sections of some finite dimensional sets and classes of smooth functions*, Isv. ANSSSR, ser. mat. **41** (1977), 334-351 (Russian).

[M] V. D. Milman, *A new proof of the theorem of A. Dvoretzky on sections of convex bodies*, Funct. Anal. Appl. **5** (1971), 28-37 (translated from Russian).

[Sza] S. J. Szarek, *On Kashin's almost Euclidean orthogonal decomposition of* ℓ_1^n, Bull. Acad. Polon. Sci. **26** (1978), 691-694.

[Sze] E. Szemerédi, *Regular partitions of graphs*, in Proc. Colloque Inter. CNRS, J.-C. Bermond et al. eds, (1978), 399-401.

[R] K. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245-252.

[T] A. Thomason, *Pseudo-random graphs*, in M. Karonski, ed., Proceedings of Random Graphs, Poznán 1985, Annals of Discrete Mathematics **33**, 307-331.