Stimulating Presentation of Theorems Followed by Responsive Proofs*

NITSA MOVSHOVITZ-HADAR

* This paper was presented at the annual conference of the Israel Mathematical Society held at the University of Tel-Aviv. February 1987

A mathematics teacher's task is often to convey a certain theorem In practice, this task usually comprises two parts: a presentation of the theorem followed by a presentation of a proof Planning how to do this is a decision-making process of selecting the method by means of which each part of the presentation should take place, and the sequential path connecting the two parts What ought we to consider as we make our choices? What are the options? The more we think about them the greater the variety.

Focussing on two theorems this paper demonstrates several ways to present a theorem. Two ways of presentation, one for each theorem, are more stimulating than the others, and therefore are claimed to set the stage for the coming proofs more successfully. In proving the two theorems, several methods of presenting a proof are illustrated The outcome of each is considered from the point of view of the learner, keeping in mind that "a good proof is one which makes us wiser' Manin, 1977 p. 51]

THEOREM PRESENTATIONS

Theorem 1: A very special square matrix

A surprising exposition

The matrix in Table 1 is an 8×8 matrix. Suppose we open a lesson with this task: "Please choose any 4×4 submatrix and select four numbers in it, such that no two are in the same row nor in the same column. Note the sum of the four numbers. Remaining in the same 4×4 submatrix, sum a few more sets of four such numbers, no two of them in the same row or column "

It does not take long before the students are taken by surprise, as they start to realize that all these sums are the same. Consequently, many students decide to try another submatrix, one of a different size, perhaps, to see whether the same phenomenon holds there too Others feel eager to examine the orginal 8×8 table and find out what it is that makes all these sums equal. Soon enough they observe that Table is 1 is not a random collection of 64 numbers. There is a fixed common-difference (of 1) between any two adjacent elements in the *same* column. But, what does this have to do with n elements each taken from a *different* row and column?

At this point, most of the students are not as yet able to answer the questions raised by the exposition However, many might be ready to conjecture that for *any* positive integer n, and for *any* $n \times n$ matrix over the real field having the properties of Table 1, the sum of *any* set of n elements, no two of which are in the same row nor in the same column, is invariant

10	11	12	13	14	15	16	17
19	20	21	22	23	24	25	26
28	29	30	31	32	33	34	35
37	38	39	40	41	42	43	44
46	47	48	49	50	51	52	53
55	56	57	58	59	60	61	62
64	65	66	67	68	69	70	71
73	74	75	76	77	78	79	80

A representative matrix for the first Theorem

Table 1

While students are busy with, say, a 4×4 submatrix of their choice, calculating the sum of a few sets of four elements satisfying the stated property, the instructor can circulate among them and quickly tell each student his or her sum This can be done either by glancing at, or by asking for, just two diagonally opposite corner-numbers To obtain the common sum, these two numbers are added and the result multiplied by two For any submatrix, the sum of two diagonally opposite corner-numbers should be multiplied by half the dimension of the submatrix For example, in Table 1 as a whole, the common sum for any eight elements as defined above is $(10+80)\times 8/2 = 360$. This is a variation of Mel Stover's Calendar Trick described by Martin Gardner [1956, p 49]. The trick suggests that there is something special about the diagonals. By the time the students are ready to come up with a generalizing conjecture, some may have realized that the n elements in either of the diagonals are particular sets of n elements satisfying the requirements. A proof is now appropriate We shall return to it later on

The surprises incorporated in this presentation, the wondering reaction they yield, and the student-initiated behaviours that follow, indicate that intellectual curiosity has been triggered. This presentation makes them want to know something about matrices, something which they

For the Learning of Mathematics **8**, 2 (June 1988)
FLM Publishing Association, Montreal, Quebec, Canada

could not have cared less about a moment earlier

Would the following presentation have the same effect?

A symbolic verbal presentation

The symbolic presentation is typical of mathematics textbooks and journals. Mathematics instructors usually feel that it is unlikely that many of the audience in the mathematics class will make any sense of this presentation. They therefore intertwine verbal interpretations to shed some meaning on the formal mathematical language. Such a presentation would look somewhat like this:

Verbal Presentation *Teachers Says*	Symbolic Presentation *Teacher writes*
The theorem we are going to prove today is very interesting. It claims that given any square matrix with real entries having two properties:	*Theorem*: Let A be an $n \times n$ matrix over the field R with entries $A_{i\,j}$ such that
The rows are arithmetic progressions all with the same common difference, and the columns are also arithmetic progressions, all with the same common difference,	$A_{i,\,j+1} - A_{i\,j} = C_1$ $A_{i+1\,j} - A_{i\,j} = C_2$ where C_1, C_2 are constants and $i, j = 1, \ldots, n$
then	then there exists a real number C_3 such that for any $A_{i_1\,j_1}; A_{i_2\,j_2}; \ldots; A_{i_n\,j_n}$
any *n* elements of which no two are in the same row or column add up to the same sum	for which for $r,s = 1, \ldots, n$ $r \neq s \Rightarrow i_r \neq i_s;\ j_r \neq j_s$ $\sum_{k=1}^{n} A_{i_k j_k} = C_3$

Taking this approach is an easy escape for the teacher with respect to the preparation time required, yet it is troublesome for the ordinary student. Such expositions are often blindly copied from lecturehall-blackboards with very little effect on students' minds. If a student is not turned on by this symbol-loaded presentation, is he or she to be blamed for that?

A presentation via an inductive inquiry

The guided discovery approach deserves serious examination here. Assigning a set of goal-oriented activities in an inductive sequence of problems is characteristic of this approach. [Zaslavsky and Movshovitz-Hadar, 1986] In this case, instruction may take the following form:

> *Introduction*: In the next set of activities you will be studying square matrices with the following properties:
> The rows are arithmetic progressions, all with the same common difference, and the columns are also arithmetic progressions, all with the same common difference. We shall call this "a fixed common difference (c.d.) matrix".

> *Step 1*:
> a. Construct a 2×2 fixed c.d. matrix
> b. For each combination of two matrix elements, add up the elements.
> c. Have you found anything interesting
> d. Examine as many cases of a 2×2 fixed c.d. matrix as you need in order to make a general conjecture. Answer: The sums of the diagonals are identical in *any* 2×2 matrix with the fixed c.d. property.)
> e. Prove or refute your conjecture

This task is quite easy and promising

> *Step 2*. Repeat step 1 for a 3×3 fixed c.d. matrix, and the combinations of three matrix elements.

The task here is much more complicated. In each 3×3 matrix there are eighty-four combinations to be calculated. It takes some careful control to allocate all of them, avoiding repetitions. Six combinations have the same sum. Isolating them from amongst the eighty-four combinations requires non-trivial observation, even if the arithmetic is error-free. Finally, finding their common property, namely finding out that these are the six combinations of three elements each taken from a different row and column, is quite a challenge. In view of the work involved it is very likely that students would be tempted to base a conjecture for a general 3×3 matrix upon one example.

To prove the conjecture it is sufficient to show that the algebraic expressions for the six combinations are all equal. We will return later to the proofs of particular cases and their relations to the general proof

> *Step 3*: Same as step 1 for 4×4 fixed c.d. matrices, and combinations of four matrix elements.

In this case, there are 1820 combinations to study in each 4×4 matrix. In addition, one is expected to notice that 24 of them, that is less than 1.5%, have the same sum. Even if we change the task and suggest that students examine only those 4-tuples which have no two elements in the same row or column, they will be dealing with twenty four combinations. In the case of 5×5 matrices, there are one hundred and twenty 5-tuples of that sort

Need we go further to convince ourselves that the prospect of discovery, however intriguing it may be, does not overcome the prospect of the tedious and time-consuming work involved? On the other hand, if we want to provide students with sufficient ground for a general conjecture about *any n×n* fixed c.d. matrix, can we take the liberty of abandoning the inductive process before this step?

This kind of presentation constitutes a two-level inductive inquiry. One level deals with a particular case, the other deals with the general case. At the particular case level the inquiry aims at solving the problem for a particular value of n: in our case, for example, solving the problem for $n = 3$ by an examination of as many instances of 3×3 fixed c.d. matrices as needed. At the general case level the inquiry aims at solving the problem for any value of *n*. In our case it is the generalization for *any n×n* fixed c.d. matrix, based upon the results obtained at the first level for the 2×2, 3×3, 4×4, 5×5 cases. The two-level generalization in the inductive inquiry process is a major source of difficulty in this kind of presentation.

Before going deeper into a discussion, let us take a look at a second theorem and its presentation.

Theorem 2: A remarkable property of prime numbers

A surprising imposition

Honsberger [1970] tells us that Sundaram's Sieve was invented in 1934 by a young East Indian student, named Sundaram, as an instrument for sifting prime numbers from positive integers. The Sieve consists of the infinite table represented by Table 2.

4	7	10	13	16	19	22	25	.
7	12	17	22	27	32	37	42	
10	17	24	31	38	45	52	59	. .
13	22	31	40	49	58	67	76	
16	27	38	49	60	71	82	93	.
.	.	.	.	.	.	.	.	.

Sundaram's Sieve

Table 2

"The remarkable property of this table is: *If N occurs in the table, then 2N + 1 is not a prime number, if N does not occur in the table, then 2N+1 is a prime number* "[Ibid. p 75]. This is astonishing indeed. Even though the entries in the table have immediately visible additive properties, they do not seem to have anything that ties them with primality, basically a multiplicative property. Moreover, it is well known that the prime numbers are rather irregular; there is no formula that generates all of them; the only well known means for finding primes is the Sieve of Eratosthenes, remembered as an awkward and peculiar algorithm limited to finite sets of integers. It is therefore the case that Sundaram's proposition, unlike that of the first theorem above, needs no special build-up to achieve a surprise. The mere presentation does it, formal and imposed as it is. One feels an inner drive to get to the proof. We will, right after a short discussion of the various presentations.

Discussion of theorem presentations

John Mason *et al* [1985] suggest that mathematical thinking is provoked by a gap between new impressions acting on old views [p. 151]. This obviously happens in the surprising presentations of the two theorems above. In the matrix case, the gap is between the common sum and the large number of combinations, which seems to have nothing in common. In the prime numbers case, it is the gap between a collection of arithmetic progressions having an obvious regularity resulted from their additive property, and the fact that prime numbers are defined by a multiplicative property and are known to have very little regularity. None of the other presentations possess gap-creating potential.

Let us, now, compare expected audience reactions to two other presentations: the Symbolic/Verbal Presentation of the first theorem, and the Surprising Imposition of the second theorem. In both cases declarative statements are forced upon the audience. However, the working in the latter is relatively easy to decode, it is meaningful and surprising. Upon reading or hearing it, one would probably wonder: "How come???" The former, in contrast, requires a lot of decoding to see its meaning. One might end the decoding process wanting to ask: "So what???". Audience reaction might be reversed if we used different presentations of the very same theorems. For example, the reaction to the first theorem, if presented as in the Surprising Exposition section, can raise the question "how come???" On the other hand a different presentation of the second theorem may leave one almost indifferent, with a taste of "so what???" We'll encounter such a presentation in the section: A Bottom-Up Development of the Proof (and of the theorem).

If mathematics teachers agree to give first priority to thought-provoking presentations, priority should be given to the ones causing some kind of surprise. "Any good teacher is a frustrated actor," an old saying says. Mathematical surprises are wonderful for those frustrated actors. The "question-mark look" in the eyes of curious students is such a delight to a teacher that no one ought to miss it due to hasty or careless planning. If not for the motivational merit of surprising presentations, they should be sought for their potential as preventive measures against teacher burnout.

PROOF PRESENTATIONS

Three proofs are presented for the second theorem, then two proofs for the first theorem.

A formal proof

In Table 2 the first row comprises all the terms of the infinite progression beginning with 4, 7, 10 ..... This progression is also used to generate the first column. Succeeding rows are then completed so that each consists of an arithmetic progression, such that the common differences in successive rows are the odd integers 3, 5, 7, 9, 11,... *Sundaram's claim is:* If the number N occurs in this table, then 2N + 1 *is not* a prime number; if N does *not* occur in the table, then 2N + 1 *is* a prime number. Honsberger [ibid., pp. 84-5] proceeds with the proof as follows:

> *Proof:* We begin by finding a formula for the entries in the table. The first number in the nth row is
>
> $4 + (n - 1)3 = 3n + 1$
>
> The common difference of the arithmetic progression comprising the nth row is $2n + 1$; hence the mth number of the nth row is
>
> $3n + 1 + (m - 1)(2n + 1) = (2m + 1)n + m$
>
> Now, if N occurs in the table, then $N = (2m + 1)n + m$ for some pair of integers m and n. Therefore,
>
> $2N + 1 = 2(2m + 1)n + 2m + 1 = (2m + 1)(2n + 1)$
>
> is composite.
>
> Next, we must show that, if N is not in the table, $2N + 1$ is prime; or, equivalently, if $2N + 1$ is not prime, N is in the table. So, suppose $2N + 1 = ab$, where $a, b,$

are integers greater than 1. Since $2N+1$ is odd a and b must both be odd, say

$$a = 2p + 1,\ b = 2q + 1$$

so that

$$2N + 1 = ab = (2p + 1)(2q + 1) = 2p(2p + 1) + 2q + 1$$

and

$$N = (2q + 1)p + q$$

But this means N appears as the qth number of the pth row in the table.

We conclude that N occurs in the infinite table represented by Table 2 if $2N + 1$ is not a prime number

Every step in this proof is clear, Sundaram's sieve is admittedly valid, and yet the manner in which the Indian student came up with his remarkable idea remains altogether mysterious. I suspect many readers feel a bit disappointed after going carefully through this proof for we still have no answer to the question, what do these arithmetic progressions have to do with primality? This proof does not make us any wiser The tension caused by the surprising declarative statement is not relieved

Let us start anew, and see if the following approach untangles the problem

A gap-bridging proof

1 The claim we wish to prove concerns the odd numbers. Let us transform every number N occurring in Table 2 to the corresponding K satisfying $K = 2N + 1$, as shown in Table 3. Consequently, the statement to be proved becomes: K occurs in the infinite table represented by Table 3 if K is not prime

9	15	21	27	33	39	45	51	
15	25	35	45	55	65	75	85	
21	35	49	63	77	91	105	119	
27	45	63	81	99	117	135	153	. .
33	55	77	99	121	143	165	187	.
.						.		.

Sundaram's Sieve Transformed

Table 3

2 As any odd integer is a product of two odd integers, the infinite multiplication table of all pairs of odd integers (Table 4), must contain all primes except 2.

3 By definition, *all* prime numbers (except 2) occur in the first row and column of this table, and *no* prime number occurs elsewhere. In addition, any odd composite integer must occur at least once outside of the first row and column

4 On the other hand, if we omit the first row and the first column of Table 4, the remainder is identical with Table 3 This is because, like any integer-multiplication table, Table 4 is, in fact, a set of row arithmetic progressions with the marginal numbers as their respective common differences

5 We conclude that Table 3 contains *all* odd composite integers and *no* primes. In other words: For any odd integer K if K is prime, then it does not occur in Table 3, and if K does not occur in Table 3, then K is prime Q.E.D

×	1	3	5	7	9	11	13	15	17	...
1	1	3	5	7	9	11	13	15	17	.
3	3	9	15	21	27	33	39	45	51	
5	5	15	25	35	45	55	65	75	85	
7	7	21	35	49	63	77	91	105	119	. .
9	9	27	45	63	81	99	117	135	153	
11	11	33	55	77	99	121	143	165	187	
			.		.			.		

Multiplication table of odd integers

Table 4

Most of the students in my math problem solving class for prospective teachers liked this proof more than the previous one. They found that they could finally make sense of Sundaram's Sieve.

This proof bridges the gap, created by the statement of the theorem, between the arithmetic progressions and prime numbers. The bridging takes place at the multiplication table of odd numbers in step 4 where arithmetic progressions and primes intersect. This proof may be called "responsive" since it responds to the stimulation created by the theorem. In general, responsive proofs usually leave most of the audience with an appreciation of the invention, along with a feeling of becoming wiser.

Another popular approach is that of guided discovery, to which we turn in the next section Diagram 2 at the end of the section compares the various presentation of theorem 2 and its proofs

A bottom-up development of the proof (and of the theorem)
Suppose, now, we have no idea whatsoever about Sundaram's Sieve. Here is the mainline of a sequence of tasks leading gradually to the discovery of the sieve.

The goal At the end of this sequence you will have discovered an algorithm separating all primes from the positive integers.

a Recall the definition of a (natural) prime number and a (natural) composite number

b. What property do all prime numbers except 2 have in common? (Ans : All are odd).

c In view of the previous finding, how can the goal be simplified? (Ans.: In order to separate the primes from the positive integers it is sufficient to separate the odd-primes from the odd-positive integers.)

d. Construct the multiplication table of the odd positive integers up to 17 (Result: See Table 4 above.)

Consider this table as a representative of the infinite table of the products of all pairs of odd integers. Study its properties:

1. What property do all entries in the infinite table have in common? (Ans.: All are odd integers as any odd number is a product of at least one pair of odd numbers.)

2. Where do *all* prime numbers occur in the infinite table? (Ans.: In the first row and column.)

3. Where do *only* composite numbers occur? (Ans.: In the complementary part of the table, that is in all but the first row and column.)

e. If we omit the first row and the first column of the infinite multiplication table of odd positive integers, what kind of integers are left in the reduced table? (Ans.: The reduced table contains *all* composite odd numbers and *only* them.)

f. Restate your findings in terms of a conditional statement:
If an odd integer K occurs in the reduced table, then . .
If an odd integer K does not occur in the reduced table, then

g. Let K designate any odd integer, then $K = 2N + 1$ for some integer N. Transform the reduced table by replacing K with the corresponding N and restate your summary in terms of N. (Ans.: The transformed table coincides with Table 2 and above the statement is Sundaram's: If N occurs in the table, then $2N + 1$ is not a prime and vice-versa.)

h. Based upon the finding in step "g" create a flow chart describing an algorithm by which you can now determine for any given positive integer N whether or not N is prime (Result: For an elaboration of this task see Hadar & Hadass, 1983.)

Clearly, this task-sequence proceeds in a bottom-up fashion, from previous knowledge about prime numbers and odd integers to the discovery of Sundaram's Sieve. It is noteworthy that the theorem is stated at the *end* of the process, at which stage it has already been proven. The sequence, therefore, is constructive. In comparison to the Gap-Bridging Proof, the two proofs are very similar yet they differ in principle. Steps "a" to "f" in this proof parallel steps 2 to 5 in the previous one. However, step 1, coming right after the presentation of the striking theorem and opening the previous proof, comes last, as step "h", in this proof, resulting in the statement of the theorum.

Upon completion of this guided discovery sequence, it is evident that students know something which they didn't know before. Moreover, unlike the situation at the end of the formal proof, most of the students now also believe in it, and are perhaps even happy to have accomplished it. So they *are* wiser. According to Manin [ibid] this is a good approach. But do they *feel* wiser? Do they appreciate the mathematics they have learned? Do they feel that this bit of knowledge is a remarkable one? The new knowledge is so well grounded in previous knowledge, it is so gradually built up through a logical sequence, that it leaves very little room for enthusiasm. Many students tend to take the end products, that is the theorem and the algorithm it yields, almost for granted. The teacher may find it necessary to point out the remarkable achievement, otherwise it might remain underestimated or even totally overlooked.

One may decide to add to the discovery sequence a task such as: "Discuss the importance of this discovery with your friends". As a result, learners may acknowledge the importance of having gained this knowledge. This, however, is very far from the feeling of aesthetic appreciation engendered by the Surprising Imposition presentation followed by the Gap-Bridging Proof.

We now return to the first theorem to examine two more methods of presenting a proof.

A structured deductive proof

A *deductive proof* usually proceeds straightforwardly from the antecedent to the consequent, through a possibly lengthy process employing valid rules of inference. The formal proof of Sundaram's Sieve, above, is a typical example of a deductive proof. Such a process rarely transmits the underlying idea of the proof. This is particularly true of the longer and more complicated proofs.

A *structured proof* proceeds in a top-down fashion, giving an overview of the logic of the proof first, and filling in the details deductively, later on. As argued by Leron [1983] a structured proof is appropriate following any method of theorem-presentation in which the theorem is presented at the outset. A structured proof is more responsive to a student's search for meaning than is a straightforward deductive proof. However, as soon as the parts that need detailed proofs become the main concern, it becomes clear that the problem of meaning remains unresolved. This section demonstrates the problem. A solution is suggested in the next section.

A structured proof for the matrix theorem, presented at the opening of this paper, may look like this:

> Given a square matrix A, as stated in the theorem above, and a set of n elements of it for which
> $r \neq s \Rightarrow i_r \neq i_r;\ j_r \neq j_s\ \ r,s = 1,2,\ldots,n$
> we will prove that:
>
> $$(1)\quad \sum_{k=1}^{n} A_{ik,jk} = \sum_{k=1}^{n} A_{ii}$$
>
> Namely, the sum of any n elements of which no two are in either the same row or column, equals the sum

of the main diagonal elements

(2) We let $C_3 = \sum_{i=1}^{n} A_{i,i}$

Combining (1) and (2) we get the claim of the theorem.

(3) Moreover, we'll show that

$$\sum_{k=1}^{n} A_{i\,j} = n/2\,(A_{i\,1} + A_{n\,n}).$$

Namely, the sum of any *n* elements is dependent solely upon the size of *n* of the matrix, and two diagonally opposite matrix elements

This is the "bird's eye view" of the proof. It is clear and sensible. At this point all that remains to do is to fill in the details of the "worm's eye view". It includes an algebraic manipulation of terms and indices from the left hand term to the right hand term in (1), using the antecedent of the theorem (see A Symbolic Presentation). This can be performed artistically by the teacher alone. Such a proof is usually highly impressive in the eyes of an experienced mathematician because the consequent appears to follow almost magically from the antecedent. However, such a proof creates tension for the inexperienced student who can rarely figure out the meaning behind the symbol manipulations. This tension, unlike that caused by the surprising imposition of a theorem followed by a responsive proof, is not likely to be relieved.

A structured proof reduces the pedagogical problem, inherent in any deductive proof, to the "worms's eye view". When the details are not trivial, structured proofs and fully deductive proofs show similar pedagogical problems of meaning. Is there an alternative? The next section suggests one.

A generic-example assisted proof

Here is an alternative to the structured proof used successfully in a problem-solving class for future high school mathematics teachers.

We confine ourselves to Table 1 which is large enough to be considered a non-specific representative of the general case, yet small enough to serve as a concrete example. Using Pimm's [1983] terminology, Table 1 is a *generic example*.

On a projected transparency of Table 1, we superimpose circles on eight of the entries chosen in an arbitrary manner to represent *any* eight numbers of which no two are in either the same row or column. This means, in particular, that in the first row the circled entry will *not* be a corner one, but say, the fifth entry. We deal here, for example, with the eight entries circled in Table 5.

We slide the circle in the first row from the fifth column to the first. Now there are two elements circled in the first column and none in the fifth column. We slide the circle in the fourth row, from the first column to the fifth. We now have a new set of eight elements as required. (See Table 6.)

10	11	12	13	**14**	15	16	17
19	20	**21**	22	23	24	25	26
28	29	30	31	32	33	**34**	35
37	38	39	40	41	42	43	44
46	47	48	**49**	50	51	52	53
55	56	57	58	59	**60**	61	62
64	65	66	67	68	69	70	**71**
73	**74**	75	76	77	78	79	80

Generic example for proof of first theorem

Table 5

What is the difference between the sum of the new set of eight numbers and the sum of the original set? The moves did not affect the sum, as there is a common difference between any two adjacent elements of the same row, and this common difference is the same for all the rows. That is to say we subtracted and added the same number.

We now deal in a similar fashion with the elements circled in rows two and eight. Consequently, we have a new

10	11	12	13	14	15	16	17
19	20	**21**	22	23	24	25	26
28	29	30	31	32	33	**34**	35
37	38	39	40	**41**	42	43	44
46	47	48	**49**	50	51	52	53
55	56	57	58	59	**60**	61	62
64	65	66	67	68	69	70	**71**
73	**74**	75	76	77	78	79	80

Sum-preserving interchanges of circled elements

Table 6

set of eight elements with the required property. *Two* of the eight elements are now in the main diagonal and their sum remains unchanged. We go on interchanging pairs one by one, until *all* eight elements are in the main diagonal. Evidently the process is sum-preserving. This completes the proof of claim (1) for the generic example. Students observing it as its happens visualize the generalization of the process and assimilate the data. It is, then, a challenge within their reach, to write down a formal proof for the general case phrased in (1) above.

Moreover, the step-by-step transformation of the intitial set of elements into the diagonal elements makes it apparent that the elements in the main diagonal form an arithmetic progression whose common difference equals the sum of the two common differences: that of the rows and that of the columns. This completes the proof that the sum of the eight elements is none other than four times the sum of the first and last elements of the main diagonal as claimed in (3) above. This also explains how and why Stover's Trick works (see A Surprising Exposition). Most students can take it from here and fill in the formal details of the general case independently. This generic-example assisted proof is therefore another kind of gap-bridging proof.

Discussion of proof presentations

Leron [ibid, p. 185] paraphrased Manin's saying [ibid]: "A good *presentation* of a proof is one which makes *the listener* (or reader) wiser". This paper suggests that a good presentation of a proof is one which not just make the learner wiser, but also make the learner *feel* wiser.

Very often, in going through a formal proof, particularly those suffering from the "let us define a function" syndrome [Avital, 1973], the student feels treated shabbily. The origin of the proof remains a mystery and the student is left with a sense of unresolved doubts about it. There is a frustrated feeling of *not* being wise enough, not only not as wise as the person who invented the proof, but not even wise enough to understand how the inventor came up with the idea. The attitude towards mathematics which is encouraged this way is: "I'll never understand it, it is not for me ". Unlike the former, a proof, such as the Gap-Bridging Proof, by giving the feeling of becoming wiser, brings about an appreciation of the ingenuity in mathematics with a sense of "I see! It is quite simple and clever. I want more of it."

As mentioned earlier, a structured proof reduced the pedagogical problem of the meaningful exposition inherent in a formal deductive proof to the inner parts of the proof. When the inner parts are not trivial, structured proofs and fully deductive proof show similar pedagogical problems. A generic-example assisted proof, as demonstrated, has the potential to respond to this difficulty.

The proof of a generic example should not be confused with a fully general proof. It only *suggests* the full proof through a generalizable concrete example. From the purely logical point of view there is no replacement for the formal proof. From the pedagogical point of view, a proof of the generic example can sometimes replace the general proof. How often should it be done? This is a matter for the philosophy of mathematics education, of course.

Discussion

While the two theorems presented in this paper are both very interesting, their various presentations are not equally stimulating. Although the proofs presented are all logically valid, they are not equally responsive to students' intellectual needs.

Six theorem presentations were demonstrated: (1) A Surprising Exposition (of invariant sum of elements in a squared matrix); (2) A Surprising Imposition (of Sundaram's Sieve); (3) A Symbolic Presentation (with lots of sigmas and suffixes); (4) A Verbal Presentation (to explain or to replace the symbolic one); (5) A Bottom-Up Development (of Sundarams's Sieve through a guided discovery); (6) An Inductive Inquiry (of successive size matrices).

Six proof presentations were illustrated: (1) A Formal Proof (as in Honsberger's book); (2) A Gap-Bridging proof (for Sundaram's Sieve); (3) A Structured Proof (for the square-matrix theorem, a top-down approach); (4) A Generic Example Assisted Proof (to accompany a structured proof); (5) A Bottom-Up Development (of Sundaram's Sieve where the proof precedes the theorem); (6) An Inductive Inquiry (where it was just indicated that proofs of particular cases may or may not be generalizable).

The last two are listed as theorem presentations as well as proof presentations for they contain elements of both. Hereinafter, they are referred to as "Guided Discovery (G-D)". Another mixture of theorem and proof presentations will be considered here. This is the method of teaching mathematics theorems which employs a surprising theorem presentation, imposition or exposition, followed by a gap-bridging proof or by a generic-example assisted proof. The two theorem presentations are stimulating. The two proofs respond to the stimulations. The term "Stimulating Responsive (S-R) method" is therefore proposed for this mixture.

"Stimulating Responsive (S-R)" vs. "Guided Discovery (G-D)"

Usually, both the Stimulating Responsive and the Guided Discovery approaches include some independent student work, some conjecturing, and some problem-solving activities. However the two are based on different theories of learning.

Guided Discovery is a *teacher-motivated* learning process. The teacher wants to teach and asks questions which the student is expected to answer in order to accumulate new knowledge. G-D is usually a sequence leading to a goal which the teacher believes is worthwhile. Students do not always perceive the goal in this way. Very often the goal is not clear to the student, and even when it is clear, it is often seen as teacher-imposed. Even if the importance of reaching the goal is discussed, students very often do not feel a need to reach it.

The S-R method is a *student-motivated* learning process. It imitates the question-posing process by which very young children learn new things. We know that a child provoked by the strange or the curious approaches an authority, usually an adult, with a question. Being anxious to get a direct answer any effort to guide him or her towards the answer often causes an impatient reaction on the child's part. Analogously, a student who has a burning question is a student attentive to the answer. The Stimulating Responsive method makes the student be the one who has a query for which he or she requires an answer. In this approach the need to reach the goal is created within presentation itself. For such inner motivation to be aroused, it is imperative that the goal be clear and even personal. Hence, a major difference between the Guided Discovery and the Stimulating Responsive methods is in the degree of inner motivation built into the sequence.

A second difference between the G-D and the S-R methods is in the nature of the surprise involved, and in the stage at which it appears. In G-D the surprise comes, if at all, at the *end* of the process. If cleverly done, this may motivate students towards the *next* activity In the S-R method the surprise comes right at the *beginning* It motivates proof activity by creating a gap, or by presenting a challenge, sometimes in the form of a conflict. Rather than motivating the next activity, it motivates the *present* one.

A third difference is in the nature of the "aha!" effect which both the G-D and the S-R methods may create. The "aha!" at the end of the G-D is an expression of excitement This is usually an excitement about the wisdom inherent in mathematics. In the case of the S-R method the "aha!" at the end is an expression of relief from the tension created by the surprise The relief results from a narrowing down of the gap, or from the satisfaction of the intellectual curiosity raised by the surprise and by the conjectures that followed.

A fourth difference is in the attitudes towards mathematics the two approaches may yield These attitudes were discussed earlier in the Bottom-Up Development section.

There is still another difference. Problems involving combinatorial thinking usually lend themselves naturally to the G-D via inductive inquiry Any problem of this sort is in fact an infinite set of problems, one for each value of *n*, for which we seek a general solution [Hadar and Hadass, 1981a]. Two risks are involved here: (1) The examination of particular cases is usually a very good way of getting at a conjecture or of testing an existing one However this is not always so Recall the discovery path in the matrix theorem presentation via an inductive inquiry It was not sufficiently intriguing to justify the tedious and time-consuming process involved in the examination of each 4×4 or 5×5 matrix, yet the latter was necessary in order to provide students with sufficient grounds for a general conjecture (2) The inductive inquiry process, involving the proof of a series of particular cases taken in an increasing order, may sometimes suggest the general idea of the proof. However, this is not always so The proofs of a few particular cases, often of the smaller values of *n* in the inductive series, may have very little to do with the general proof The proofs of the matrix theorem for the cases of 2×2 and 3×3 matrices (see the section on Presentation Via an Inductive Inquiry above) demonstrate this limitation

An analysis of teaching moves [Hadar and Hadass, 1981b] is sometimes helpful in planning an inductive inquiry sequence so as to avoid these risks. The generic-example assisted proof provides an alternative way It was exemplified by the 8×8 case (Table 1). This case, needless to say, does not belong in the inductive inquiry sequence Nevertheless, it is small enough to serve as a concrete example, yet large enough to be considered a non-specific representative, of the general case. The proof for the 8×8 case (see Generic-Example Assisted Proof, above) is kind of "transparent", one can see the general proof through it because nothing specific to the 8×8 case only enters the proof.

Table 7 gives a brief summary of the differences discussed in this paper

Criterion	Stimulating Responsive	Guided Discovery
Presentation	Stimulating	Assigned, suggested
Inner Motivation	Inherent ("How come?)	Not Necessary ("So what?)
Stage of surprise	Initial	Terminal
Task-motivated	Present one	Next one
Nature of "aha"	Satisfaction, relief	Excitement
Nature of proof	Gap-reducing	Linear, bottom-up
Hints	Generic example (if applicable)	Instances, inductive series (if applicable)
Student's role	Posing questions	Seeking answers
Teacher's role	Provoking and answering	Posing questions
Perception of goal	Student s need	Acceptable
Process of reaching	Responsive	Guided (risks tedium)
State at end	Feeling wiser	Wiser
Resulting attitude toward math	Appreciation of beauty and ingenuity	Acknowledgement of importance

Differences between two methods of presentation
Table 7

Closing remarks about presentation planning

To help us make a difference in what happens in teacher education and in classrooms, Underhill [1986] suggests the following four statements:

A *Knowing* is believing
B *Learning* is developing and altering beliefs.
C. *Teaching* is helping others develop and alter beliefs.
D *Behaviour* is human activity aimed at operationalizing beliefs [p. 16]

If Underhill's statements are accepted, the Stimulating Reponsive method is one of their applications to mathematics The S-R method takes more time and effort in the planning stage. Its planning is somewhat like architectural design, in that it includes considerations of mathematical and pedagogical aesthetics, and of long-term attitudinal effects, as well as operational considerations. However, the time and effort are rewarding, they make a noticable difference for students and teachers alike

The relationship of a mathematician, who creates mathematics, to a mathematics teacher, who conveys mathematics to people, is somewhat like the relation of a music composer to a music performer, or that of a playwrite to actors and director. Artists invest a lot of time and effort in polishing their performances Mathematics teachers, likewise, must persistently seek ways to present mathematics such that it is in harmony with the intellectual and psychological needs of their audience This is as crucial to mathematics theorems and their proofs as the way music is performed is crucial to musical compositions. Therefore presentation planning is worth the time and effort it requires. The Stimulating Responsive presentation captures audience attention by surprise, maintains the motivation by the inner drive it creates, and at the end makes the audience enjoy the "music" and appreciate the "composer".

A few questions yet to be answered

A theorem is like a maze. Its proof is the path from the entrance to the center of the maze In a two dimensional maze, many people go from both ends until the two pencil-

Continued on page 30