7-2020

Gifted and Talented: With Others? Separately? Mathematical Analysis of the Problem

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:

Technical Report: UTEP-CS-20-73

Gifted and Talented: With Others? Separately? Mathematical Analysis of the Problem

Olga Kosheleva[1] and Vladik Kreinovich[2]
[1]Department of Teacher Education
[2]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

Crudely speaking, there are two main suggestions about teaching gifted and talented student: we can move them to a separate class section, or we can mix them with other students. Both options have pluses and minuses. In this paper, we formulate this problem in precise terms, we solve the corresponding mathematical optimization problem, and we come up with a somewhat unexpected optimal solution: mixing, but with an unusual twist.

1 Practical Problem

Two main suggestions. What is the best way to organize teaching that takes into account that in each class, some students are ahead of others – such students are usually called gifted and talented?

There are two main suggestions; see, e.g., [1] and references therein. The first suggestion is to select gifted and talented students into a special class section. This suggestion helps these students study to the best of their potential, without being dragged back by their less successful peers. On the other hand, in this suggestion, the rest of the students are deprived of their leaders who could help them and encourage them to catch up.

The second suggestion is to keep all the students in the same class. The advantage is that students who are currently somewhat behind will be infected by the enthusiasm of the gifted and talented ones and thus, catch up. The disadvantage is that in this suggestion, the learning progress of gifted and talented students is slowed down by the rest of the class, and they may not achieve their full potential.

Which option is better? Usual arguments in favor of each of these options are qualitative, based more on ethics and morality than on mathematical models. These arguments can be, in a nutshell, easily summarized.

On the one hand, we want all people to be well educated. This is not easy, this requires some sacrifices. We all pay taxes to support local schools, parents do a lot of things for their children to allow them to have time and energy for studying, and, yes, students in whom we all have invested so much have a moral obligation to help others.

On the other hand, modern economy no longer needs people who do a dull menial work, this work is now done by machines. Current economy is based on and promoted by creativity. Every person has a talent, and our moral obligation is to find and grow this talent – and not hinder its development.

Both viewpoints sound convincing, and if we only use qualitative moral arguments, we will never come up with a decision of what to do.

What we do in this paper. In this paper, we formulate this problem in precise terms. Then, we can formulate this problem as an optimization problem. We then solve this optimization problem, and we show what is the solution.

2 Analysis of the Problem

Is it genes or is it attitude? Of course, if we strongly believe that, e.g., good math (or any other) abilities are determined by a person's genes, then, yes, we should nurture kids with the corresponding rare combination of genes. Otherwise, they will not grow into successful inventors, and our economy will fall behind economies of other countries, countries that do nurture their gifted and talented kids.

The belief that, e.g., math abilities are largely determined by the genes is very strong in the US. In the US, many students are not good in math – in spite of all the money thrown into education; for those who do not know, much more money per schoolkid than in any other country. Many folks try to explain this by the fact that not everyone is born with mathematical abilities.

We come originally from Russia, so we know better. In the Soviet Union, when a drunken alcoholic would come to a store to buy some cheap wine for his remaining few rubles, this person may not have been steady on his feet, he may not have remembered his home address or today's date – but he always knew exactly how much change he must get for his purchase. In contrast, in the US, when the power sometimes goes out and the cash machines stop working, many teenage salesfolks – who have recently studied math in school – have hard time computing how much change to give.

Clearly, this is not the genes, the difference is in the attitude. In Russia, everyone was required to get a high school diploma, and a common understanding was that every person can learn basic (and even no so basic) math. It did not even occur to people that some folks may be genetically predisposed to not be able to learn school math – and 99.9% did.

Yes, there were exceptions, the most famous exception was the most famous 19 century Russian poet Alexander Pushkin. When in school, he excelled in all disciplines, but mathematics was always his weak point. And it did not improve later in life: he was the most popular poet in Russian, his poems sold like the proverbial hot cakes, he started a literary journal that had thousands of subscribers. With good arithmetic skills, he could have become a millionaire – but, in reality, when he died, the government had to pay several hundred thousands rubles to cover his debts.

We need to nurture attitude. What our arguments lead to is that to make sure that everyone learns all the needed school material, the most important need is to nurture *attitude*.

How can we nurture attitude? Why do kids start loving different subjects? A lot of this comes from the teachers. Teachers are usually very enthusiastic for their subject: if you do not feel some enthusiasm for math, you will not devote your life to teaching kids, year after year, how much is 7 times 8. This is true, by the way, for most professions: you need to love your profession to be a policeman and risk you life catching criminals, you need to love your profession to deal – like medical doctors do – with blood and coughs etc., you need to love your profession to be a farmer, to spend almost all your day plating crops or raising cows far away from civilization – this can go on and on.

So, the teacher tries his or her best. Some students get some enthusiasm from the teacher, but some don't. What usually helps is other students. Let us make a comparison which is very appropriate for this epidemic year: enthusiastic students infect others with their enthusiasm – and, of course, students who lack enthusiasm infect others with their lack of enthusiasm.

As a result of this mutual influence, some classes gain a culture of interest and enthusiasm where a student's success is encouraged – while others gain an opposite culture where straight A students are ridiculed and even harassed as "teacher's pets."

Different cultures can be observed even in prestigious highly competitive institutions where students, by definitions, are what would be called gifted and talented in the US. When in Russia, we studied at Math departments of two highly selective institutions: Vladik studied at St. Petersburg University (ranked second in Russia), while Olga studied at Novosibirsk University (ranked third in Russia, but actually very close, at practically the same level). Students in both places were equally good, studied similar tough subjects equally hard, probably spent the same big amount of time on each subject, but the culture was different.

In St. Petersburg, studying was respected. So, when a student got an A, he would say: I deserve it, because I studied hard – and would even braggingly overstate the number of study hours.

In Novosibirsk, studying was not respected at all. So, when a student got an A, she would say: It was my lucky day, I did not study at all – even though Olga has seen her studying all night long.

How to describe this in precise terms. To describe all this in precise terms, let us consider a simple model – similar to simple models describing how epidemics spread.

For each student, let us describe his/her level of enthusiasm toward a given subject by x. The 0 value corresponds to indifference, positive values correspond to enthusiasm, while negative values correspond to the attitude that dampens the enthusiasm.

On average, each student with enthusiasm level x affect every other student. This effect changes the other student's enthusiasm level, from the original level x' to a new level $f(x', x)$. The average enthusiasm is, honestly, rather small – otherwise, we would not have the problem that we are trying to deal with. So, in most cases, both values x' and x are small. Thus, we can use the usual idea (see, e.g., [3, 4]): expand the expression $f(x, x')$ in Taylor series and keep only linear terms in this expansion. A general linear dependence has the form

$$f(x', x) = a + b \cdot x' + c \cdot x. \tag{1}$$

In the absence of x's enthusiasm, i.e., when $x = 0$, we have $f(x', 0) = x'$. Substituting the expression (1) into this formula, we conclude that $a = 0$ and $b = 1$, i.e., that

$$f(x', x) = x + c \cdot x \tag{2}$$

for some constant c. When we have several students, with levels $x_1, \ldots, x_n$ influencing the student x', we similarly get the new enthusiasm level

$$f(x', x_1, \ldots, x_n) = a + b \cdot x' + c_1 \cdot x_1 + \ldots + c_n \cdot x_n,$$

then $a = 0$, $b = 1$, and

$$f(x', x_1, \ldots, x_n) = x' + c_1 \cdot x_1 + \ldots + c_n \cdot x_n. \tag{3}$$

When only one student in the class has non-zero enthusiasm, i.e., when $x_i \neq 0$ and $x_j = 0$ for all other j, we get $f(x', x_1, \ldots, x_n) = x' + c_i \cdot x_i$, but we should get formula (2). Thus, we should have $c_i = c$ for all i. So, the formula (3) takes the form

$$x' \to f(x', x_1, \ldots, x_n) = x' + c \cdot \sum_{i=1}^{n} x_i. \tag{4}$$

According to this formula, the enthusiasm level of each student in the class is changed by a value proportional to the overall enthusiasm of all the students except for this one. In other words, in a class with enthusiasm levels $x_1, \ldots, x_n$, the new level of enthusiasm x'_i of each student is equal to

$$x'_i = x_i + c \cdot \sum_{j \neq i} x_j. \tag{5}$$

Here,

$$\sum_{j \neq i} x_j = \sum_{j=1}^{n} x_j - x_i,$$

so the formula (5) takes the form

$$x_i' = (1-c)\cdot x_i + c\cdot \sum_{j=1}^{n} x_j. \tag{6}$$

What do we want. As we have mentioned, our main objective is to make sure that every student has a sufficient level of enthusiasm, i.e., that $x_i' \geq x_0$ for all i – i.e., equivalently, that $\min(x_1', \ldots, x_n') \geq x_0$ for some threshold level x_0. In this is achievable for the minimal level x_0, a natural idea is to see if we can reach an even higher level of enthusiasm – i.e., to achieve the largest possible value of $\min(x_1', \ldots, x_n')$.

Thus, we arrive at the following precise formulation of the problem.

3 Precise Formulation of – and Solution to – the Optimization Problem

Precise formulation of the problem. Let us consider the whole population of a big school district. Let us assume that the enthusiasm levels of all N students are $x_1, \ldots, x_N$. Let n be a typical size of the class. Our objective is then to divide all the students – i.e., the set $\{1, \ldots, N\}$ – into N/n disjoint subsets $I_1, \ldots, I_{N/n}$ so that the value $\min(x_1', \ldots, x_N')$ becomes the largest possible, where for each $i \in I_k$, we have:

$$x_i' = (1-c)\cdot x_i + c\cdot \sum_{j\in I_k} x_j. \tag{7}$$

Towards a solution: idea of the proof. Let us denote the optimal level of enthusiasm achieved by each student by x_0. Then, if the original level of enthusiasm was $x_i < x_0$, then for the group I_k containing the i-th student we should have

$$\sum_{j\in I_k} x_j \geq \frac{1}{c}\cdot x_0 - \frac{1-c}{c}\cdot x_i. \tag{8}$$

So, if the class I_k contains two such students with drastically different initial values $x_i < x_{i'}$, then the i-th student will be pushed to the level x_0, but the other student, with the larger initial enthusiasm value $x_{i'}$, will be pushed way beyond x_0. Instead of doing this, we can move this student to another class, with higher initial level of enthusiasm, then its value will be increased only to x_0 – and the resulting un-used influence, instead of concentrating it on the student $x_{i'}$, we can spread it around and lift everyone's enthusiasm level.

So, in each class, we combine students with approximately the same value $x_i < x_0$: students with very low level go into one class, students with a somewhat higher level go into another class, etc.

How about students with $x_i > x_0$? When the value x_j is much smaller than x_0, we need large values of $x_i > x_0$ to compensate. So, we need to reserve such large values x_i for compensating the students with very small values $x_j < x_0$. So, it is optimal to assign largest values $x_i > x_0$ to a group that contain the smallest values $x_j < x_0$, etc. Thus, in each group, the values $x_i > x_0$ are also very close.

Thus, in the resulting solution, each class has:

- several low-enthusiasm students, with approximately the same value $x_i < x_0$, and
- several high-enthusiasm students, with approximately the same value $x_j > x_0$.

The proportion of both types of students – e.g., the proportion h of high-enthusiasm students – must be determined from the fact that for previously low-enthusiasm students, the new level of enthusiasm will be equal to x_0, i.e., that we should have

$$(1-c)\cdot x_i + c\cdot(n\cdot h\cdot x_j + n\cdot(1-h)\cdot x_i) = x_0.$$

From this, we conclude that

$$c\cdot n\cdot h\cdot(x_j - x_i) = x_0 - (1-c)\cdot x_i - c\cdot n\cdot x_i,$$

i.e.,

$$h = \frac{x_0 - (1-c)\cdot x_i - c\cdot n\cdot x_i}{c\cdot n\cdot(x_j - x_i)}. \qquad (9)$$

Resulting solution. Once we select x_0, we do the following:

- First, we match the largest values $x_i > x_0$ with the smallest values $x_j < x_0$, selecting the proportion of high-enthusiasm students according to the formula (9).
- Then, we match the largest of the remaining values $x_i > x_0$ with the smallest of the remaining values $x_j < x_0$, etc.

If at the end, we got some low-enthusiasm students unassigned, this means that we have been too optimistic, the value x_0 must be decreased. Similarly, if at the end, we got some high-enthusiastic students unassigned, this means that can use these remaining students to increase the value x_0.

Thus, we can use bisection (see, e.g., [2]) to find the optimal value x_0:

- In the beginning, all we know is that the largest achievable value x_0 is located somewhere between the smallest values

$$\underline{x}_0^{(0)} = \min_i x_i \text{ and } \overline{x}_0^{(0)} = \max_i x_i.$$

- At the beginning of each step k, we have the values $\underline{x}_0^{(k-1)}$ and $\overline{x}_0^{(k-1)}$ about which we know that x_0 is somewhere in between. Then, we run the above scheme with $x_0 = \dfrac{\underline{x}_0^{(k-1)} + \overline{x}_0^{(k-1)}}{2}$.
- If, as a result, we have several low-enthusiasm students unassigned, this means that we have chosen too large value x_0 for our testing, so we know that the actual value x_0 is smaller than the midpoint. Thus, we know that the optimal value x_0 belongs to the interval $\left[\underline{x}_0^{(k)}, \overline{x}_0^{(k)}\right]$, where
$$\underline{x}_0^{(k)} = \underline{x}_0^{(k-1)} \text{ and } \overline{x}_0^{(k)} = \frac{\underline{x}_0^{(k-1)} + \overline{x}_0^{(k-1)}}{2}.$$
- If, as a result, we have several high-enthusiasm students unassigned, this means that we have chosen too small value x_0 for our testing, so we know that the actual value x_0 is larger than the midpoint. Thus, we know that the optimal value x_0 belongs to the interval $\left[\underline{x}_0^{(k)}, \overline{x}_0^{(k)}\right]$, where
$$\underline{x}_0^{(k)} = \frac{\underline{x}_0^{(k-1)} + \overline{x}_0^{(k-1)}}{2} \text{ and } \overline{x}_0^{(k)} = \overline{x}_0^{(k-1)}.$$
- At each step, the width of the interval decreases by a factor of two. We stop when the width $\overline{x}_0^{(k)} - \underline{x}_0^{(k)}$ of the interval $\left[\underline{x}_0^{(k)}, \overline{x}_0^{(k)}\right]$ containing the optimal value x_0 becomes smaller than some pre-determined accuracy ε. In this case, any point from the interval $\left[\underline{x}_0^{(k)}, \overline{x}_0^{(k)}\right]$ approximates the optimal value x_0 with the desired accuracy ε.

How do we know it works? The result seems somewhat counterintuitive – we did not any of the two usual options, we got a completely new arrangement. How do we know it works?

One argument is based on our experience of students in Russia. In each class, the best-performing student (straight-A student, in US terms) was teamed with the worst-performing student to help and advise – and it worked well!

Another argument, also accidentally from Russian experience. One of our colleagues, Dr. Andrei Finkelstein, actively participated in the educational TV program aimed at middle-school and high-school kids, especially those who need help in addition to what they learn at school. He witnessed many discussions and plans, and what always surprised him is why when have a renowned scientist explaining something very basic, e.g., the three Newton's laws of motion, it led to much better resulted that when a similar TV lesson was taught by a professional teacher. From our viewpoint, this is exactly what we mentioned: that the most important thing is not so much the transfer of knowledge, but rather the transfer of enthusiasm:

- If there is no enthusiasm, even the best teacher cannot force the knowledge into the students' minds.

- However, if there is an enthusiasm, then the students will learn even by themselves, there is less of a need of a pedagogically skilled teacher.

And this is the perfect example of our seemingly counterintuitive idea – we match the most enthused physicist (a renowned person has to be very enthusiastic) with the least enthused students – who are so low-enthused that they need extra help.

What about slowing down? Shall not the presence of low-enthused (and thus, low-performance) students in the class drag down the gifted and talented ones? It would if they were given the same tasks, but this is *not* what we propose.

Our only purpose of bringing low-enthused and high-enthused folks together is to raise the students enthusiasm – and thus, make them want to learn. Of course, the tasks for the two groups of students will be somewhat different – and good teachers do it already in classes, by giving interested students additional tasks and/or additional parts of the task.

But would not low-performing students feel bad? They probably would if they were low-performing in all classes, but, as we mentioned earlier, we share the usual belief that everyone has a special talent.

Pushkin may been a low-performer in math, but in humanity courses, he would have been on top. One of us (Vladik) was high-performer in math, but, e.g., on PE (Physical Education), in the opinion of his PE teacher, he was performing on the 3-year-old level. In Art, the teacher have him a completely un-deserved B for his (honestly, awful) picture of a horse, just because she realized (and saw) that he tried very hard.

This is normal, everyone will be in a high-performing group in some disciplines and in a low-performing group in others, so what's to feel bad about?

Acknowledgments

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

The authors are greatly thankful to Branka Radulovic for her encouragement and help.

References

[1] S. Assouline and A. Lupkowski-Shoplik, *Developing Math Talent: A Guide for Educating Gifted And Advanced Learners in Math*, Prufrock Press, Waco, TX, 2005.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, MA, 2009.

[3] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[4] K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2017.