Who Says We R_0 Ready for Change?

Nicolás M. Crisosto

UC Berkeley

Co-Authors
Carlos Castillo-Chávez

Department of Biometrics, Cornell University

Christopher Kribs-Zaleta

Departments of Mathematics and Curriculum & Instruction, University of Texas at Arlington

Stephen Wirkus

Department of Mathematics and Statistics, California State Polytechnic University, Pomona

August 2001

Abstract

This paper uses a theoretical model of cooperative learning to examine the survivability of communities of learners. A primary assumption in this article is that individual-based learning tends to use local information to increase an individual's "fitness" while collaborative learning, based on the sharing of information, knowledge, and resources, increases group fitness. We frame our discussion about the importance of cooperative learning at the community and individual level using theories of intellectual development based on the views of Vygotsky ("individualistic") and Piaget ("social"), and mediated by concepts and ideas from the fields of epidemiology and evolutionary biology. Our approach is motivated by evolutionary biology metaphors and modeled via epidemiological (contact) processes. Furthermore, using a simple cooperative learning model, we address the belief that sharp community thresholds characterize separate learning cultures such that one must cross a tipping point to move from one culture to the other. Our results allow us to discuss the impact of individual learning on community intellectual development and its resilience to change.

Introduction

We frame our discussion about the importance of cooperative learning by providing an elementary presentation of Confrey's examination (1994, Feb. 1995, June 1995) of theories of intellectual development based on the views of Vygotsky ("individualistic") and Piaget ("social"). Basically, Vygotsky believed that development begins at the social level and moves towards individual internalization, while Piaget held that development proceeds from the individual to the social world. Confrey (1994, Feb. 1995, June 1995) tries to bridge these different philosophies (highly simplified by the above dichotomous characterization) through "the evolutionary biology metaphor", which incorporates environmental concerns. This reduced version of Confrey's metaphor of intellectual development assumes that learning is driven by individual and social forces. It assumes that learning takes place in a landscape that is not independent of the nature of individual and social (group or community) interactions. Finally, Confrey's theory implicitly assumes that a scale of observation has been chosen so that the nature of each specific question determines the unit of selection and level of inquiry.

The evolutionary biology metaphor used in this paper implicitly assumes a unit (e.g., the individual) and a level (e.g., the classroom). Hence, it is suitable for the study of processes like learning at various levels (e.g., the individual, group or community level). Since we are interested in the process of learning at multiple levels, it is important to develop ways of extrapolating our understanding of learning from one scale to another (e.g., impact of individual learning on community knowledge). Accordingly, the development of multiscale approaches is critical for our understanding of the mechanisms behind intellectual development. As such, the impact of learning on the intellectual landscape is based on the fact that learning is not a unidirectional process. Namely, local changes may impact the structure of the overall intellectual landscape and vice versa (coevolution). Here, we focus on the impact of cooperative learning on community knowledge and, consequently, the influence that individuals have on the community learning landscape is important.

Our approach is based on evolutionary biology metaphors and modeled by epidemiological (contact) processes. The underlying assumption (using evolutionary jargon) is that individual-based learning uses local information to increase an individual's "fitness" while collaborative learning shares community knowledge and support. Hence, on the average, collaborative learning will increase the mean fitness of the group and possibly the fitness of most individuals in the group. We also consider the "tipping point" concept (see Gladwell, 2000), which is the belief that sharp community thresholds characterize separate learning cultures. This idea asserts that one must cross a tipping point to move from one culture to the other and, in particular, the crossing of a threshold substantially alters the learning structure of a community.

Metaphors will be derived that enhance the importance of cooperative learning through the introduction and analysis of a simple mathematical model for cooperative learning. The analysis of this model focuses on the impact that cooperative learning has on the resilience of a community learning landscape. The paper is organized as follows: Section 1 introduces a working definition of cooperative learning and briefly reviews some of the literature on this subject; Section 2 introduces a simple model for cooperative learning; Section 3 presents some of the analysis of the model and discusses its consequences; Section 4 outlines the conclusions

of our analysis and discusses some extensions; and Section 5 describes an example application of the model. We close with some conclusions.

1 Cooperative Learning: a brief overview

Norwood (1995) provides the following definition of cooperative learning:

> "Cooperative learning is a set of instructional strategies which bring students of all performance levels together to work in small, mixed-ability learning groups...for problem solving experiences. The students in these groups are not only responsible for learning the material being taught in class, but also for helping their group members learn the material."

Our use of the concept of cooperative learning will follow Norwood's definition. It is important to note that the development of an environment that is conducive to, and supportive of (for a long period of time), cooperative learning may be difficult at first (Whicker et al., 1997). Here, it is assumed that cooperative learning, wherever and whenever it takes place, is carried out in the context of Norwood's definition. We hope to show that cooperative learning brings students together into situations where the impact of individual learning on others is high enough that a strong culture of learning is established (community intellectual resilience).

There is plenty of data that demonstrates the positive impact that cooperative learning can have on a community of individuals. In fact, it is known that cooperative learning promotes achievement as well as other positive affective outcomes at the elementary and middle grade levels (Whicker et al., 1997). Students also benefit from improved social skills by working in groups (Slavin, 1984, 1987). Working together, students learn to be tactful, to manage conflicts effectively and to respect the opinions of others (Whicker et al., 1997; Augustine et al., 1990). In particular, as collaborative skills become of increasing importance in social and professional life, students' academic experiences are of more value when they include exercises in cooperative efforts. As Brown et al. (1989) point out:

> "Students who are taught individually rather than collaboratively can fail to develop skills needed for collaborative work. In the collaborative conditions of the workplace, knowing how to work collaboratively is increasingly important. If people are going to learn and work in conjunction with others, they must be given the situated opportunities to develop those skills."

Of course, the fact that we live in a global economy has re-emphasized the importance of addressing scientific, economic, ecological, health and environmental problems on larger scales. Problems tend to require multidisciplinary skills—the type of skills that can be found on a team. Collaborative work is fundamental to many of the problems faced by society today. Where will these interdisciplinary teams come from? In the U.S., these teams will have to include a larger proportion of minority students, but current inequities in the U.S. educational system exclude their full participation.

Minority students in urban school districts face dropout rates nearly twice the national average (Education Week, 1998) and are disproportionately attending the nation's high

schools with the weakest promoting power (correlated with high dropout rates) (Balfanz, 2001). Educational strategies such as cooperative learning, with the capacity to raise the achievement level of groups of students as well as individuals, are vital because cooperative learning enhances the education of all. Empirical evidence shows that students who are doing well academically and who participate in the educational enterprise as peer tutors learn more than those who do not (Ginsburg-Block & Fantuzzo, 1998). Even the simplest forms of collaborative learning improve the intellectual growth of individuals (e.g., being exposed to high-level conversations of mathematics during group work leads to higher gains in learning (Cossey, 1997)). Thus, cooperative learning affords increased intellectual development to the individual members as well as the community as a whole.
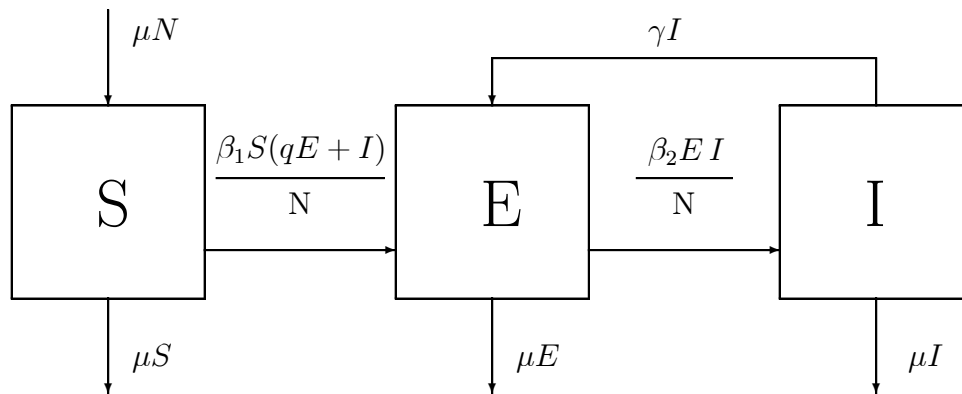
2 Cooperative Learning Model

Figure 1: Flow chart for cooperative learning model

$S(t)$	population of least cooperative individuals at time t
$E(t)$	population of moderately cooperative individuals at time t
$I(t)$	population of highly cooperative individuals at time t
$N(t)$	total population (constant)
β_1	the maximal effective I peer pressure rate by I on S
q	proportion of E class individuals who interact with the S population
$\beta_1 q$	the maximal effective E peer pressure rate by E on S
$1/\mu$	the average life span of an S, E, and I individual
β_2	the maximal effective I peer pressure rate by I on E
γ	the rate of loss of cooperation (energy and leadership) by I class

Table 1: Population variables and parameter definitions

Keeping in mind our working definition of cooperative learning, we consider the model given in Figure 1, with the parameters as listed in Table 1. The dynamics of cooperation in this model are assumed to be driven by peer pressure type interactions between individuals who have been immersed in a collaborative learning environment. The μN gives the rate at

which individuals enter the system (proportional to the total population N). The population is subdivided into three classes: S denotes those individuals who have agreed to be part of a cooperative learning environment; E denotes those who have learned the value of cooperation and use it to "convert" others; I denotes the class of leaders, that is, the group of individuals who have taken on their shoulders the spread and survival of cooperative behavior. Then $N = S + E + I$, and if we let μS, μE, and μI represent the exit rates from the system by group membership, the recruitment and departure rates from the system are equal, that is, the total population N is constant. Hence, demography will not play a role in our simple model.

The movement between classes is governed by the frequency of each type and the conversion rates β_1 and β_2. S individuals become E individuals through their interactions with E and I individuals; the interactions depend on their frequencies, $\frac{E}{N}$ and $\frac{I}{N}$. Here we assume that E interactions are less effective than I interactions in teaching S individuals to be more cooperative, by a factor of q (where $0 \le q \le 1$), which is why the impact of E on S is given by $\frac{qE}{N}$. Since the interactions are multiplicative, the process is nonlinear. As such, β_1 multiplied by $\frac{(qE+I)}{N}$ represents the combined peer pressure on S individuals by the I and E classes. To move from the E class to the I class, E individuals respond to the influence of the I class population, that is, $\beta_2 \frac{I}{N}$ represents the I peer pressure exerted on E individuals to join the elite of the system. The rate γI describes the loss of cooperation (energy and leadership) in the I class. The box-flow diagram describes explicitly the rates of flows between each class. The specific set of nonlinear differential equations that describes this system and its analysis can be found in the appendices (beginning with Model A, the special case where $\gamma = 0$).

3   Cooperative Learning Model Outcomes

We are primarily interested in the study of all possible qualitative outcomes of our cooperative learning model, that is, the intellectual landscapes in which our population under study might end up. For this reason, the model's transient dynamics are ignored. It is assumed that the system is at equilibrium, remaining in the same steady or equilibrium state for some long period of time. We focus on the changes in the nature of these steady states as parameters (degree of peer pressure, etc.) are varied. The current state of a system like this typically depends on two non-independent factors: initial conditions and thresholds. In this section we describe first the possible end states, and then the threshold quantities involved. The next section will describe how and when the initial conditions come into play.

The key characteristics of this model for cooperative learning are that there exist five possible outcomes for the population over time, depending on the values of the initial conditions and thresholds described below (see Appendices A-F for the mathematical analysis). One of the following must occur:

(A) All the population remains a part of the least cooperative group (the "novices").

(B) The population divides between the least cooperative and moderately cooperative groups.

(C) The population distributes between all three classes.

(D) Either the population distributes among all three classes or the entire population remains in the novice group, depending on the initial numbers of cooperative individuals introduced into the population.

(E) The population distributes either among all three classes or only between the least and moderately cooperative groups, depending on the initial number of highly cooperative individuals introduced into the population.

There are two basic threshold parameters associated with the outcomes of this model. The first is the *low-peer pressure threshold value*, R_d, which measures the average effectiveness of E peer pressure on S individuals, the ability of the E class to establish the cooperative learning environment. R_d is given by the product of the E on S peer pressure, times E effectiveness q, times the average residence time of an individual in the system. This interpretation assumes that the system has no I individuals and that the population is composed mostly of S individuals. Intuitively, if R_d is greater than one then the E class grows by recruiting S individuals. In epidemiological terms, E individuals have managed to successfully invade the S population on their own. Here, all but a fraction $1/R_d$ of the population will become successful cooperative learners.

The second threshold parameter is the *high-peer pressure threshold value*, R_0, which measures the ability of the I class to establish itself by recruiting from the E class, given the size of the latter determined by R_d. R_0 is given by the product of the I-on-E peer pressure times the proportion $(1 - 1/R_d)$ of the population already in the E class, times the average residence time of an individual in the system. This interpretation assumes that the system has few I individuals, and that the population is composed mostly of S and E individuals (with $R_d > 1$). Intuitively, if R_0 is greater than one, then the E and I classes will grow by recruiting from the S and E populations, respectively. In terms of our biological metaphor, I individuals would then have managed to successfully invade the S and E populations.

However, the situation here is not as simple as a "tipping point", and Gladwell's (2000) approach is not enough to explain the cooperative dynamics of our model. Indeed, initial conditions play a critical role on the outcomes, and as parameters are varied it will be often observed that community learning is a process that is hard to undermine once established. Furthermore, it will be seen that investing in cooperative learning is sound because it creates communities of learners that can survive in situations where they might not naturally arise (community resilience).

4   Interpretation of Mathematical Analysis

Our goal has been to analyze the behavior of this theoretical model for peer-driven interactions and the general idea of cooperative learning. Individuals begin in the S class, being part of the least contributing group. Over time, through the efforts of the I class and to a lesser extent of the E class, a portion of the S class becomes more proficient and cooperative, resulting in their promotion to the E class. In addition to interacting with the S class, E class individuals also develop through the presence of the I class and eventually enter the I class, where as experienced cooperative learners they take an active part in supporting and fostering the learning community.

As the mathematical analysis (see appendices) indicates, this model exhibits a phenomenon known as a backward bifurcation, which indicates that under certain conditions two different end states coexist, and initial conditions determine which of the two actually occurs. This is significant here because of the implication that, under those conditions, although the *high-peer pressure threshold value*, R_0, may be less than one, high levels of cooperation can still be maintained if enough I class individuals are introduced into the system. So while there is a tipping point (where $R_0 = 1$), dependent on the parameter conditions, beyond which I class individuals are able to maintain a constant, sustainable influence on the system, in this case the initial proportion of I class individuals in the system can be as important as the working conditions, or more so. In other words, there are learning conditions for which the natural tipping point has not been reached, but under which, once there has been an initial investment of resources to introduce a sufficient proportion of highly cooperative individuals, their efforts along with those of the E class combine and reach a turning point, overcoming working conditions that would normally predict failure (as measured by discontinuing the promotion of participants into the highly cooperative class). It also means that once a successful level has been reached, destroying the system is difficult.

In fact, analysis shows that it can even be possible to create a sustainable cooperative learning environment involving all three classes of individuals under conditions where normally not even the intermediate E class would be able to sustain itself ($R_d < 1$). This even more remarkable contrast, again dependent on how many experienced cooperative learners are brought into the system at the beginning, leads us to examine more closely the conditions under which such a dichotomy is possible (outcomes D and E above). If the I mentor class is not able to sustain itself through recruitment from an E class sustained by S-E interaction ($R_0 < 1$), but it *is* capable of sustaining the E class by working with novice S individuals ($R_d > q$), then the relevant condition (inequality (16) in Appendix D), under which a large enough infusion of experienced cooperative learners results in a stable learning environment, requires that the mentor class be able to interact enough with novices that the size of the intermediate E class grows into a pool large enough to generate new I class mentors before the current generation leaves the system. Here it is precisely simultaneous cooperation at every level which can sustain the community: the S-E interaction alone is not strong enough to produce an intermediate level of learners large enough to generate mentors, but the mentors' added interactions with the novices increases the intermediate group enough for new mentors to arise. In order for this teamwork to succeed, however, there must already be enough mentors in the system to foster growth; without them, that crucial added boost to the novices is missing. This requires that the learning community have enough support to bring in enough expert learners.

To further investigate the nature of this turning point we consider two questions: (1) What happens if highly cooperative individuals become less cooperative before exiting the system, i.e., what if there is backsliding from being highly cooperative to being moderately cooperative? and (2) What happens if mid-level individuals do not positively interact with the novices, i.e., if the motivated individuals do all the work?

Model B, analyzed in Appendix E, responds to the first question by considering the effect of the parameter γ, defined previously as the rate at which the energy of the I class is lost, as master cooperative learners give up their leadership roles. What we find is that as I class individuals lose their "spark", it becomes nearly impossible to reach a turning point. The

requirements for a synergistic victory over the working conditions are so strict that the initial investment of resources required is likely to be too high, not benefiting the system in the long run. Therefore, it is important to continually support motivated and successful individuals in cooperative learning situations; otherwise a program could weaken dramatically. Also, we see that providing long term resources aimed at maintaining the effectiveness of the I class can be essential, and in the long run less costly and more effective than having to periodically rescue a system in crisis.

Model C, analyzed in Appendix F, looks at the second question by setting the E class's contribution q to zero. The results are clear and provide information about the interactions within the entire system. If there are not interactions among all the classes, i.e., if, in addition to learning from the leaders in the learning community, moderately cooperative individuals are not encouraged to work with and mentor the novice individuals, then there is no natural turning point. In essence, the system is not a cooperative learning environment without everyone interacting with everyone else. Even introducing high proportions of motivated individuals will only generate a sustainable cooperative environment if the efficiencies of both transitions are exceptionally high.

5   An Example

Our original motivation for studying cooperative learning has been the tremendous success of the Mathematical and Theoretical Biology Institute (MTBI), which is directed by one of the authors (CCC), and with which the remaining authors help at various levels. MTBI is a summer research program for undergraduates—primarily those from underrepresented backgrounds, including Latinos, Native Americans, African-Americans, and women of all groups. Typically, 24-28 new students from various socioeconomic and academic backgrounds participate each year and 6-8 also return from the previous year. The students are housed together on the Cornell University campus for the duration of the eight week program. Although the overall goal of MTBI is to increase the number of underrepresented minorities pursuing Ph.D.'s in mathematics and other science fields, giving them a solid research experience is also important. To this end, all of the research (students work in groups of 3-4 students supervised by 1-2 faculty) is on the frontier of research in the fields of epidemiology, and is of high enough quality that it is published in Cornell University's Biometrics Department Technical Reports. Moreover, many of our students (who are typically from non-research institutions) have gone on to pursue Ph.D.'s in math or the sciences and are doing quite well.

In order to reach the frontier of epidemiology in four weeks, the students attend courses in discrete, continuous, and stochastic dynamical systems in the morning, computer laboratory courses introducing them to necessary computational software in the afternoon, and are given homework to do in the evening which covers the material learned that day. The amount of homework given is too much for the vast majority of students to complete in the allotted time if they work alone. Initially, we constantly work against the students' misconception that "they must be able to solve a problem on their own before they can contribute to the group" (Frankenstein, 1987). The students returning from the previous year help facilitate communication among the students, help answer homework questions,

and continually encourage the students to help each other with their work. A large common area is set up which is conducive to working together. It usually takes about a week of fostering on the part of the faculty and returning students before the majority of students realize not only that cooperative learning is the only way to accomplish everything that is required, but that each and every individual can actually learn more through collaboration.

In terms of the mathematical model derived in this paper, we assume that all new students enter as individuals in the S class. The students returning from the previous year are selectively chosen by two of the authors (CCC and SW) and all belong to either the E or I classes. In a given year, there are typically a few new students who also belong to the E class, and these quickly move into the I class for the remainder of the program. But there are also S individuals that move first to E and then to I. In any case, most of the students end up in the E class by the end of the second week (some still remaining in the S class and others having moved to the I class). It is also essential that this happens because the returning students begin intense work on their research projects at that point (previously having only attended their own lectures and done homework) and are no longer able to devote nearly as much time to the facilitation of cooperative learning among the new students.

Before going into more detail, it is important to quote Confrey (June 1995):

> "The introduction of emotional intelligence into discussions of mathematics education allows one to assert that both facilitating and debilitating emotions play a significant role in learning, and that emotional qualities of classroom interactions will exert a significant influence on what is learned."

To this end, we have weekly group meetings in which students share about themselves and express concerns that they may have. We realize that emotion enters into learning, and we do everything we can to make a positive cooperative atmosphere a reality. We often try to have a few of the returning female students talk with the new female students so they may feel more comfortable expressing their concerns, and then those concerns are addressed by the faculty and returning students.

These meetings help make a smooth transition between the classroom part of MTBI and the research part of MTBI. After four weeks, as mentioned, the vast majority of students are in the E and I classes. We then have them divide into groups of 3-4 and choose a research project that interests them. We again quote Confrey (June 1995) for her insightful comments on elimination of the oppressive view of abstraction:

> "Allowing mathematics to continue to require students to disengage from their personal sources of experience and to learn a system of rituals that makes little sense to them but which will admit them to the ranks of the elite is one of the most effective ways of maintaining this oppression."

As mentioned previously, one of the main goals of MTBI is to increase the number of underrepresented minorities pursuing Ph.D.'s and this can only be done (in our opinion) by making mathematics exciting and appealing to the students. We thus let them choose their own research project in the field of mathematical biology. Whether they have an interest in HIV, tuberculosis, or education, they use techniques learned during the first four weeks to model a problem that they find interesting. The quality work that results is amazing, and

it is their excitement about the topic that makes it all worthwhile. Yes, they get frustrated many times (as is part of research) with the lack of data in the literature, the difficulty of the mathematics that they are using, etc. But with the help of the faculty, they are encouraged to keep the larger goal in mind. As the problems the students work on are truly unique, the suggestion of Frankenstein (1987) that teachers and students must truly be co-researchers is heeded. Perhaps most importantly, the idea and practice of cooperative learning is continually reinforced during the research period. Students' freedom to research topics meaningful to them and collaboration with faculty and other students alike, as peers, serve to stop the intellectually oppressive view of abstraction described above.

At the end of the eight week program, the students present their work to the Cornell University academic community in a colloquium. Most leave feeling good about their work, good about math, and good about the importance of working with others. This is evidenced in the number of students that are inspired to continue on in mathematics and the sciences, who expressed to us that they had no idea they could accomplish so much in such a short period of time. As one student said from the summer of 2000:

> "Before I came to MTBI, I wasn't seriously considering grad school. I was getting discouraged from feeling overwhelmed at school and had lost my passion for learning for the sake of learning...working and learning at MTBI inspired in me a renewed sense of self confidence and desire to continue my education."

This is a successful cooperative learning environment which thrives under conditions—a short span of time and high workload—where it would not arise without its substantial returning core of experienced cooperative learners, working alongside the rest.

6   Conclusion

Essential for students to attain the academic standards expected of them are the basic elements of the educational process: available and sufficient materials and quality instruction. Yet many of today's students are expected to achieve given overcrowded classrooms, lack of textbooks and other study guides, and at times with a teacher who does not have expertise in the given subject matter. One proposed solution is the effort underway to reduce the teacher-student ratio in several states, but the costs are high and the process slow. Moreover, the associated battle against teacher shortages does little to ensure that recruits will have developed the skills necessary to be effective educators. In this context, understanding cooperative learning as an educational strategy may provide an additional course of action.

The model for cooperative learning we have considered above is only a simple sketch of a complex human process — learning — and its value lies in the insights gained from examining the role of structure and hierarchy in the learning process. Analysis of our model shows that minimal investments in cooperative learning can establish a small, resilient community of learners. We see that distributing, rather than compartmentalizing, the spread of knowledge may allow a community of learners to succeed despite adverse conditions. We believe one reason for this effect is that cooperative learning provides teachers and students an opportunity to work together, distributing academic decision making. In this way, cooperative learning promotes the deconstruction of pedagogical power structures by encouraging

teachers and students to share and even trade roles in the educational process. This benefits both students and educators by unlocking and focusing the creativity and motivation of students.

As the model demonstrates, a key characteristic of the cooperative learning environment is the positive interaction among all the students, especially between the novice and intermediate students. In this model, rigidly hierarchical educational systems preclude the second turning point described above — where the successful establishment of a cooperative learning environment depends on the proportion of motivated students initially participating, instead of on purely environmental conditions — thereby converting resilient community structures into fragile constructs. In contrast to a hierarchical learning environment, within an atmosphere of cooperative learning, students' individual learning and enthusiasm have an opportunity to contribute positively to the establishment of a strong culture of learning. In particular, the multi-level interactions make the effects of successful investment in cooperative learning difficult to destroy, adding to the community's intellectual resilience.

While such learning communities are able to withstand changes in some of the conditions, we also find that what affects community intellectual resilience most is the loss of the highly motivated class's cooperation and lowering the proportion of moderately cooperative individuals participation. Losses in the energy and leadership of the highly motivated students causes the cooperative learning environment to be weakened against changes in the learning conditions. In other words, if the highly motivated students interact less within the cooperative learning environment, pressures like lack of resources or the teacher's time are more likely to break down the cooperative learning environment. Similarly, too few moderately cooperative students participating brings the cooperative learning environment closer to becoming a hierarchical learning situation which is more dependent on the learning conditions.

In conclusion, with so many studies showing that cooperative learning has a positive impact on the academic achievement of students, by analyzing a simple mathematical model, we have examined the characteristics of cooperative learning environments in order to better understand what may make this educational strategy effective. Inspired by cooperative learning seen every summer in MTBI, we proposed a model of cooperative learning using epidemiological ideas. We showed that under the right conditions a group of cooperative learners can be established, and examined its resilience to change. We interpreted this to indicate that if one wants to establish cooperative learning, it is necessary to invest the materials (both human and capital) necessary to achieve a cooperative atmosphere, and that, although maintenance is necessary, maintaining this atmosphere is not nearly as difficult as achieving it in the first place.

References

[1] Augustine, D. K., Gruber, K. D., and Hanson, L.R., Cooperation Works! *Educational Leadership* 47: 4–7, 1990.

[2] Balfanz, R. and Legters, N., How Many Central City High Schools Have A Severe Dropout Problem, Where Are They Located, and Who Attends Them? Initial Estimates Using the Common Core of Data. Center for Social Organizations of Schools at Johns Hopkins University, 2001.

[3] Bianchini, J. A., Where Knowledge Construction, Equity, and Context Intersect: Student Learning of Science in Small Groups. *Journal of Research in Science Teaching* 34 (10): 1039–1065, 1997.

[4] Brown, J.S., Collins, A., and Duguid, P., Situated Cognition and the Culture of Learning. *Educational Researcher* 18: 32–42, 1989.

[5] Confrey, J., A Theory of Intellectual Development, Part 1. *For the Learning of Mathematics* 14 (3): 2–8, November 1994.

[6] Confrey, J., A Theory of Intellectual Development, Part 2. *For the Learning of Mathematics* 15 (1): 38–48, February 1995.

[7] Confrey, J., A Theory of Intellectual Development, Part 3. *For the Learning of Mathematics* 15 (2): 36–45, June 1995.

[8] Cossey, R., Mathematical Communication: Issues of Access and Equity. Unpublished dissertation, Stanford University, 1997.

[9] Frankenstein, M., Critical mathematics education: an application of Paulo Freire's epistemology, in *Freire for the classroom: a sourcebook for liberatory teaching*, I. Shor, ed., Portsmouth, NH: Heinemann, 1987. pp. 181–209.

[10] Ginsburg-Block, M. D., and Fantuzzo, J. W., An Evaluation of the Relative Effectiveness of NCTM Standards-Based Interventions for Low-Achieving Urban Elementary Students. *Journal of Educational Psychology* 90 (3): 560–569, 1998.

[11] Gladwell, M., *The Tipping Point*, New York: Little, Brown & Co., 2000.

[12] Norwood, K. S., The Effects of the Use of Problem Solving and Cooperative Learning on the Mathematics Achievement of Underprepared College Freshmen. *Primus* 5 (3): 229–252, 1995.

[13] Orfield, G. and Gordon, N., Schools More Separate: Consequences of a Decade of Resegregation. The Civil Rights Project at Harvard University, 2001.

[14] Roth, W.M., Interactional Structures during a Grade 4-5 Open-Design Engineering Unit. *Journal of Research in Science Teaching* 34 (3): 271–302, 1997.

[15] Slavin, R. E., Students Motivating Students to Excel: Cooperative Incentives, Cooperative Tasks, and Student Achievement. *The Elementary School Journal* 85 (1): 53–63, 1984.

[16] Slavin, R. E., Cooperative Learning and the Cooperative School. *Educational Leadership* 44 (11): 7–13, 1987.

[17] Smith, E., Haarer, S., and Confrey, J., Seeking Diversity in Mathematics Education: Mathematical Modeling in the Practice of Biologists and Mathematicians. *Science & Education* 6: 441–472, 1997.

[18] Whicker, K. M., Bol, L., and Nunnery, J. A., Cooperative Learning in the Secondary Mathematics Classroom. *The Journal of Educational Research* 91 (1): 42–48, 1997.

A   Model and Equations

From the model described in Figure 1 and Section 2, we derive the following equations, where variables and parameters are as given in Table 1:

$$\frac{dS}{dt} = \mu N - \beta_1 S \left(\frac{qE + I}{N}\right) - \mu S,\tag{1}$$

$$\frac{dE}{dt} = \beta_1 S \left(\frac{qE + I}{N}\right) - \beta_2 E \left(\frac{I}{N}\right) - \mu E + \gamma I,\tag{2}$$

$$\frac{dI}{dt} = \beta_2 E \left(\frac{I}{N}\right) - \mu I - \gamma I,\tag{3}$$

where $N = S + E + I$ is the total population at time t.

By adding (**??**), (**??**), and (**??**) we get that $\frac{dN}{dt} = 0$ which shows that the total population is constant. Hence, without loss of generality, we set $N = 1$ and consider the reduced model via the substitution $S = 1 - E - I$.

B   Model A: $\gamma = 0$

We begin by solving for the equilibria using the two-dimensional model when $\gamma = 0$, the special case where there is no reversion from I to E. From (**??**), we have:

$$\begin{aligned}
0 &= \beta_2 EI - \mu I \\
\Rightarrow 0 &= I(\beta_2 E - \mu) \\
\Rightarrow I &= 0 \text{ or } E = \frac{\mu}{\beta_2}.
\end{aligned}$$

Taking $I = 0$ and $S = 1 - E - I$ in (**??**), we get:

$$\begin{aligned}
0 &= \beta_1(1 - E)qE - \mu E \\
\Rightarrow E &= 0 \text{ or } E = 1 - \frac{\mu}{\beta_1 q}.
\end{aligned}$$

So our first two equilibrium points are $X_1 = (1, 0, 0)$ and $X_2 = \left(\frac{\mu}{\beta_1 q}, 1 - \frac{\mu}{\beta_1 q}, 0\right)$. We note that $I = 0$ in both of these equilibrium points. Since ultimately we would like to have as many individuals in the I class as possible, we consider the equilibrium points without I individuals the "cooperation free" equilibria.

In order for X_2 to have contextual meaning, we must have $\frac{\beta_1 q}{\mu} > 1$; otherwise we have a negative population. Therefore, we define the threshold quantity $R_d = \frac{\beta_1 q}{\mu}$. R_d measures the ability of members of the E class to "convert" or mentor S-class novices into the E class before leaving the community; $R_d > 1$ means E individuals mentor the S class well enough to convert more than one S individual on average to an E individual over their lifetime within the learning community, when $S \approx 1$. We call R_d the *low-peer pressure basic reproduction number*.

B.1 Stability of Cooperation Free Equilibria

We analyze the stability of the equilibrium points by linearizing around each within the two-dimensional system. First we compute the general Jacobian matrix recalling the substitution $S = 1 - E - I$:

$$\begin{bmatrix} \beta_1(q - 2qE - I - qI) - \beta_2 I - \mu & \beta_1(1 - qE - E - 2I) - \beta_2 E \\ \beta_2 I & \beta_2 E - \mu \end{bmatrix}$$

Evaluated at $X_1 = (1, 0, 0)$, the Jacobian is:

$$\begin{bmatrix} \mu(R_d - 1) & \beta_1 \\ 0 & -\mu \end{bmatrix}$$

Since this matrix is upper triangular, the eigenvalues are $\lambda_1 = \mu(R_d - 1)$ and $\lambda_2 = -\mu$. Thus, applying the Routh-Hurwitz criteria, X_1 is locally asymptotically stable when $R_d < 1$ and unstable otherwise.

Note that when not only $R_d = \frac{\beta_1 q}{\mu} < 1$, but in fact $\frac{\beta_1}{\mu} < 1$, we can show that X_1 is globally asymptotically stable by using the Lyapunov function $V = E + I$:

$$\begin{aligned} V(0) &= 0 \\ V &> 0 \ \text{if} \ E + I \neq 0 \\ \frac{dV}{dt} &= \beta_1 S(qE + I) - \mu(E + I) < \beta_1 \left(S - \frac{\mu}{\beta_1} \right)(E + I) \\ &< \beta_1 \left(1 - \frac{\mu}{\beta_1} \right)(E + I) \\ &< 0 \ \text{if} \ \frac{\beta_1}{\mu} < 1. \end{aligned}$$

Similarly, the Jacobian at $X_2 = \left(\frac{1}{R_d}, 1 - \frac{1}{R_d}, 0 \right)$ is:

$$\begin{bmatrix} \mu(1 - R_d) & \beta_1 \left(1 - \frac{1}{R_d} \right) \left[\frac{R_d}{R_d - 1} - (q + 1) - \frac{\beta_2}{\beta_1} \right] \\ 0 & \mu \left(\frac{\beta_2}{\mu} \left(1 - \frac{1}{R_d} \right) - 1 \right) \end{bmatrix}$$

This is also an upper triangular matrix; the eigenvalues are simply $\lambda_3 = \mu(1 - R_d)$ and $\lambda_4 = \mu \left(\frac{\beta_2}{\mu} \left(1 - \frac{1}{R_d} \right) - 1 \right)$. We note that λ_3 is negative when $R_d > 1$. As for λ_4, it is negative when $\frac{\beta_2}{\mu} \left(1 - \frac{1}{R_d} \right) - 1 < 0$. Otherwise, if $\lambda_4 > 0$, X_2 is a saddle point.

Based on the above characterization of X_2, we define $R_0 = \frac{\beta_2}{\mu} \left(1 - \frac{1}{R_d} \right)$. This way we have two stages of cooperation free status. The first stage shifts from having no E or I individuals based on whether R_0 is negative or positive (i.e., $R_d < 1$ or $R_d > 1$), to having E class individuals but still no I individuals. The second stage shifts from not having I individuals to having population in the I class based on whether $R_0 < 1$ or $R_0 > 1$.

Implicit in R_0 we have $1 - \frac{1}{R_d}$ which we can now understand as the proportion of S class individuals becoming E class individuals when the I class is just arising (i.e., $I \approx 0$). Thus, we define R_0 as the *high-peer pressure reproductive number*. $R_0 > 1$ implies that an I individual on average is responsible for promoting more than one E individual into the I class over her/his lifetime when $S \approx \frac{1}{R_d}$ and $E \approx 1 - \frac{1}{R_d}$.

2

C Endemic Equilibria

We use $E_3 = \frac{\mu}{\beta_2}$ to determine the remaining equilibria. In this case we must use (**??**) to solve for the corresponding S_3 value and then check $I = 1 - S - E$ for the corresponding I_3 value.

$$
\begin{aligned}
0 &= \beta_1 S(qE + I) - \beta_2 EI - \mu E \\
\Rightarrow \; 0 &= \beta_1 S\left(q\frac{\mu}{\beta_2} + I\right) - \beta_2 \left(\frac{\mu}{\beta_2}\right) I - \mu \left(\frac{\mu}{\beta_2}\right) \\
\Rightarrow S_3 &= \frac{\mu(\mu + I\beta_2)}{\beta_1(q\mu + I\beta_2)}
\end{aligned}
\tag{4}
$$

Using $E_3 = \frac{\mu}{\beta_2}$ and (**??**) in $I = 1 - S - E$, we solve for I_3:

$$
\begin{aligned}
0 &= 1 - \frac{\mu}{\beta_2} - \frac{\mu(\mu + I_3\beta_2)}{\beta_1(q\mu + I_3\beta_2)} - I_3 \\
0 &= I_3^2 + I_3\left(\frac{\mu}{\beta_1} + (1+q)\frac{\mu}{\beta_2} - 1\right) + \left(\frac{\mu^2}{\beta_1\beta_2} - \frac{q\mu}{\beta_2} + \frac{q\mu^2}{\beta_2^2}\right)
\end{aligned}
\tag{5}
$$

From (**??**), for certain conditions there may exist one or two positive I_3 values for the same E_3 value.

Based on (**??**), we let $x = I_3$, $A = 1$, $B = \left(\frac{q\mu}{\beta_2} - 1 + \frac{\mu}{\beta_1} + \frac{\mu}{\beta_2}\right)$, and $C = \left(\frac{\mu^2}{\beta_1\beta_2} - \frac{q\mu}{\beta_2} + \frac{q\mu^2}{\beta_2^2}\right)$. Then (**??**) becomes $f(x) = Ax^2 + Bx + C = 0$. First we shall show that for $R_0 > 1$, there is a unique endemic equilibrium. We observe that $R_0 > 1 \Longrightarrow f(0) < 0$ from the following:

$$
\begin{aligned}
f(0) &= \frac{\mu^2}{\beta_1\beta_2} - \frac{q\mu}{\beta_2} + \frac{q\mu^2}{\beta_2^2} \\
&= \frac{q\mu^2}{\beta_2^2}\left(\frac{\beta_2}{q\beta_1} - \frac{\beta_2}{\mu} + 1\right) \\
&= \frac{q\mu^2}{\beta_2^2}\left[1 - \frac{\beta_2}{\mu}\left(1 - \frac{\mu}{q\beta_1}\right)\right] \\
&= \frac{q\mu^2}{\beta_2^2}(1 - R_0)
\end{aligned}
\tag{6}
$$

Next we observe that $f(1) > 0$.

$$
\begin{aligned}
f(1) &= 1 + \frac{q\mu}{\beta_2} - 1 + \frac{\mu}{\beta_1} + \frac{\mu}{\beta_2} + \frac{\mu^2}{\beta_1\beta_2} - \frac{q\mu}{\beta_2} + \frac{q\mu^2}{\beta_2^2} \\
&= \frac{\mu}{\beta_1} + \frac{\mu}{\beta_2} + \frac{\mu^2}{\beta_1\beta_2} + \frac{q\mu^2}{\beta_2^2}
\end{aligned}
\tag{7}
$$

Since $f(0) < 0$ and $f(1) > 0$, we know that the graph $f(x) = 0$ of (**??**) crosses the x-axis once to the left of zero and once in $(0,1)$, proving the existence of a unique positive equilibrium point, $X_3 = (S_3, \frac{\mu}{\beta_2}, I_3)$, when $R_0 > 1$.

3

To check the stability of this equilibrium point, we consider the Jacobian of the two-dimensional model evaluated at $E_3 = \frac{\mu}{\beta_2}$ where $I_3 > 0$ (instead of solving for I_3 explicitly due to the complexity of the form of the solution). The Jacobian evaluated at X_3 is:

$$\begin{bmatrix} \beta_1 q - \frac{2\beta_1 q\mu}{\beta_2} - I_3(\beta_1 + \beta_1 q + \beta_2) - \mu & \beta_1 - (1+q)\frac{\beta_1\mu}{\beta_2} - 2\beta_1 I_3 - \mu \\ \beta_2 I_3 & 0 \end{bmatrix}$$

with determinant and trace:

$$\text{determinant} = -\beta_2 I_3 \left[\beta_1\left(1 - (1+q)\frac{\mu}{\beta_2} - 2I_3\right) - \mu \right] = \beta_1\beta_2 I_3(B + 2I_3), \qquad (8)$$

$$\text{trace} = \beta_1 q - \frac{2\beta_1 q\mu}{\beta_2} - I_3(\beta_1 + \beta_1 q + \beta_2) - \mu. \qquad (9)$$

To satisfy the Routh-Hurwitz criteria we require the determinant > 0 and the trace < 0. We have that

$$\text{determinant} > 0 \iff I_3 > \frac{-B}{2A}. \qquad (10)$$

Since $R_0 > 1 \Rightarrow C < 0$, and $A = 1 > 0$, $I_3 = \frac{-B+\sqrt{B^2-4AC}}{2A} > -\frac{B}{2A}$. Similarly:

$$\text{trace} < 0 \iff I_3 > \frac{\beta_1\beta_2 q - 2\beta_1 q\mu - \mu\beta_2}{\beta_1\beta_2(1+q) + \beta_2^2}. \qquad (11)$$

It can be shown that this condition holds true for $R_0 < 2$ or $I_3 > -\frac{B}{2A}$, so this condition is also always satisfied for $R_0 > 1$.

D   Backward Bifurcation

We turn our attention to when $R_0 < 1$. In this case, under certain conditions there can exist two positive solutions to (**??**). These solutions, with appropriate stabilities, arise when the system exhibits a backward bifurcation.

Note that the condition $R_0 < 1$ is equivalent to:

$$\frac{1}{\mu} < \frac{1}{\beta_1 q} + \frac{1}{\beta_2}. \qquad (12)$$

Next we consider the quadratic equation (**??**) defined above. With $R_0 < 1 \Rightarrow C > 0$, for there to exist two positive solutions requires also (a) that the vertex be in $(0, 1)$, meaning $B < 0$ and $f'(1) > 0$, and (b) that there are two real zero crossings, i.e., $B^2 - 4AC > 0$. From (a) we have

$$f'(x) = 2x + \frac{\mu}{\beta_1} + (1+q)\frac{\mu}{\beta_2} - 1$$
$$\Rightarrow f'(1) = 1 + \frac{\mu}{\beta_1} + (1+q)\frac{\mu}{\beta_2}, \qquad (13)$$

4

which is always positive, and

$$0 \;<\; -\frac{1}{2}\left(\frac{\mu}{\beta_1} + (1+q)\frac{\mu}{\beta_2} - 1\right)$$
$$\frac{1}{\mu} \;>\; \frac{1}{\beta_1} + \frac{1}{\beta_2}(1+q). \tag{14}$$

From (b) we have

$$0 \;<\; \left(\frac{\mu}{\beta_1} - 1 + \frac{\mu}{\beta_2}(q+1)\right)^2 - 4\left(\frac{\mu^2}{\beta_1\beta_2} - \frac{q\mu}{\beta_2} + \frac{q\mu^2}{\beta_2^2}\right)$$
$$0 \;<\; \frac{\mu^2}{\beta_1^2} - 2\frac{\mu}{\beta_1} + 2\frac{\mu^2}{\beta_1\beta_2}(q-1) + 1 + 2\frac{\mu}{\beta_2}(q-1) + \frac{\mu^2}{\beta_2^2}(q-1)^2$$
$$0 \;<\; \left(\frac{\mu}{\beta_1} - 1 - \frac{\mu}{\beta_2}(1-q)\right)^2 - 4\frac{\mu}{\beta_2}(1-q)$$
$$2\sqrt{\frac{\mu}{\beta_2}(1-q)} \;<\; \left|\frac{\mu}{\beta_1} - 1 - \frac{\mu}{\beta_2}(1-q)\right|. \tag{15}$$

When $B < 0$ so that (??) holds, we have that $\frac{\mu}{\beta_1} - 1 - \frac{\mu}{\beta_2}(1-q) < 0$ in (??), and we need the positive root so we multiply by (-1) as follows:

$$2\sqrt{\frac{\mu}{\beta_2}(1-q)} \;<\; -1\left(\frac{\mu}{\beta_1} - 1 - \frac{\mu}{\beta_2}(1-q)\right)$$
$$\frac{\mu}{\beta_1} \;<\; 1 - 2\sqrt{\frac{\mu}{\beta_2}(1-q)} + \frac{\mu}{\beta_2}(1-q)$$
$$\frac{\mu}{\beta_1} \;<\; \left(1 - \sqrt{\frac{\mu}{\beta_2}(1-q)}\right)^2$$
$$\sqrt{\frac{\mu}{\beta_1}} \;<\; 1 - \sqrt{\frac{\mu}{\beta_2}(1-q)}$$
$$\left(\sqrt{\frac{1}{\beta_1}} + \sqrt{\frac{1}{\beta_2}(1-q)}\right)^2 \;<\; \frac{1}{\mu}. \tag{16}$$

It can be shown that, given inequalities (??) and (??), (??) can be replaced with the simpler $qR_d < 1$. Thus, when the parameter values for conditions (??), (??) and either (??) or $qR_d < 1$ are met, we can expect the system to exhibit a backward bifurcation, meaning there are two simultaneous endemic equilibria when $R_0 < 1$. Note that if (??) or $qR_d < 1$ is broken, a cooperative equilibrium is globally stable, while if (??) is broken no I class can survive. Thus we focus our interpretation on (??) in Section 4 of the main text.