Theoretical Description of Teaching-Learning Processes: A Multidisciplinary Approach

Clelia M. Bordogna[1,2] and Ezequiel V. Albano[1]

[1]Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas, (INIFTA), CONICET, UNLP, Sucursal 4, Casilla de Correo 16, (1900) La Plata, Argentina

[2]IMApEC, Facultad de Ingeniería, UNLP, La Plata, Argentina

(Received 9 August 2000; revised manuscript received 16 January 2001; published 27 August 2001)

A multidisciplinary approach based on concepts from sociology, educational psychology, statistical physics, and computational science is developed for the theoretical description of teaching-learning processes that take place in the classroom. The emerging model is consistent with well-established empirical results, such as the higher achievements reached working in collaborative groups and the influence of the structure of the group on the achievements of the individuals. Furthermore, another social learning process that takes place in massive interactions among individuals via the Internet is also investigated.

DOI: 10.1103/PhysRevLett.87.118701 PACS numbers: 87.23.Ge, 01.40.Fk, 02.70.Rr, 89.70.+c

The study and understanding of complex systems is a promising field of multidisciplinary research that has attracted the attention of an increasing number of scientists [1]. Learning processes are quite complex and their study has also developed into a very active area of investigation [2]. Early studies on these processes have been conducted by psychologists and sociologists [3,4]. Recent thought on educational psychology suggests that learning processes take place while people participate within social communities [5]. Also, learning is not restricted to any type of intentional education but involves all kinds of social activities [2,5]. However, the topic is so difficult and complex that there are still many controversies and open questions. Within a more restricted social context, teaching-learning processes that take place in the classroom (TLC) are being extensively investigated. For example, the processes of learning and understanding physics and mathematics have become the focus of cognitive research, because these subjects are based on well-defined conceptual frameworks [6]. Therefore, an increasing number of physicists and mathematicians have also been attracted to the study of cognitive processes [7].

The aim of this work is to propose and study a model for TLC that can be treated mathematically, solved numerically, and analyzed statistically. Furthermore, it is shown that the model provides a suitable framework for the study of other social learning processes, such as learning via the Internet. Therefore, this novel and multidisciplinary approach links psychological and sociological theories of impact [8], education psychology [2], mathematics, computer science, and statistical physics [9].

The framework adopted is a generalized Ising model, similar to those used to treat spin systems and neural networks. These kinds of models have been employed to describe other social activities and behavioral processes in sociology [10] and economy [11]. So, let us now define and discuss relevant concepts. The cognitive impact (CI) acting on an individual is the overall result of those interactions with his/her environment (teachers, peers, bibliography, etc.), capable of modifying his/her knowledge and the self-elaboration of such influence. The individual can also become a source of CI to other individuals by persuading and supporting [12]. The persuasiveness, $P_{ji} \geq 0$, describes the degree to which the ith individual can persuade the jth individual. Also, the support, $S_{ij} \geq 0$, describes the degree to which the ith individual supports the statements of the jth individual. Within an interactive group, both S_{ij} and P_{ij} become enhanced when individuals share similar ideas about the subject under examination; they have social and cultural affinities, etc. The knowledge of the jth individual, $[\sigma_j(t)]$, at time t, is defined as a dynamic variable such as $-1 \leq \sigma_j(t) \leq 1$, where $\sigma_j(t) = 1$ corresponds to optimum knowledge.

Based on these considerations, we propose that the CI of the teacher on the jth student $[CI^{\mathrm{TS}}(j,t)]$, can be written as

$$CI^{\mathrm{TS}}(j,t) = P_{jT}[1 - \sigma_j(t)\sigma_T], \tag{1}$$

where $\sigma_T > 0$ and P_{jT} are the knowledge of the teacher and his/her ability to persuade the jth individual, respectively. P_{jT} depends on many factors, characteristic of both the teacher itself and the teacher-individual relationship, such as, e.g., the rhetorical ability and the persuasive skills of the teacher, the didactic presentation of the subject of study, etc. Notice that CI^{TS} is minimal for $\sigma_j = 1$ and $\sigma_T = 1$, because it corresponds to the impact between two individuals having the same (maximum) knowledge. Also, CI^{TS} is maximal for $\sigma_j = -1$ and $\sigma_T = 1$, due to the largest difference in the knowledge.

Within groups of N individuals, the CI of the student-student interaction (supervised by the teacher) $CI^{\mathrm{SS}}(j,t)$, is given by

$$CI^{\mathrm{SS}}(j,t) = \sum_{i=1, i\neq j}^{N} \{P_{ij}(t)[1 - \sigma_i(t)\sigma_j(t)] - S_{ij}(t) \times [1 + \sigma_i(t)\sigma_j(t)]\} \operatorname{sgn}[\sigma_i(t)/\sigma_T], \tag{2}$$

where the first (second) term accounts for the mutual persuasiveness (support). The structure of these two terms is similar to that of Eq. (1) and it is plausible since it is expected that mutual support will be larger when the individuals have similar knowledge ($\sigma_i \sigma_j > 0$) while persuasiveness is expected to play a more relevant role in the opposite case ($\sigma_i \sigma_j < 0$). It is also assumed that both S_{ij} and P_{ij} are composed of intrinsic and extrinsic (or interactive) factors, so $S_{ij}(t) = S_{ij}^o[\sigma_T + \sigma_i(t)]$, and $P_{ij}(t) = P_{ij}^o[\sigma_T + \sigma_i(t)]$, where the intrinsic terms, S_{ij}^o and P_{ij}^o, depend on many factors such as the strength of psychological coupling, affinity of social status, education, rhetorical abilities, personal skills, etc. The extrinsic term is provided by a comparison established by the individual with the teacher who assumes a leadership role. This term is included to account for the fact that the model attempts to describe supervised collaborative group work [13]. In fact, the term $\sigma_T + \sigma_i(t)$ means that both persuasiveness and support between individuals could be either reinforced or weakened when the knowledge of the teacher is taken as a reference level. In addition, the term $A = \mathrm{sgn}[\sigma_i(t)/\sigma_T]$ in Eq. (2) explicitly accounts for the plausible fact that an individual with knowledge below the average ($\sigma_i < 0$) has little chance to cause an increment of the knowledge of another individual which is above the average ($\sigma_j > 0$). Also, due to this term, in the inverse case ($\sigma_i > 0, \sigma_j < 0$), the jth individual has a great chance to increase his/her knowledge. It should be noticed that CI^{SS} may be either positive, negative, or zero.

The CI of the bibliography and other sources of information $[CI^{\mathrm{BS}}(j,t)]$ are given by

$$CI^{\mathrm{BS}}(j,t) = A(j)Q(j)[1 - \sigma_j(t)], \tag{3}$$

where $0 \leq A(j) \leq 1$ is the ability of the jth individual to understand the available material that involves the individual's own capacity to perform critical analysis, and to establish relationships between topics, etc. Also, $0.1 \leq Q(t) \leq 1$ is the quality of such material.

The knowledge is a dynamic variable influenced by the CI. So, during a time interval Δt, the knowledge changes as follows: $\sigma_j(t + \Delta t) = \sigma_j(t) \pm \Delta\sigma$, where for the calculation σ_j is assumed to be discrete so that $\Delta\sigma$ is the "quantum" of knowledge. Notice that $\sigma_j(t)$ has an upper bound given by the maximum knowledge of the available sources, e.g., teachers (σ_T), bibliography Q, etc. Also, $\sigma_j(t)$ may improve (become worse) with the probability $P_j = \tau_j/(1 + \tau_j)$ and $(1 - P_j)$, where τ_j is a generalized Metropolis rate [9] given by

$$\tau_j = e^{\beta_{\mathrm{TS}} CI^{\mathrm{TS}}(j,t) + \beta_{\mathrm{SS}}(N) CI^{\mathrm{SS}}(j,t) + \beta_{\mathrm{BS}} CI^{\mathrm{BS}}(j,t)}, \tag{4}$$

where each process has its own "noise" given by $1/\beta_{\mathrm{TS}}$, $1/\beta_{\mathrm{SS}}(N)$, and $1/\beta_{\mathrm{BS}}$, respectively. In fact, for the teaching-student relationship, the noise may be due to disorder in the classroom, inappropriate teaching material, lack of attention of the students, obscure explanations, etc. For the student-student interactions the noise $1/\beta_{\mathrm{SS}}(N)$ appears due to disordered discussions, misunderstandings, the lack of a well-organized participative activity, etc. In this case, the dependence on the number of students N has been considered to account for the division of the impact observed upon interactions within groups [8]. Finally, $1/\beta_{\mathrm{BS}}$ may be due to inappropriate selection of the bibliographic material, lack of attention, etc.

It is worth mentioning that in order to understand the plausibility of Eqs. (1)–(3) it is essential to analyze them in connection to Eq. (4) [14]. Also, in the absence of any source of knowledge ($CI^{\mathrm{TS}} \equiv 0$ and $CI^{\mathrm{BS}} \equiv 0$), the term CI^{SS} may be suitable for the description of the spontaneous emergency of knowledge, either positive or negative with the same probability. The study of this phenomenon is beyond the aim of this work; instead we have analyzed the case where such symmetry is broken by a source of knowledge. The proposed model is simulated by means of a standard Monte Carlo technique [9].

In order to study the influence of the structure of the collaborative groups on the TLC process we have assumed $\sigma_T = 1$, $P_{jT} = 1 \quad \forall\, j$, and $\Delta\sigma = 0.1$, with $i, j = 1, \ldots, N_T$, where N_T is the total number of individuals. Also, S_{ji}^o and P_{ji}^o are assumed to be randomly distributed in the interval $(0, 1)$, so their average value over the whole classroom is close to $1/2$. It has also been assumed that the students can be classified into three different sets, namely "high-achieving (HA) students" with $\langle\sigma\rangle_{\mathrm{HA}} = 0.5$, "average-achieving (AA) students" with $\langle\sigma_{\mathrm{AA}}\rangle = 0$, and "low-achieving (LA) students" with $\langle\sigma_{\mathrm{LA}}\rangle = -0.5$. Figure 1 shows the time evolution of the knowledge of the students corresponding to three different cases with $\beta = \beta_{\mathrm{TS}} = \beta_{\mathrm{SS}} = 4$ and $\beta_{\mathrm{BS}} = \infty$. In case I

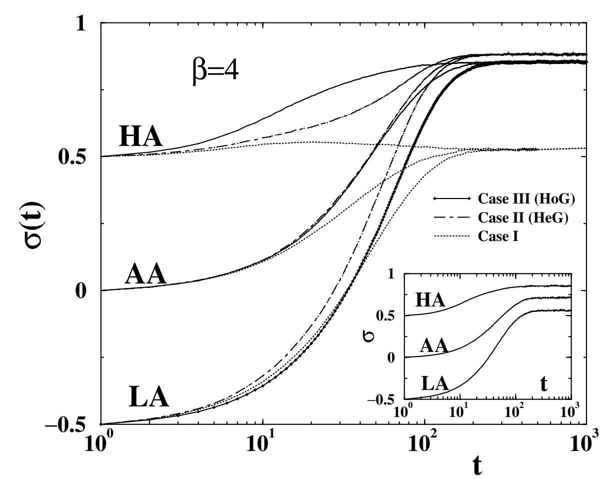

FIG. 1. Plots of the time evolution of the knowledge of HA, AA, and LA individuals in three different environments. In case I individuals only attend lectures of the teacher. In cases II and III the individuals are also engaged in collaborative group work forming heterogeneous and homogeneous groups, respectively. The inset shows results for HoG where the interactions between individuals are weakened according to the ratio HA:AA:LA → 1:2:10. The total number of individuals is $N_T = 96$ and results are averaged over 10^5 different cases. More details are provided in the text.

the students only attend the lectures of the teacher. Cases II and III correspond to students that not only attend such lectures, but are also engaged in collaborative work forming groups of $N = 3$ individuals [15]. However, those groups are formed in two different ways: in case II the groups are heterogeneous (HeG) and the members are chosen at random, but in case III the groups are homogeneous (HoG) and all the members of each group are selected so that they have similar initial achievements. An analysis of Fig. 1 clearly indicates the occurrence of three distinct time regimes: For the short time regime ($t < 5$) the knowledge is almost independent of both the employed method and the structure of the groups. In the intermediate time regime (roughly for $5 < t < 80$) it is clear that the performance of HA students is higher when they form HoG, while LA students perform better when they are in HeG groups. This finding is very well documented by numerous studies, see, e.g., [16,17], and references therein. Finally, for the long time regime (say $t > 80$) it is clear that the achievement of the students involved in collaborative work is much better than that of those attending only the lectures, in agreement with extensive studies [18]. Also, an interesting result is that HeG have obtained better (final) knowledge than HoG. The results shown in Fig. 1 can be discussed in light of the available evidence: (i) The fact that the composition of the groups may cause LA students to learn at the expense of HA students has already been recognized [16]. The dilemma faced by teachers in order to determine optimal grouping strategies becomes evident, due to the fact that the performance of all kinds of students cannot be optimized simultaneously. However, our results show that achievements may depend on how long the course is. (ii) Our result from the best performance of HeG for long instructional times is subtle, but evidence showing this behavior has been reported [19]. It should be noticed that this difference depends on the noise vanishing for $1/\beta = 1/2$, while for $1/\beta = 1$ the opposite trend is observed.

Figure 1 shows that homogeneous LA classes will, over time, rise to a similar level of knowledge as all other combinations of classes. However, tracking research has found, in like cases, dramatically different results [20]. This kind of behavior can also be accounted for by the model just changing the strength of the interactions. In fact, assuming that persuasiveness and support are weakened for HoG according to the ratio HA:AA:LA $\rightarrow$ 1:2:10, different final outputs are obtained, as shown in the inset of Fig. 1.

Extensive test studies have demonstrated that in a large number of cases it is difficult to find differences between the traditional teaching method and the modern approach of group work [18]. In order to help understand these observations, we have also performed simulations with groups in different environments. The initial knowledge of the students is assumed to be uniformly distributed, $-1 \leq \sigma_j(t = 0) \leq 1$ and HeG of $N = 3$ students are considered [15]. Figure 2 shows plots of the maximum knowledge achieved after a long instructional time (σ_M) as a function $1/\beta$, with $\beta_{TS} = \beta_{SS}$. It is found that using the traditional method σ_M decreases steadily when the noise is increased. In contrast, the knowledge achieved in collaborative groups is more robust and exhibits a sharp drop only for $1/\beta \approx 6$. It is found that collaborative work always improves the achievements (see the inset of Fig. 2). However, the achievements of very good teachers (smaller values of $1/\beta$) can only slightly be improved by the groups, a fact which makes it difficult to detect differences using tests. This is also difficult in the other extreme case, e.g., for bad teachers and noisy groups. Also, there is an intermediate regime where the difference becomes maximum and tests have great chances to be useful [18].

Nowadays, the Internet has become an attractive media which favors the interactions among a large number of individuals. The application of the proposed model to treat this social learning process is straightforward. First, it is assumed that the process is no longer supervised by a teacher, so the first term of Eq. (4) and the standard of knowledge provided by the teacher in Eq. (2) have to be neglected. It is also assumed that, while a large number of individuals ($N_T = 1024$) are engaged in the process, only subsets having N_L individuals are actually linked among them. Also, the members of each subset are selected at random during each time step, in order to allow the branching of ideas. The division of the impact [8] is considered taking $\beta_{SS}(N_L) = \beta(3)/(N_L/3)^{\alpha}$, where α is an exponent and it is assumed that the best achievement is obtained for $N_L = 3$ [15]. Figure 3 shows plots of the maximum knowledge (σ_M) achieved by the individuals as a function

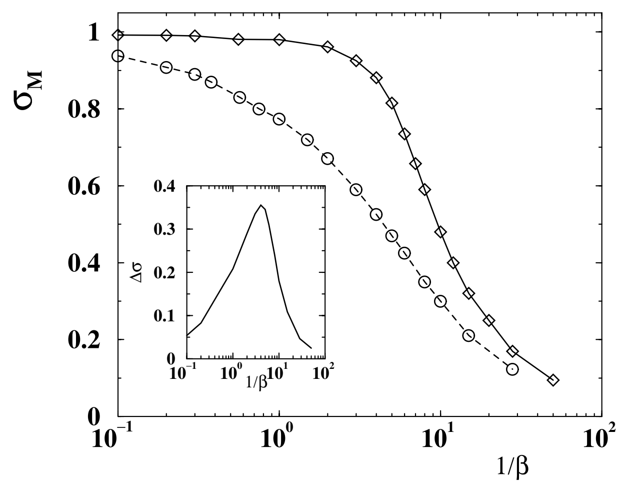

FIG. 2. Plots of the maximum achieved knowledge (σ_M) versus the noise $1/\beta$, as obtained for (a) ○, individuals only attending lectures of the teacher; (b) ◇, individuals as in (a) but also engaged in collaborative work forming groups of three members. The inset shows the difference between cases (b) and (a). The total number of individuals is $N_T = 96$ and results are averaged over 10^5 different cases.

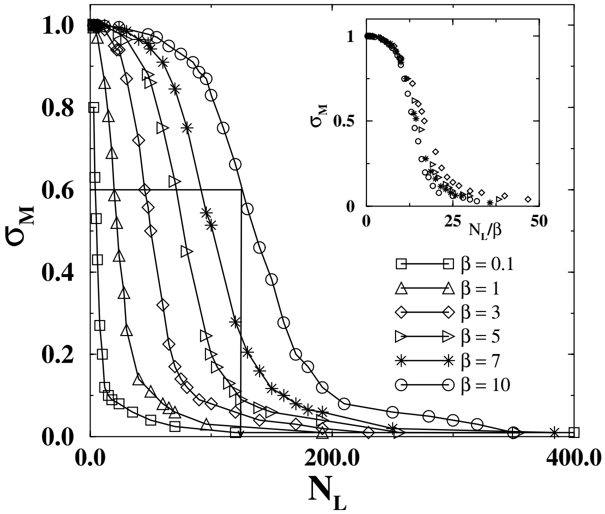

FIG. 3. Plots of the maximum achieved knowledge (σ_M), corresponding to a learning process using the Internet, versus the number of linked individuals (N_L) and obtained for different values of the noise, as indicated in the figure. Setting a threshold for the knowledge $\sigma_M^T = 0.6$ (horizontal full line), gives the maximum connectivity $N_L^M \simeq 124$ (vertical arrow). Results averaged over 10^4 different cases. The inset shows the data collapsing that were obtained by plotting σ_M versus N_L/β.

of N_L for $\alpha = 2$. The main results shown in Fig. 3 are (i) keeping the noise constant, σ_M decreases steadily with the connectivity, reflecting the fact that for $\alpha = 2$ the division of the impact always prevails. (ii) For each value of β, every threshold previously defined as an acceptable knowledge (σ_M^T) is compatible with a maximum possible connectivity (N_L^M), as shown in Fig. 3 for $\beta = 10$ and $\sigma_M^T = 0.6$. It is also found that $N_L^M \propto \beta$, a result that can be interpreted as a compromise between noise and connectivity: the greater the noise the smaller the amount of individuals that can be successfully linked. All these results can be summarized by means of a universal curve, as shown in the inset of Fig. 3. The collapsed curve gives $N_L^M/\beta = 14 \pm 2$ for $\sigma_M^T = 0.6$, which sets the maximum connectivity for such a threshold. (iii) Neglecting the division of the impact ($\alpha \leq 1$) we have obtained "naive" results, namely, optimum knowledge ($\sigma_M \simeq 1$) can always be achieved, irrespective of the noise, simply by increasing the connectivity toward an unrealistic large number of individuals.

Summing up, a theoretical framework for the study of learning processes that take place in the classroom has been developed. The theory is flexible and can be applied to other areas of social learning, as in the case of learning via the Internet.

This work was supported by UNLP, CONICET, and ANPCyT (Argentina).

[1] See, e.g., the special section devoted to complex systems published in, Science **284**, 79–109 (1999).
[2] *Handbook of Educational Psychology,* edited by D. C. Berliner and R. C. Calfee (MacMillan, New York, 1996).
[3] J. Piaget, *The Childs Conception of World* (Harcourt, Brace, New York, 1929).
[4] L. S. Vygotsky, *Mind in Society,* edited by and translated by M. Cole, V. Steiner, S. Scribner, and E. Souberman (Cambridge University Press, Cambridge, MA, 1978).
[5] J. Lave and E. Wenger, *Situated Learning* (Cambridge University Press, Cambridge, MA, 1999).
[6] See, e.g., D. Hestenes, Am. J. Phys. **55**, 440 (1987).
[7] See, e.g., a set of interesting articles collected in the Supplement 1 to the Am. J. Phys. **67**, S3–S62 (1999); **68**, S3–S63 (2000).
[8] B. Latané, Am. Psychol. **36**, 343 (1981).
[9] *Application of the Monte Carlo Method in Statistical Physics,* edited by K. Binder (Springer-Verlag, Heidelberg, 1987).
[10] M. Lewenstein, A. Nowak, and B. Latané, Phys. Rev. A **45**, 763 (1992); K. Kacperski and J. A. Holyst, Physica (Amsterdam) **269A**, 511 (1999); Y. Kamp, Phys. Rev. E **62**, 1725 (2000).
[11] D. Chowdhury and D. Stauffer, Eur. Phys. J. B **8**, 477 (1999).
[12] D. Helbing, *Quantitative Sociodynamics* (Kluwer, Dordrecht, 1995).
[13] N. Webb and A. S. Palincsar, in "Group Processes in the Classroom," Ref. [2], p. 841.
[14] If $CI > 0$ ($CI < 0$) the *j*th individual has a large (small) chance to enhance his/her knowledge, as found for $\sigma_i > 0, \sigma_j < 0$ ($\sigma_i < 0, \sigma_j > 0$).
[15] The optimal group size depends on the nature of the task and the experience of the group members. Group sizes between two and six individuals are recommended in various contexts; see, e.g., J. L. Larkin *et al.,* Science **208**, 1335 (1980), and Refs. [13,19].
[16] N. Webb, K. Nemer, C. Chizhik, and B. Sugrue, Am. Educ. Res. J. **35**, 607 (1998).
[17] L. Fuchs, D. Fuchs, C. Hamlett, and K. Karns, Am. Educ. Res. J. **35**, 227 (1998).
[18] R. Hake, Am. J. Phys. **66**, 64 (1998).
[19] P. Heller and M. Hollabaugh, Am. J. Phys. **60**, 637 (1992).
[20] J. Oakes and M. Lipton, *Teaching to Change the World* (McGraw-Hill, Boston, 1999).