The “flat peer learning” agent-based model

Philippe Collard¹

Received: 5 December 2020 / Accepted: 21 April 2021 / Published online: 30 May 2021
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2021

Abstract

This paper deals with *peer learning* and, in particular, with the phenomena of *exclusion*; it proposes to model a group of learners where everyone has his own behaviour that expresses his way of following a curriculum. The focus is on individual motivations that avoid disadvantage certain individuals while optimising behaviour at the community level; in this context, the approach is based on the belief that the induced learning dynamics can be clarified by the contribution of agent-based modelling and its entry into the field of peer learning simulation. *Flat learning* means here that every learner features the same initial skill level, along with the same opportunities to learn both independently and with the help of peers. To address this topic the paper proposes the *Flat Peer Learning* agent-based computational model inspired by the Vygotsky’s social and learning theory. The paper shows that even if strict equity could be guaranteed, educators would still be faced with the dilemma of having to choose between optimising the learning process for the group or preventing exclusion for some.

Keywords Peer learning · Social exclusion · Multi-agent-based modelling

Introduction

While peer learning has been successfully applied in various discipline-based education research agenda, much is left to understand about the underlying dynamics of skill improvement among learners. For instance, little is known about how both peer learning and independent learning processes affect one another in a group of learners [15, 16, 23].

The originality of the approach resides in a focus on looking for causes that go beyond the role of intelligence, family environment, social status,... or even the fact that schools guarantee a level playing field for all children; rather, but without

✉ Philippe Collard
philippe.collard@unice.fr

¹ Laboratoire I3S, UMR 7271, CNRS (MDSC team), Université Côte d’Azur, CS 40121, 06903 Sophia Antipolis Cedex, France

ignoring the relevance of it all, the paper explores the causes that may be basically inherent in the very process of knowledge diffusion in a network of peer learners [37]. Beyond the empirical evidence that peer learning works, the question arises as to how it works. Answering this type of question should enable practitioners to design increasingly adaptable and effective forms of peer learning. As Abrami et al. [5] wrote “For many years peer learning was under-theorised, supported by old sayings such as to teach is to learn twice [...] A number of researchers have conducted work with strong implications for building theory in peer learning; however, a plethora of theories does not help the hard-pressed practitioner.”. One way to supplement existing empirical evaluations is to develop abstract models featuring explanatory and predictive capabilities with respect to the knowledge diffusion dynamics. Such approaches have been previously leveraged in psychology and sociology where agent-based modelling (hereinafter ABM) has established itself as an useful tool [28, 32, 33].

This paper proposes an ABM approach to model a group of learners able to increase their skill level through independent study and by engaging in peer learning; it is based on *flat learning* and *reciprocal tutoring*.

Flat peer learning

Although recent years have been marked by a greater awareness of equitable practices in peer learning, it is not easy to know/measure the implication of these efforts towards greater equity [21].

Most often, from the earliest age, children are grouped together in a classroom according to their age; although this limits the differences in the initial levels, it is unfortunately not sufficient to ensure strict equity and, moreover, this practice does not guarantee similar learning abilities. The model we proposed is based on *strict equity* which requires that every learner features the same initial skill level, along with the same capability to learn independently or with the help of peers. As no such perfectly homogeneous group would naturally occur in a real educational settings, this aspect may seem unrealistic at first sight; however, it is precisely this impossibility which justifies using simulations to focus exclusively on the intrinsic properties of the process, independently of the particularities of any specific group of learners. We use the term *flat-learning* here to underline the fact that such a learning system does not exist in real life, even though it may be a goal to which one could aspire.

Reciprocal tutoring

Interest in reciprocal mentoring has also greatly increased in recent years, as it allows everyone to be both helper and helped, avoiding discrimination on the basis of ability and status [5, 18]. It is again the concern not to differentiate the actors in the peer learning process that motivates us to make the hypothesis of *reciprocal tutoring*. Such approach avoids having questionable results because they may be biased by a multitude of confounding factors arising from the learners' individual

characteristics; simulating an homogeneous group eliminates such factors and helps us to isolate the very properties that are inherent to the underlying learning dynamics.

Although this paper focuses exclusively on reciprocal learning, we do not deny the importance of teacher-to-student interactions; indeed, the two approaches are complementary and a combination, closer to the daily reality of learners and pedagogues, is supposed to generate a form of synergy [9, 11, 30].

Purposes

In real peer learning situations, there are both endogenous and exogenous causes that interact with each other; while being aware of this, it is however relevant to focus on the role of the former and in particular to highlight what in the learning process is due to endogenous causes independently of the exogenous ones.

The primary model's purpose is to gain insights on the exclusion phenomenon whereby a learner ends up not being able to improve his skills and falls behind his classmates or within a group engaged in distance learning. On-line learning has developed over the last decade and, since the COVID-19 epidemic, it has become even more central to people's lives. In addition, the constraint of spatial separation has led learners, pupils or students, to go outside the school or academic framework; in this context relational networks have played a central role in promoting learning interactions between people. The concomitance of these two trends shows that the issues raised by peer learning, and in particular that of dropping out, are essential today and are likely to remain relevant in the future [36].

Simulations will answer the following questions: (i) does the exclusion phenomenon manifest itself despite homogeneity among learners? (ii) compared to independent learning, to what extent does peer learning support the learning process? (iii) what factors affect the emergence of exclusion? (iv) can we maximize the global learning performance and, at the same time, minimize exclusion? These questions are different from both educational studies focusing on establishing the impact of a given pedagogical intervention on a given group of students, or educational studies leveraging data mining or machine learning techniques on interactions and performance data [6, 20]. The approach is in the vein of the work on the one-dimensional, probabilistic and totalistic λ CA cellular automata where peer learning have been reformulated as a synchronization problem [14]; because of the strict-equity assumption, it exemplifies a type of research questions that can only be investigated by means of artificial simulations.

Previous works

Here, we present some previous works to establish comparisons according to the different modelling approaches.

The paper “*An agent-based model for teaching–learning processes*” proposed an ABM “for describing the increase the knowledge by accumulating the information needed to complete a learning task or objectives” [30]. Simulations allow

to evaluate the performance of learning in the classroom. From this results the authors propose to built a gas model analogy and thus use such models to interpret the resulting learning process. The common features with the present contribution are: (i) use of an agent-based modelling/simulation approach; (ii) see learning as an accumulative phenomenon; (iii) choose a simplified model of reality. The differences are: (i) studied the influence of interaction with the teacher; (ii) validated on a classroom only; the question then arises whether this model could be extended to a non-homogeneous network of learners; (iii) use of an analogy with a physical process; (iv) does not distinguish between independent learning and peer learning; (v) does not address the phenomenon of exclusion.

Koponen et al. [25] proposed “*An agent-based model of discourse pattern formation in small groups of competing and cooperating members*”. The authors approach discourse patterns in a small group formation through the ABM where patterns are the outcomes of peer-to-peer comparison events. The dynamics result from both *competition* and *cooperation* between agents; it has been shown that low competitiveness leads to egalitarian triads and that increased cooperation favours the formation of such triads. The common features with the present contribution are: (i) use of an agent-based modelling/simulation approach; (ii) refers to peer-to-peer interactions; (iii) consider a situation of *cooperation* between agents. while the differences are: (i) find exclusion cases due to learning differences; (ii) has very specific objectives relating to discourse in group (no learning); (iii) use small size group (four up to seven agents); (iv) consider a situation of *competition* between agents for a common resource (i.e. compete for the floor).

The paper “*Collective learning modeling based on the kinetic theory of active particles*” proposes a “systems approach to the theory of perception and learning in populations composed of many living entities” [11]. From these two processes, the authors derive a mathematical structure which reveals their complexity and modelling uses methods derived from the kinetic theory. Apart from the main topic of peer-learning in a classroom, this article presents many points of differentiation from the present article: (i) describe peer and teacher interactions; (ii) combine perception and learning; (iii) heterogeneity is a central assumption; (iv) use mathematical model rather than ABM; (v) based on kinetic theory and game theory; (vi) only validated on a classroom; the question then arises whether this model could be extended to a non-homogeneous network of learners; (vii) does not focus on the exclusion phenomenon.

Bordogna et al. [9] proposed “*A cellular automata model for social-learning processes in a classroom contexts*”. Drawing on ideas inspired by sociology, educational psychology, statistical physics and computer science, the authors propose to model “teaching-learning processes that take place in the classroom”. They focus on the role of collaborative groupwork and, in particular, on the size of such groups to track the effectiveness of the learning process. The common features with the present contribution are: (i) use of an agent-based modelling approach; (ii) take inspiration from sociology and educational psychology; (iii) point out that collaboration between learners is a key point. The differences are: (i) the presence of teachers; (ii) only validated on a classroom; the question then

arises whether this model could be extended to a non-homogeneous network of learners; (iii) take inspiration from statistical physics.

The paper “*Theoretical description of teaching-learning processes: A multidisciplinary approach*” proposes a “systems approach to the theory of perception and learning in populations composed of many living entities” [8]. Although it has much in common with the previous article, its originality lies in the study of the learning process that results from interactions between individuals *via* the Internet; however, the very structure of the network is not considered.

The remainder of the paper is articulated in five main parts: “[The flat peer learning model](#)” presents the peer learning model; “[Simulation with independent learning](#)” and “[Simulation with peer learning](#)” present the simulations and results; “[Discussion](#)” discuss on the peer learning network; then we conclude with a summary of our findings and discuss future work.

The flat peer learning model

The *Flat peer learning* model (hereinafter *FPL*) models the interactions of a population of agent-learners connected with each other. Each agent features an internal state that is updated *via* two stochastic processes, *independent* and *peer* learning.

Specifics of the model

The model supposes that agents are: (i) autonomous units, free to interact with other agents; (ii) reactive with a form of memory; (iii) heterogeneous regarding their state; (iv) interdependent as they influence others in response to the influence that they receive [31].

Patterns

The learners are represented by a set of M agents; each one occupies a node in a peer network (hereinafter Pn). The learner’s positions in the network define a neighbourhood between individuals; for each one, this is the set of other learners who possibly can help him or that he can help. The potential influence of a learner is the number of connections he has with others; this number is represented by the *node-degree* in Pn . The agents are indexed by the integers, so that a_i is the agent-number i , N_i his neighbourhood and $|N_i|$ his degree.

In the following we will consider successively two patterns for the *FPL* model, a regular *lattice* and a *scale-free* structure; the first one models an homogeneous space zone like a *classroom*, while the second refers rather to a *relational network*.

Discrete states

The current state of agent a_i is his skill level or knowledge with respect to an arbitrary task and is denoted by $state_i$. The model allows for agents to progress through

a fixed number of levels. The set of states values $\Sigma = \{\text{level}_1, \dots, \text{level}_L\}$ is a set of finite cardinal totally ordered with order defined by the *succ* function such as $\forall j$ as $1 \leq j < L$, $\text{succ}(\text{level}_j) = \text{level}_{j+1}$ and $\text{succ}(\text{level}_L) = \text{level}_L$. This defines a total order on the states by:

$$\text{state}_i < \text{state}_j \text{ iff } \exists n \in \mathbb{N}^* \text{ with } \text{state}_j = \text{succ}^n(\text{state}_i)$$

where the integer n means an iterate with n steps.

Process overview

At each time step t , each agent a_i updates his current state state_i^t according to a local transition rule:

$$\begin{aligned} \Phi : \quad \Sigma^{|N_i|+1} &\rightarrow \Sigma \\ sN_i \cup \{\text{state}_i^t\} &\mapsto \text{state}_i^{t+1} \end{aligned}$$

where sN_i is the set of states of all the agents in N_i ; so, the function Φ takes as input the neighbourhood states and the own state of the agent.

At a time, for all the agents, the updates are synchronous. An agent can not progress by more one level at each time step (i.e. $\text{state}_i^{t+1} = \text{state}_i^t$ or $\text{succ}(\text{state}_i^t)$); as the level of an agent never diminishes, there is an irreversible *ratchet effect*. The entire process represents the evolution of the system from the initial configuration (all agents have state value level_1) to a configuration in which all agents have reached the target state value level_L . For each agent, we define his own performance as his learning-time that is the number of time steps necessary to reach the target skill level.

The model is build on the assumption that knowledge is accomplished following a linear path passing through intermediate levels. You have to be aware that in real life this is not always the case; even fixing level_1 and level_L there could be many different paths joining them and each learner can be more comfortable with one path than another one. Although this assumption induces a limitation, it is necessary, at least in a first approach, to highlight the exogenous causes of the peer learning process.

Independent learning

This part of the model captures the ability for learners to improve their skill level by studying the material, or practicing, on their own. Independent learning thus refers to the ability of any given learner to improve his skill level by one during one-time step; the probability p for such an improvement to occur is fixed and identical for every learner.

For instance, if p is set to 0.1, each learner has a one-in-ten chance of improving his skill level during one-time step. This boils down to any given learner improving one skill level every 10 time steps, on average. So, if we consider 10 time steps to represent the duration of the learning episode (e.g. a semester-long course), then a value for p that is much lower than 0.10 would indicate that many students will not achieve mastery of the skill being taught by the end of the course; conversely, if p is

much higher than 0.10, most students will achieve mastery by the end of the learning episode.

To model the independent learning process we present two approaches based, first on probabilistic equation, then on agent-based modelling.

Probabilistic equation model

For each person, the learning process is independent and follows a *Negative binomial distribution*. Such distribution, pertains to the trial number t at which the first L successes have been obtained, each trial being the realization of a *Bernoulli* variable with success probability p . Therefore the probability for a learner to reach the target value level _{L} at step t is:¹

$$\binom{t-1}{L-1} p^L (1-p)^{t-L} \quad (1)$$

Following Wolfram [43] we deduce that a learner reaches the target level _{L} on average after $\frac{L}{p}$ time steps with a standard-deviation of $\frac{\sqrt{L \times (1-p)}}{p}$.

Agent-based model

First and foremost, let us recall that with independent learning a single run entails the simulation of M independent processes. For each learner, the neighbourhood is not taken into account and thus the function Φ takes its values from Σ only. During each time step t , the level of a learner a_i is updated on the basis of a stochastic dynamics: a random variable x_p is generated on the uniform interval $[0..1]$ and compared to the p value; so, the probabilistic transition function Φ is fully defined by:

$$\text{state}_i^{t+1} = \begin{cases} \text{succ}(\text{state}_i^t) & \text{if } (x_p < p) \\ \text{state}_i^t & \text{else} \end{cases} \quad (2)$$

Let us note that, once a learner reaches the top level, Φ is the identity function.

Peer learning

Now, we combine the ability for each learner to improve by himself his skill level with the ability to increase this one by interacting with other learners. As peer learning is described as a way of moving beyond independent to mutual learning [10], we are particularly interested in exploring the extent to which interactions can improve the learning process both from the point of view of each participant but also at the level of the group considered as a whole. Obviously, this objective will be all the more meaningful in that the learners have a measured capacity to progress by themselves.

¹ Note that formula (1) holds true only if level₁ = 0.

Peer learning rules

The $\mathcal{FPL}$ model captures interactions between learners. Any given agent a_i may only engage in peer learning with agents a_k checking the following conditions:

1. $a_k \in N_i$
2. a_k features a higher skill level than a_i (i.e. $\text{state}_i < \text{state}_k$);
3. a_k is 'within the reach' of a_i

More knowledgeable other

To specify the third condition, we will refer to the notion of peer formalised by the developmental psychologist *Lev Vygotsky* using the concept of *More Knowledgeable Other* (hereinafter MKO) [2, 13, 35, 38].

Vygotsky suggests that knowledge is developed through social contact and that learning takes place through interactions between teachers and students as well as between students themselves² [23, 39]. In a real learning scenario, a MKO is anyone who can help a learner with regard to a particular task; such a person may be a teacher, a parent, an older adult, a coach or a peer. Following this, the MKO of a learner is defined as the subset of all other learners that may help him. We propose to specify the MKO according to the following educational strategy: let δ an integer larger or equal to 1, for each agent-learner a_i his MKO is the set:

$$\text{MKO}_i = \{a_k \in N_i \mid \text{state}_k = \text{succ}^\delta(\text{state}_i)\} \quad (3)$$

The integer parameter δ means an iterate with δ steps; it will be referred to as the *level gap*; for simplicity, it is assumed that its value is identical for all learners. This strategy means that a neighbour (condition 1) can help if he is better (condition 2) with a skill level gap equal to δ (condition 3). Although this strategy is logical and understandable way, once applied by all learners, the consequences are difficult to predict.

² In this paper, we consider interactions between students only.

Peer learning dynamics

During each time step t , if a given agent a_i did not already improve his skill level *via* independent learning, thanks to peer learning he may still do so if, at least, there is one other agent in MKO_i^t . In such a case, his level can be improved *via* a new process based on both a fixed probability q (assumed to be the same for all learners) and the number of MKOs among his neighbours; the idea is that the more neighbours who can help him a learner has, the more likely he is to progress. This new process supplements the above-described independent learning dynamic and captures the idea of peer learning.

So the $\mathcal{FPL}$ model is based on initially setting the population of agents, each with the initial state level₁. The dynamics then proceeds in a series of discrete time steps. During each time step t , the agent's states are updated simultaneously based on two stochastic processes: for each agent a_i , two random variables x_p and x_q are generated on the uniform interval $[0; 1]$ and compared to p and q values respectively; so, the probabilistic transition function Φ is fully defined by:

$$\text{state}_i^{t+1} = \begin{cases} \text{succ}(\text{state}_i^t) & \text{if } x_p < p \\ \text{succ}(\text{state}_i^t) & \text{if } x_p \geq p \text{ and } x_q < 1 - (1 - q)^{|\text{MKO}_i|} \\ \text{state}_i^t & \text{else} \end{cases} \quad (4)$$

The simulation of a complete learning episode consists of repeating such elementary step until all agents reach the target state value level_L. It should be noted that as soon as an agent has reached the maximum level, he can no longer progress, but can continue to help those who are still learning.

To supplement the description of the model, algorithm 1 provides the pseudo-code and to ensure reproducibility of results, the full source code will be available upon request from the authors. Let's note that the model implements well the *principle of equity* because (i) initially all agents have the same state value; (ii) the three parameters p, q, δ that control the learning processes are identical for all agents; (iii) and the peer network is static.

Algorithm 1 $\mathcal{FPL}$ model (variables and pseudo-code)

```

global [
   $M$                                 ▷ number of learners
   $Pn$                                 ▷ peer network
   $L$                                 ▷ number of levels
   $p$                                 ▷ independent learning probability
   $q$                                 ▷ peer learning probability
   $\delta$                                 ▷ level gap
]

```

learners-own [$state$; $learning-time$; MKO]

Require: for each learner i , $state_i = level_1$, $learning-time = 0$

```

procedure MAIN
  while ( $\exists$  one learner  $k$  with  $state_k \prec level_L$ ) do GoONESTEP
  end while
end procedure

procedure GoONESTEP
  for each learner  $k$  with [ $state_k \prec level_L$ ] do UPDATE( $MKO_k(Pn, \delta)$ )
  end for
  for each learner  $k$  with [ $state_k \prec level_L$ ] do LEARN( $k$ )
  end for
end procedure

procedure LEARN( $k$ )
  if ( $randomFloat\ 1.0 < p$ ) then IMPROVE( $k$ )
  else
     $helped? \leftarrow false$ 
    for each learner in  $MKO_k(Pn, \delta)$  do
      if ( $randomFloat\ 1.0 < q$ ) then
         $helped? \leftarrow true$ 
      end if
    end for
    if ( $helped?$ ) then IMPROVE( $k$ )
    end if
  end if
   $learning-time \leftarrow learning-time + 1$ 
end procedure

procedure IMPROVE( $k$ )
   $state_k \leftarrow succ(state_k)$ 
end procedure
Ensure: for each learner  $i$ ,  $state_i = level_L$ 

```

Entities and variables

The model parameters can be seen as the variables of a particular entity named *observer*; the value of such a variable is fixed before an execution and does not vary during the learning process [22]. Table 1 summarizes the “observer variables”. The other entities are the *learners*; each has its own variables which can vary during the learning process. Table 2 summarizes the “learner variables”.

Table 1 Observer variables (global parameters)

Name	Meaning	Independent	Ind. + peer
M	Number of learners	1024	1024
L	Number of levels	50	50
p	Ind. learning probability	]0; 1[	0.3
P_n	Peer network	–	Lattice/SFN
q	Peer learning probability	–	]0; 1[
δ	Level gap	–	[1; 20]

Table 2 Learner variables

Name	Meaning	Value
State	Current level	[1; L]
Learning-time	Current time steps	$\mathbb{N}^*$
MKO	Current more-knowledgeable-other	Set of learners

Netlogo simulations

In the following sections we will present simulations of the $\mathcal{FPL}$ model. Experiments will be performed with an implementation of the model in the *NetLogo* multi-agent programmable environment [3, 41]. The observer entity corresponds to the agent-observer and the learner entities to the agent-turtles; so, in *Netlogo* code, the observer-variables will be the *global* variables and the learner-variables the *turtles-own* variables.

Scales

We will refer to as *one single run*, the process starting with all the skill levels set to level₁ and ending when all the learners has reached the target value level _{L} .

As the model is probabilistic, all presented quantitative results will be averaged over 100 runs, unless otherwise noted. We use the *coefficient of variation* (hereinafter *cv*), that is the ratio of the *standard deviation* of a sample to its *mean*, to choose this sample size. As proposed by Lee et al. [26], “the sample size at which the difference between consecutive *cv*’s falls below a criterion, and remains so is considered a minimum number of runs”. For example, with a grid network, with $p = 0.3$, $q = 0.3$ and $\delta = 4$, as the outcome drawn from sample sizes in {10, 100, 500, 1000}, yields *cv* in {0.0042, 0.0045, 0.0045, 0.0045}, 100 runs is a reasonable choice [27]. To avoid effects due to a small sample size, the number of agent-locations M will be oversized to 1024.

Global measures of performance

For each learner, his own performance is defined as his *learning-time* that is the number of time steps necessary to reach the target skill value level_L from level₁.

As the dynamics can lead to heterogeneous groups with respect to level, simulations will focus on monitoring two aggregate measures over the entire population:

- The *learning cost* (hereinafter cL) is the *mean* of the *learning-time* over all the learners.
- The *exclusion cost* (hereinafter cE) is the *standard-deviation* of the *learning-time*.

Of course, to know whether the mean and standard deviation are representative, it will be necessary to look at the learning time distribution; if this is indeed the case, the first measure will be a good indicator of learning effectiveness for the group considered as a whole, whereas the second will be an indicator of the extent of dropping out related to the number of learners who are significantly behind their peers and/or also a measure of overachievers. Ideally, we would like to minimize both cL and cE ; with self-directed learning, the cost depends on p only, while with peer learning, costs are based on p , q and δ .

Simulation with independent learning

To obtain a basis for comparison, we begin by simulating independent learning alone. The parameters used are described in table 1 (column 2).

Figure 1 shows the influence of the probability p on the independent learning capabilities (p varies from 0.1 to 0.9 with a 0.1 step). It can be observed that experimental and theoretical values fit very well³: as p increases, (i) the *Learning-cost* decreases as $\frac{L}{p}$ and (ii) the *Exclusion-cost* decreases as $\frac{\sqrt{L \times (1-p)}}{p}$. With low independent learning capability (say $p \leq 0.3$), in spite of an initial homogeneous population, it can be observed a kind of *learning drift* which leads to heterogeneous levels, and to a relatively long learning duration.

One must remain aware that the model is a simplification of independent learning in real life where things are more complicated. For example, the successive attempts to learn something depend on the previous ones, the second time you deal with a homework because you did not achieve to complete it the first time, you do not start from scratch because you already have thought about it. However our aim here is not so much to *stick to reality* as to show the value of peer learning. On the basis of these initial results, we are therefore now going to investigate how to improve the overall performance of a group of learners by enabling peer learning interactions.

³ In fact, this comes as no surprise because this result just confirm the negative binomial law

Fig. 1 Independent learning alone: influence of the probability p $M = 1024$; $L = 50$ (mean over 1000 runs)

Simulation with peer learning

Here, we consider the complete $\mathcal{FPL}$ model with both independent and peer learning.

To determine the extent to which peer learning interactions may reduce the learning and the exclusion costs, we will refer to the concept of *Zone of Proximal Development* (hereinafter ZPD) elaborated by Vygotsky [17, 40]. Educational research literature defines the ZPD as the difference between the ability of learner to perform a specific task under the guidance of his MKO and the learner's ability to do that task independently; basically, the theory explains that learning occurs in ZPD [42]. Research inspired by this concept, and its more recent generalization to *Scaffolding* techniques [44], is therefore of particular relevance when attempting to sketch a model of peer learners interactions. By taking inspiration from this concept, $zpdL$ and $zpdE$ are defined as the respective *gain* due to peer interactions in the costs of learning and exclusion:

$$zpdL(p, q, \delta) = cL(p, 0, \delta) - cL(p, q, \delta) \quad (5)$$

$$zpdE(p, q, \delta) = cE(p, 0, \delta) - cE(p, q, \delta) \quad (6)$$

Obviously, without peer learning (i.e. $q = 0$) the two gains are null. As the aim is to minimize the learning cost while avoiding wildly differing skill levels, we have to both maximize the two gains $zpdL$ and $zpdE$.

Unless otherwise noted, parameters are set as described in table 1 (column 3). By setting p to 0.3, we will consider situations where the independent learning capabilities are relatively low. There are two global parameters to monitor the capability to learn from the peers: for each learner a_k , (i) the level gap δ which allows to

determine at each time step t the set MKO_k^t , and (ii) the probability q to learn from a peer in MKO_k . Of course, the MKO depends on the underlying peer network also; in the following two kinds of network will be considered, first a lattice then a scale-free structure.

Peer learning on a lattice

To model an homogeneous space zone like a classroom a 2-D spatial lattice is considered [4].

Regular network

A regular square lattice of 32×32 cell-agents is used where each cell represents a learner. To avoid some side effects and to guarantee the homogeneity of the agents, *periodic boundary conditions* is imposed (i.e the world wraps both horizontally and vertically); so, for each learner, the neighbourhood is composed of the agents around him and his *degree* is invariant with value 8.⁴

Learning-time distribution

First of all, we have to look at the experimental learning-time distribution. Figure 2 plot the experimental distributions for $p = 0.3$, $q = 0.3$ and different values of the δ parameter ($\delta \in \{1; 4; 7; 10\}$).

It can be observed that in all cases, except for $\delta = 1$, the distributions fit well a theoretical *Normal distribution* with the same average and the same standard-deviation (red curves in Fig. 2); let us note that the mean of the distribution increases with δ . As the distribution is symmetrical the drop-out and the over-performing learners are in the same order of magnitude.

On the contrary, with $\delta = 1$, the distribution is asymmetrical with a *tail* for high values: there are many over-performing learners and few drop-out learners and the vast majority of learners have a learning-time close to the average (Fig. 2a).

ZPD versus MKO

Figure 3 plot the values of the Zone-of-Proximal-Development *versus* the More-Knowledgeable-Others for different values of the q parameter. Let's remember that the MKO is characterized by the level gap δ (Eq. 3). For $q = 0.1$, results are presented by a bar graph and for others q values by lines; the horizontal line ($y = 0$) corresponds to $q = 0$ and serves as a reference.

⁴ The Moore neighbourhood is used.

Fig. 2 Lattice network: learning-time distributions (Normal distribution in red). $M = 1024$; $L = 50$; $p = 0.3$; $q = 0.3$; $\delta \in \{1; 4; 7; 10\}$ (significant runs)

- It can be observed that the gain $zpdL(q, \delta)$ in learning cost decreases with δ to reach a value close to zero for $\delta = 20$; this is true for all the non-zero values of q (Fig. 3a).
- For exclusion, the situation is quite different and even unexpected since, if $q < 0.6$, the gain first increases with δ up to a maximum, then gradually decreases (Fig. 3b). The value of δ for which $zpdE(q, \delta)$ reaches its maximum will be noted $\delta_{opt}(q)$; for instance, $\delta_{opt}(0.1) = 4$ (see the bar graph on Fig. 3b).

Peer learning on a scale-free network

Here the peer network is not based on spatial proximity but rather on social relationships. Previously we have made the assumption that the peer network is defined in such a way that every learner-node has the same number of neighbours but this

Fig. 3 Lattice network: ZPD vs. MKO with $M = 1024$; $L = 50$; $p = 0.3$; $q \in \{0; .1; .2; .3; .6\}$ (mean over 100 runs). The gain in learning-cost decreases with δ to reach a value close to zero; if $q < 0.6$, the gain in exclusion-cost first increases with δ up to a maximum, then gradually decreases

disregards many situations where real networks do not share this feature. In particular, in the context of e-learning or online learning the peer network rather looks like a relational network.

We are aware that this leads to a violation of the principle of strict equity because the position of the agents in the network, will differentiate one learner from the other; although the agents have the initial same level and progress from level to level according to the same laws, it will be interesting to study the influence of the *degree* on the learning process and, in particular, to highlight the role played by the *hubs* in such a dynamics.

Scale-free network

As many relational networks are scale-free network (hereinafter *SFN*), the constraint is relaxed by studying peer networks where each node may have its own *degree* [7, 29]. Thus, it is assumed that the degree distribution follows a *power law*; that is, the fraction $n(k)$ of learners having k neighbours goes approximately as:

$$n(k) \approx k^{-\gamma} \quad (7)$$

where γ is a parameter whose value is typically in the range [2; 3]. Let us note that ABM is well appropriate to model such an heterogeneous population.

Thus some learners have a huge numbers of neighbours whereas a lot of learners have just some; the most connected are the *hubs* and the least connected are the *leaves*. In the following, questions about the role of such individuals in the learning process will be asked. Another important characteristic is that a *SFN* can be generated by a random process called *preferential attachment* [12] where new nodes attach to old ones with a probability proportional to its degree; this feature will be used to synthesize such networks for simulations (Fig. 4).

Fig. 4 A scale-free network. $M = 1024$; $\gamma \approx 2.4$; $\text{max-degree} = 75$ (learner node-size is correlated to degree)

Learning-time distribution

A clearer vision is needed for the experimental learning-time distribution; in particular, the question is whether it fits with a theoretical *Normal* distribution.

Figure 5 plot the experimental distributions for $p = 0.3$, $q = 0.3$ and $\delta \in \{1; 5; 10; 15\}$. It can be observed that (i) the distributions are quasi symmetrical and fit well a theoretical *Normal* distribution with the same average and the same standard-deviation (red curves in Fig. 5); (ii) the average increases with δ ; and (iii) the smallest standard-deviation is obtained for $\delta = 5$.

ZPD versus MKO

Figure 6 plot the values of the Zone-of-Proximal-Development *versus* the More-Knowledgeable-Others for different values of the q parameter. Results are presented with bars for $q = 0.3$ and for others q values with lines; the horizontal line ($y = 0$) corresponds to $q = 0$;

- Once again, for all values of q , it can be observed that the gain $zpdL(q, \delta)$ in learning cost decreases with δ to reach a value close to zero for $\delta = 20$ (Fig. 6a).

Fig. 5 Scale-free network: learning-time distributions (Normal distribution in red). $M = 1024$; $L = 50$; $p = 0.3$; $q = 0.3$; $\delta \in \{1; 5; 10; 15\}$ (significant runs)

- For exclusion, the situation is yet again unexpected since, if $q < 1$, the gain $zpdE(q, \delta)$ first increases with δ up to a maximum, then gradually decreases; for instance, $\delta_{opt}(0.3) = 5$ (Fig. 6b).

Discussion

Comparing a regular lattice with a scale-free network, results are qualitatively equivalent: the learning cost decreases with δ while the exclusion cost first increases and then finally decreases; as expected, the gains provided by the peer learning process is cancelled out for high δ values. From a quantitative point of view, however, it can be noted that the gains due to peer learning are less important with a scale-free than a regular network.

Fig. 6 Scale-free network: ZPD vs. MKO with $M = 1024$; $L = 50$; $p = 0.3$; $q \in \{0; .3; .6; .8; .95\}$ (mean over 100 runs). The gain in learning-cost decreases with δ to reach a value close to zero for $\delta = 20$; if $q < 1$, the gain in exclusion-cost first increases with δ up to a maximum, then gradually decreases

In both cases, the key point is that the gap value δ_{opt} that maximizes the gain $zpdE$ (and thus minimizes the cost of exclusion) leads to a quite long mean learning-time over the population. It is therefore impossible both to maximise the overall learning performance and to minimise the phenomenon of exclusion.

The scale-free network topology reinforces this phenomenon as (i) for $\delta < 3$ the gain $zpdE(q, \delta)$ is even negative—it means there is more exclusion with peer learning than without⁵ – and (ii) $zpdE(q, \delta)$ reaches its maximum for a higher δ value (e.g. $\delta_{opt}(0.3) = 5$ instead of 3 for a lattice). An explanation for all this can be found by looking more closely at the very structure of the network: in a regular lattice all the agents have the same degree and therefore play the same game whereas this is no longer the case with a *SFN* where there are few hubs with huge numbers of neighbours and many leaves with one neighbour only. In a regular lattice, due to its homogeneity, knowledge diffusion may be *isotropic*; in contrast in a scale-free network, as the hubs are obligatory crossing points, some sub-networks will be favoured while others will be penalized; of course, this is an aggravating factor that promotes exclusion.

Figures 7, 8, 9 and 10 are related to *SFN*; they show the impact of the degree on the peer learning process. Figures 7 and 8 display a scatter-plot where each point corresponds to one learner at the end of the process with coordinates (degree, learning-time). Figure 9 show, for each time step t , the current level of the learners.

- In Fig. 7, the independent learning probability p and the peer learning gap δ are fixed respectively to 0.3 and 1 and the peer learning probability q is in

⁵ However, this does not mean that the worst performances with independent learning simulations are better than the worst performances in the case of peer learning.

Fig. 7 Scale-free network: learning-time vs. degree with $M = 1024$; $L = 50$; $p = 0.3$; $q \in \{.3; .6; .8; .9\}$; $\delta = \delta_{\min} = 1$ (Four significant runs). red line = mean learning-time; blue lines = mean + $0.99 \times$ standard deviation and mean - $0.99 \times$ standard deviation

the set $\{0.3; 0.6; 0.8; 0.9\}$. Results show that, regardless of q , the minimal gap ($\delta = 1$) favours as a priority the hubs in such a way they learn very early and so become quickly unreachable to help other learners.

- In Figs. 8 and 9, p and q are fixed respectively to 0.3 and 0.6 (thus $\delta_{\text{opt}} = 5$) and δ is in the set $\{1; 5; 10; 20\}$. Results confirm that the gain in learning-time decreases with δ . More, it can be observed that (i) for $\delta = \delta_{\min} = 1$ the hubs learn very early and thus become quickly unable to help other learners; (ii) for the optimal value $\delta = \delta_{\text{opt}} = 5$, the hubs always learn first, but here there is a ripple effect on the latecomers, resulting in far fewer drop-outs; (iii) for $\delta = 10$ and $\delta = 20$ the peer learning effect is low and there is a weak correlation between degree and level. Let us note that the max degree is approxi-

Fig. 8 Scale-free network: learning-time vs. degree with $M = 1024$; $L = 50$; $p = 0.3$; $q = 0.6$; $\delta \in \{1; 5; 10; 20\}$ (Four significant runs). red line = mean learning-time; blue lines = mean + $0.99 \times$ standard deviation and mean - $0.99 \times$ standard deviation

matively of 75 and we consider a learner to be a hub as soon as his degree is above 30.

Finally, a 3-D animation of the learning process is made: the scale-free network is initially created on the bottom x - y plane ($z = 1$); then, during the learning process agents are climbing from the plane $z = 1$ to plane $z = L$; for each agent, at each time step, the z -axis represents his current state. Figure 10 display two snapshots of such an animation at a given time; it illustrates the role played by hubs in the peer learning process. If $\delta = \delta_{\min} = 1$, there are a strong drop-out for low degree agents because the hubs have progressed very fast (Fig. 10a). In contrast, with $\delta = \delta_{\text{opt}} = 5$,

Fig. 9 Scale-free network: level vs. time for hubs (blue points; degree ≥ 30) and leaves (red points; degree = 1). $M = 1024$; $L = 50$; $p = 0.3$; $q = 0.6$; $\delta \in \{1; 5; 10; 20\}$ (significant runs)

everyone climbs forward together and there is small differences in level between agents throughout the peer learning process (Fig. 10b).

Conclusions and future work

This paper proposed an agent-based approach to model peer learning dynamics occurring in a group of learners either positioned in a classroom or member of a inter-related group involved in a distance learning session. We start from the premise that to capture *the essential characteristics* of the learning process, eliminating additional effects, it is useful to develop and simulate simplified models consisting of totally or partially homogeneous entities. Here, the aim was not to

(a) $\delta = \delta_{min} = 1$

(b) $\delta = \delta_{opt} = 5$

Fig. 10 3-D Snapshot of the scale-free population at a given time during the learning process (z -axis represents the level and node-size is correlated to degree). $M = 500$; $p = 0.3$; $q = 0.95$ (Two significant runs)

minimize, let alone deny, the role of intelligence, family environment, social status... in the learning process but rather to explore the causes that may be basically inherent in the very process of knowledge diffusion in a network of peer learners [37]. Although it is relevant to look for the causes of exclusion in factors that differentiate between people, agent-based simulations of the $\mathcal{FPL}$ model have shown that this phenomenon may also arise solely on the basis of the learning process itself.

Even though the model is probabilistic, it assumes that the learning process is strictly equitable in the meaning that all learners are equivalent in terms of their initial skill levels, and also in the way they improve their level. As it is impossible to set up in a sustainable way such a strict process in real life, the agent-based modelling approach is one possibility to explore intrinsic properties of a peer learning strategy while eliminating confounding factors related to individual characteristics.

Although the value of peer learning is now recognised and its practical implications have already been considered [1, 34], it is still important that models and computational simulations suggest ways in which learning strategies can be put into practice. A crucial parameter that controls the exclusion phenomenon is the *level gap* that represents the acceptable skill level difference allowing one learner to help another. On this basis, the simulations produced some tangible results:

- there is an optimal value for the level gap to avoid exclusion; but, as this occurs at the expense of the mean learning-time, a *dilemma* arises: one could be tempted to prioritize global performance, and thus risk leaving some learners in the lurch, or, conversely, avoid excluding learners at the cost of diminishing the global performance;
- to prevent exclusion, a learner should never seek peers whose skill level is exactly one level above his own;
- all other things being equal, the exclusion phenomenon is more pronounced with a scale-free than a lattice network;
- In a scale-free network the hubs play a central role in the way knowledge spreads among learners and this role can be adequately monitored by means of the level gap.

The results presented in this paper open some perspectives for future works:

- First, the impact of static characteristics such as the definition of “acceptable peers” and the p and q probabilities needs to be examined in more detail; for example, one could relax the constraint on the uniqueness of the parameters p , q and δ (i.e. rather consider distributions of values over the population).
- The successive attempts to learn something depend on the previous ones, the second time you deal with a homework you do not start from scratch because you already have thought about it; to take this phenomenon into account the probability p should be a decreasing function of the number of steps.
- Although the $\mathcal{FPL}$ model allows *reciprocal peer learning*—i.e. all agents are learners, and all may potentially help another neighbour as peer—situations

- including *teacher* or *pedagogical virtual agent* interactions need to be considered [9, 24, 30]. The two dynamics are complementary and should interact in synergy.
- Because in the real world agents are non homogeneous, it would be interesting to compare the results with some heterogeneity (for instance, a *Gaussian* or a *Poisson* distribution for the initial level of agents) and discuss as the latter will impact the outcome.
 - Although here a simple scale-free network is used, in future work it will be useful to study the influence of the γ parameter and the presence of loops in the network.
 - More broadly, one of the interests of these results is to suggest ways for real learning strategies that make it possible to control the deleterious effect of dropping out while improving the overall performance of the learning process. From this perspective, what should serve as a guideline is *individualization* for micro-decisions and real-time *adaptability* for the learning network. Beyond the approaches which consist in favouring individuals in difficulty and/or slowing down the most advanced, this could be done by individualizing and managing the learning gap δ , for example, by correlating it to individual degree and/or level. In the same way, the learning network could usefully evolve under impetus and/or advice from educators through the implementation of a dynamic *relocation* strategy of the learners [19].

References

1. Abdu, R., & Schwarz, B. B. (2020). Split Up, but Stay Together: Collaboration and cooperation in mathematical problem-solving. *Instructional Science*.
2. Abrahamson, D., & Wilensky, U. (2005). *Piaget? Vygotsky? I'm game!: Agent-based modeling for psychology research*. Vancouver: Annual meeting of the Jean Piaget Society.
3. Abrahamson, D., Wilensky, U., & Levin, J. (2007). *Agent-based modeling as a bridge between cognitive and social perspectives on learning*. Chicago: Annual meeting of the American Educational Research Association.
4. Abrahamson, D., Blikstein, P., & Wilensky, U. (2007). Classroom model, model classroom: Computer-supported methodology for investigating collaborative-learning pedagogy. *Proceedings of the Computer Supported Collaborative Learning Conference (CSCL)*, 8(1), 46–55.
5. Abrami, P. C., Poulsen, C., & Chambers, B. (2004). Teacher motivation to implement cooperative learning: Factors differentiating users and non-users of cooperative learning. *Educational Psychology*, 24, 201–216.
6. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Larusson & B. White (Eds.), *Learning analytics*. New York: Springer.
7. Barabási, A.-L., & Bonabeau, E. (2003). *Scale-Free Networks* (pp. 50–59). Elsevier: Scientific American.
8. Bordogna, C., & Albano, E. (2001). Theoretical description of teaching-learning processes: A multidisciplinary approach. *Physical Review Letters*, 87, 118701.
9. Bordogna, C., & Albano, E. (2002). A cellular automata model for social-learning processes in a classroom context. *European Physical Journal B*, 25(3), 391–396.
10. Boud, D. (1988). Moving towards autonomy. In D. Boud (Ed.), *Developing student autonomy in learning*. London: Kogan Page.
11. Burini, D., DeLillo, S., & Gibelli, L. (2015). Collective learning modeling based on the kinetic theory of active particles. *Physics of Life Reviews*, 16(1), 123–139.

12. Clauset, A. (2011). The preferential attachment mechanism, inference, models and simulation for complex systems: CSCI 7000-001 *Lecture*, October (20). http://tuvalu.santafe.edu/~aaronc/courses/7000/csci7000-001_2011_L14.pdf.
13. Chailin, S. (2003). The Zone of Proximal Development in Vygotsky's analysis of learning and instruction, *Vygotsky's educational theory and practice in cultural context*, pp. 39–64, Kozulin, A., Gindis, B., Ageyev, V. & Miller, S. (Eds.), Cambridge University.
14. Collard, P. (2019). λ CA: A peer learning cellular automaton. *Journal of Cellular Automata*, 14(3–4), 263–288.
15. Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69, 970–977.
16. Crouch, C. H., Watkins, J., Fagen, A. P., & Mazur, E. (2007). Peer instruction: Engaging students one-on-one, all at once. *Research-based reform of university physics*, 1(1), 40–95, American Association of Physics Teachers College Park.
17. Fani, T., & Ghaemib, F. (2011). Implications of Vygotsky's zone of proximal development (ZPD) in teacher education: ZPTD and self-scaffolding. *Procedia-Social and Behavioral Sciences*, 29, 1549–1554.
18. Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A componential analysis. *Journal of Educational Psychology*, 81, 173–177.
19. Fernandes, A. C., Huang, J., & Rinaldo, V. (2011). Does where a student sits really matter? The impact of seating locations on student classroom learning. *International Journal of Applied Educational Studies*, 10(1), 66–75.
20. Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences*. 22(4), 521–563.
21. Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, 81, 371–383.
22. Grimm, V., et al. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*. <http://jasss.soc.surrey.ac.uk/23/2/7.html>.
23. Jacobs, G., Hurley, M., & Unite, C. (2008). How learning theory creates a foundation for SI leader training. *Journal of Peer Learning*, 1, 6–12.
24. Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. *International Journal of Artificial Intelligence in Education*, Springer, New York, 26(1), 25–36.
25. Koponen, I. T. & Nousiainen, M. (2018). An agent-based model of discourse pattern formation in small groups of competing and cooperating members, *Journal of Artificial Societies and Social Simulation*, 2(1).
26. Ju Sung, L., et al. (2015). The complexities of agent-based modeling output analysis. *Journal of Artificial Societies and Social Simulation*, 18(4), 4. <http://jasss.soc.surrey.ac.uk/18/4/4.html>.
27. Lorscheid, I., Heine, B.-O., & Meyer, M. (2012). Opening the 'black box of simulations: Increased transparency and effective communication through the systematic design of experiments. *Computational and Mathematical Organization Theory*, 18, 22–62.
28. Macal, C. M., & North, M. J. (2005). Tutorial on agent-based modeling and simulation. *Proceedings of the 2005 Winter Simulation Conference (IEEE Cat. No. 05CH37732C)*.
29. Newman, N. (2010). Networks, an introduction, *Oxford University Press*, ISBN="9780199206650".
30. Ormazábal, I., Borotto, F. A., & Astudillo, H. F. (2021). An agent-based model for teaching-learning processes. *Physica A*, 565, 125563.
31. Primiero, G. (2019). A minimalist epistemology for agent-based simulations in the artificial sciences. *Minds and Machines*., <https://doi.org/10.1007/s11023-019-09489-4>.
32. Smith, E. R., & Conrey, F. C. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, 11, 87–104.
33. Squazzoni, F. (2012). Agent-based computational sociology, isbn: 978-0-470-71174-3, Ed. Wiley.
34. Stahl, G. (2015). A decade of CSCL. *International Journal of Computer-Supported Collaborative Learning*, 10(4), 337–344.
35. Sundararajan, B. (2010). Emergence of the most knowledgeable other (MKO): social network analysis of chat and bulletin board conversations in a CSCL System. *A CSCL System Electronic Journal of e-Learning*, 8(2), 191–208.

36. Sun, A. Q., & Xiufang, C. (2016). Online education and its effective practice: A research review. *Journal of Information Technology Education: Research*, 15, 57–190.
37. Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, 25(6), 631–645.
38. Van der Veer, R., & Valsine, J. (1991). *Understanding vygotsky, a quest for synthesis*. Oxford: Basil Blackwell.
39. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. *M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Eds. Cambridge, MA: MIT Press*.
40. Warford, M. K. (2011). The zone of proximal teacher development. *Teaching and Teacher Education*, 27(2), 252–258.
41. Wilensky, U. (2019). Center for connected learning and computer-based modeling, Northwestern University, Evanston, IL. <http://ccl.northwestern.edu/netlogo/>.
42. Woolfolk, A. (2004). *Educational psychology (9Nd Edition)*. Boston: Allyn and Bacon.
43. Wolfram Math World. (2000). *DiscreteDistributions*. <http://mathworld.wolfram.com/NegativeBinomialDistribution.html>.
44. Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Child Psychiatry*, 17, 89–100.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.