Assessing Group Composition in e-learning According to Vygotskij's Zone of Proximal Development

Maria De Marsico[1], Andrea Sterbini[1], and Marco Temperini[2]

[1] Dept. of Computer Science, Sapienza University of Rome, Italy
{demarsico,sterbini}@di.uniroma1.it
[2] Dept. of Computer, Control, and Management Engineering
Sapienza University of Rome, Italy
marte@dis.uniroma1.it

Abstract. In this paper we build on previous work exploring a formal way to assess the composition of learning groups. We start from our existing framework, designed to provide support to personalization in e-learning environments, comprising an implementation of the Vygotskij Theory of proximal development. In such theory, effective individual learning achievements can be only obtained within the boundaries of a cognitive zone where the learner can proceed without frustration, though with support from teacher and peers. In this endeavor, the individual development cannot disregard social-collaborative educational activities. Previously we gave operative definitions of the Zone of Proximal Development for both single learners and groups; here we aim at assessing the viability of a partition of students in groups over a common task.

Keywords: Individual Zone of Proximal Development, Group Zone of Proximal development, Personalized learning path, Social collaborative e-learning.

1 Introduction and Motivations

Recent research on e-learning has especially focused on personalization and adaptivity. Several investigations mainly addressed tailoring the learning experience to personal-individual characteristics (such as learning styles and achievements) [1-5]. However, also planning and coordination of work in groups has deserved attention from recent related research [6-10], due to the increasing interest in investigating effective strategies to motivate and enhance study activity through social-collaborative learning [11,12]. The present work introduces a first attempt to formally define a "viability" measure for the composition of learning groups. It is a further step along our research line aiming at the concrete implementation of the principles underlying the concept of Zone of Proximal Development (ZPD) from Vygotskij theory [13] in a platform for adaptive e-learning.

The present discussion stems from two previous achievements. In [14] we presented a framework for the dynamic configuration of personalized learning paths for both individual students and groups, in a way adaptive to a continuous assessment and

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2014, Part II, LNCS 8514, pp. 277–288, 2014.

student model updating. Such framework can be integrated in present state-of-the-art e-learning platforms, such as Moodle [15], to avoid much of the cumbersomeness of the implementation of an entirely new system. Later [16], we focused on refinements related to individual profiling. The quality of personal achievements (skill acquisition) can vary according to features which are not presently accounted for in widespread learning environments, such as the time required to acquire a skill, or the linearity (rate of success/failure results) of the acquisition. Moreover, the difficulty of an activity is determined both by the skills required to tackle it, and by the quality/firmness of their possession by the student. Those factors too may vary from student to student (or between different phases in the development of the learner, as mirrored by the evolution of the individual student model). Quite straightforwardly, all this changes the level of "difficulty" which each single student is subjected to.

If considering these aspects may be of paramount importance in individual personalization, it deserves even more attention when composing working groups. From one side, we can "exploit" the definition of the group ZPD, and from the other side we can include into it a group-bounded version of the formal definitions of "difficulty" which might be actually encountered in tackling a learning activity. In many cases the group composition follows personal preferences of students with respect to their mates. However, from a pedagogical point of view such choices might not be the most appropriate ones. Members that are too "strong" may be demotivated by having to adapt to "weakest" ones, and, symmetrically, less bright members of the group may be frustrated by having to forcedly follow the smarter colleagues. And of course grouping and separating "brighter" students from "weaker" may be detrimental for the class as a whole and leave too many students behind.

While in face-to-face activities the experience and sensibility of the teacher might guide the choice of group partition, a similar guidance is very hard to achieve automatically in distance settings. We argue that a suitable extension to groups of the formal strategies applied to individual students in our previous work can help "measuring" the appropriateness of a given partition in groups of a class of students. Rather than exploiting any "viability" measure for group composition in order to determine from scratch the best partition of the class (as it might be computationally heavy), we try to define such an analysis on already stated class subdivision. We intend to tackle this problem in a formal way, within our e-learning framework. In this endeavor, the value of such framework is twofold: it is detached by any present concrete e-learning platform, in particular from the Moodle prototype presented in [9], and provides a ground layer to realize the pedagogical principles of the theory of Vygotskij, by extending the formal bases for the practical implementation of concepts such as the Zone of Proximal Development in an e-learning setting.

2 Framework Core Definitions

We report here for reader's convenience only the most basic concepts that make up the formal definition of our framework. Further details can be found in [14].

2.1 The Learner and the Learning Activities

A learner *l* is represented in the system by a student model, that in this barebones description is constituted by a set of his/her achieved *skills* (Student Knowledge – *SK(l)*). A skill [14] represents knowledge possessed by the learner, and is qualified by a measure of *certainty* (about its possession by *l*) $c \in [0...1]$):

$$SK(l) = \{<s_1, c_1>, \ldots, <s_{nl}, c_{nl}>\}$$

A learning activity *la* is defined by: *la.Content* (learning material); *la.A* (Acquisition: skills provided by *la)*; *la.P* (Prerequisites: skills "needed" to tackle *la*); *la.Effort* (an estimate of the cognitive load associated to *la*).

The completion of a *la* by learner *l* will trigger the insertion of *la.A* into the set of skills in *SK(l)*, with an assigned *certainty* depending on the student's performance in, say, a final test. In this way, *SK(l)*, is continuously updated during course. The updates reflect the evolution of *l* following the tackled learning activities.

In particular C_{ENTRY} is the default value assigned to a newly acquired skill. Further successful assessments for *s* increase its certainty, *c*, while unsuccessful ones decrease it. When *c* in $<s, c>$ decreases below a level C_{DEMOTE} $<s, c>$ is removed, and further activities will be needed to acquire it back; a value of *c* exceeding $C_{PROMOTE}$, states that the skill is firmly acquired, and no further assessment will be required. As for C_{ENTRY}, C_{DEMOTE}, $C_{PROMOTE}$ the teacher can confirm platform defaults, or assign them differently.

A *learning path* can be defined as a set $LP=\{la_i\}_{i \in \{1...n\}}$. A certain *LP* entails an overall set of acquired skills *LP.A*, and an overall set of requirements *LP.P* such that

$$LP.A = \cup_{i\in \{1...n\}} la_i.A \qquad\qquad LP.P = \cup_{i\in \{1...n\}} la_i.P \setminus LP.A$$

and an overall effort $\quad LP.Effort = \sum_{i \in \{1...n\}} la_i.effort$

A personalized course delivery for a learner *l*, is a learning path built basing on the initial $SK(l)_{INIT}$ and adaptively updated basing on *SK(l)* evolution. In particular a student *l* is able to access a certain activity iff all skills in *la.P* are in *SK(l)*.
For SK(*l*) = $\{<s_1, c_1>, \ldots, <s_{nl}, c_{nl}>\}$, its *s-projection* is its set of skills:

$$s\text{-}proj(SK(l)) = \{s_i, \text{ with } <s_i, c_i> \in SK(l)\} = \{s_1, \ldots, s_{nl}\}$$

The relations between Prerequisite and Acquired sets induce a partial order on the learning activities: if $la.A \cap \underline{la}.P \neq \emptyset$, some skills needed by <u>*la*</u> are acquired through *la,* so that *la* has to precede <u>*la* in any learning path</u>. The framework lets the learner choose the "next learning activity" in the course as freely as possible. To be educationally feasible, this must take into account both the present learner's knowledge and the partial order among activities. To allow this, we attempt a formalization of the concept of Zone of Proximal Development (ZPD) from Vygotskij theory [13].

2.2 A Formal Definition of ZPD

Given a learner *l*, working on a configured course $\underline{LP} = \{la_1, \ldots, la_{tl}\}$, some significant cognitive areas related to student's learning state, and defined by Vygotskij, have been formalized in our framework. The area of Autonomous Problem Solving (*APS*) is the area of firm knowledge in the present state of knowledge *SK(l)*:

$$APS(l) = \{s, <s, C_{PROMOTE}> \in SK(l)\}. \text{ Of course, } APS(l) \subseteq s\text{-}proj(SK(l)).$$

The *ZPD* for the learner is a zone where (s)he has no firm achievements yet, but that can be explored with some help from the teacher or from peers. On the contrary, the zone of Unreachable Problem Solving (UPS) is the area (of the course) where it is not pedagogically safe for the learner to enter, given the present level of skills:

$$UPS(l) = \underline{LP}.A \setminus (APS(l) \cup ZPD(l))$$

Given a learning path LP, its knowledge domain is $KD(\underline{LP}) = \underline{LP}.A \cup \underline{LP}.P$. In particular the set difference, $KD(\underline{LP}) \setminus s\text{-}proj(SK(l))$ is the set of all skills in the course knowledge domain, that are not yet acquired in *SK(l)*. A subset of these skills constitutes the ZPD of the student *l*, denoted as *ZPD(l)*. We are interested in identifying such subset, as composed by those skills that are at an "affordable" cognitive distance from the present *SK(l)*. In doing this, our aim is not to pack an additional bag of skills which can possibly be acquired, but rather to take into account the genuine interpretation of ZPD as a region of cognitive *development* [17].

Firstly, for each skill *s* outside *SK(l)*, we define the set of possible learning (sub)paths *LP'* in LP, that can eventually allow to acquire *s*, and that can start from the current state of skills:

$$Reach(s, SK(l), \underline{LP}) =$$
$$= \{G=\{la_i\}_{i \in \{1...nG\}} \subseteq \underline{LP} \mid s\in la_{nG}.A \wedge G.P \subseteq s\text{-}proj(SK(l)) \cup G.A\}$$

where the last condition relating G.P to G.A expresses the possibility that the prerequisites of some $la_i \in G$ might be acquired through a previous $la_j \in G$. We define the distance of *s* from the present *SK(l)* as

$$D(s, s\text{-}proj(SK(l)), \underline{LP}) =$$
$\underline{G}.Effort$, where G is an element of minimal overall effort in *Reach*(s, SK(*l*), LP).

The set $Support(s, SK(l), \underline{LP}) = \underline{G}.P \cap s\text{-}proj(SK(l))$ denotes the skills already possessed by the learner that are necessary to reach *s* along a minimal-effort path in LP. We designate such a set as the *support set* to reach *s*.

We assume that a higher certainty for the skills in the support set can facilitate the learner in reaching *s*. Furthermore, certainty in the support set can affect the distance from the *SK(l)* that we can span, and yet still consider *s* in the *ZPD(l)*. In other words, supposing that $D(s, s\text{-}proj(SK(l)), \underline{LP}) \geq D(s', s\text{-}proj(SK(l)), \underline{LP})$ while the overall certainty of the $Support(s, SK(l), \underline{LP})$ is higher of $Support(s', SK(l), \underline{LP})$, *s* might be reachable while *s'* might not, despite the closer distance. The effort required along the way to the target has a further role in determining the maximum reasonable distance. According to these preliminary considerations, we attempted to define such distance in a reasonable yet challenging way, which may stimulate the student without causing frustration, being dynamically tuned to his/her evolving cognitive state. To take into account both average certainty of the support set and expected average effort along the path towards a certain skill, in [14] we introduced functions *AvgCertainty()* and *AvgEffort()*. In the initial definition we considered only the pure level of certainty of a skill, and the level of effort estimated by the teacher. However, in the follow-up of the creation of our framework we realized that it is not pedagogically realistic to assign to each effort (or resp. certainty) a constant weight, as the effort in acquiring a skill may depend on the firmness of required knowledge, and the certainty of a skill may be weighted by the variable paths through which it has been secured. Therefore, in [16] we refined both *AvgCertainty()* and *AvgEffort()*. A kind of more realistic "average certainty" is obtained by a weighed sum through a backward computation starting from a skill to be possibly included in *ZPD*, and going back towards its support set. Skills with different certainty may contribute differently, and different skills presenting the same certainty might contribute differently too, depending on both the consolidation of a skill in time and the ways that certainty has been reached by the learner. For a given learner *l* and a given skill s_i, with certainty c_i in *SK(l)* the weight for s_i is:

$$w_i = \frac{age(s_i) * age(cert(s_i))}{\left(age(s_i) - age\left(cert(s_i)\right)\right) * \left({}^{ntests}/_{npostests}\right)}$$

$age(s_i)$ being the age of the skill, $age(cert(s_i))$ the age of the present value of certainty, $(age(s_i)$- $age(cert(s_i)))$ an estimate of the time to reach $cert(s_i)$ and *ntests/npostests* the ratio between the number of tests and the number of positive increments of certainty, i.e., an estimate of the linearity of the learning process. *AvgCertainty()* is then:

$$AvgCertainty(Support(s, SK(l), \underline{LP})) = \left(\sum_{<si,ci> \in Support(s,\ SK(l),\ \underline{LP})} w_i \cdot c_i\right) / \ Card(Support(s, SK(l), \underline{LP}))$$

Notice that this is not a true weighted average, since the sum of weights is not 1. In a similar way, "average effort" can be obtained by a forward computation starting from a support set towards any reachable skill to be possibly included in the ZPD. Each activity has an effort value in its definition, however each student may experience a different one. We can assume that, for each *la* in *G*, a subset of skills in both *la.P* and *la.A* are already in *SK(l)* with its certainty level. Skills already possessed, both in *la.P* and *la.A* sets, can decrease the effort actually experienced, and symmetrically a poor performance in pre-requisites may increase it. As for skills in *la.A*,

$$wa(la.A) = Card(la.A) \cdot C_{ENTRY}, / \sum_{s \text{ in } la.A} f(s, SK(l))$$

where

$$f(s, SK(l)) = \begin{cases} c & if < s, c > \in SK(l) \\ C_{ENTRY} & otherwise. \end{cases}$$

Reminding that the expected level of certainty for a newly acquired skill is C_{ENTRY}, notice that if no skill is already possessed, $w(la) = 1$ and originally defined effort is still valid. On the other hand, a level of certainty already achieved, or a low value, are able to respectively decrease or increase the effort to acquire that skill. A symmetric argument holds for prerequisite skills in *la.P*:

$$wp(la.P) = Card(la.P) \cdot C_{promote} / \sum_{s \text{ in } la.P} f(s, SK(l))$$

where *f(s, SK(l))* is defined in the same way as above. Reminding that he best supportive level of certainty for a prerequisite skill is $C_{PROMOTE}$, if all the skills are already possessed with certainty $C_{PROMOTE}$ (as it should preferably be) then *wp(la.P) = 1* and effort is not affected. Values better than $C_{PROMOTE}$ means a firmer achievement and decreases effort, and the contrary. For a certain student *l* and a given activity *la*, the weight of the activity in computing the "average effort" on a learning path will be

$$w(la) = \frac{wa(la.A)+wp(la.P)}{2}$$

We can now define :

$$A1 = AvgCertainty(Support(\text{s}, \text{SK}(l), \underline{\text{LP}})) =$$

$$(\textstyle\sum_{<si,ci>\in Support(s,\ SK(l),\ \underline{LP}))} w_i \cdot c_i) \,/\, \text{Card}(Support(\text{s}, \text{SK}(l), \underline{\text{LP}}))$$

$$A2 = AvgEffort(\text{G}^{\min}, Support(\text{s}, \text{SK}(l), \underline{\text{LP}})) =$$

$$\textstyle\sum_{la\in G^{min}} w(la)\cdot la.effort \,/\, Card(\text{G}^{\min})$$

and finally

$$DTreshold(\text{s}, \text{SK}(l)) = (\text{A1/A2}) \cdot Eff(R) \cdot dF.$$

The term *A1/A2* can be considered as the amount of certainty per unit of effort which is currently available to the student, so that the higher this ratio, the farther the student can explore; *Eff(R)* is the average effort over the learning activities in the learning domain; *dF* is a daring factor that can be configured by the teacher. Finally:

$$\text{ZPD(l)} = \{\text{s}\in KD(\underline{LP}) \setminus \text{APS}(l), \text{ such that}$$
$$\text{D(s, } s\text{-}proj(\text{SK}(l)), \underline{\text{LP}}) \leq DTreshold(\text{s}, \text{SK}(l))\}$$

The final result is a *ZPD()* with a variable radius, i.e., a radius which is not the same for all students but depends on their present state of knowledge. We think that this operational definition of an individually tuned *ZPD* can support a true implementation of a zone of development, coherently with the concept originally portrayed by Vygotskij, although in a framework where the word "development" is intended as acquisition of techniques and knowledge skills from previously possessed ones.

2.3 ZPD for Groups

When we have to select an appropriate *LP* for a given group of students, we have to first determine the overall group's state of skills (Group Knowledge - *GK*), and *ZPD*, starting from the individual ones. If we identify the group *ZPD* with the pedagogically admissible set of learning activities that the group can tackle, such set should be defined so as to maximize members' gain from the collaborative activities.

We first compute the group's *GK* as the union of the members' *SK*, where each skill has group-certainty equal to its average certainty in the members' *SK*:

$$GK(ST) = \{<s,c> \,/\, \forall l\in ST\ (<s,c_l>\in SK(l) \wedge c=((\textstyle\sum_{l\in ST,\ <s,cl>\in SK(l)} c_l\ ) \,/\, Card(ST))\}$$

In order to take into consideration the possible reciprocal support in a *group-autonomous* achievement, we also modify the definition of the *APS* of the group, by considering a "pseudo-intersection": skills that are not firmly possessed by *all* members are included in the *APS(ST)* iff they are in *APS(l')* for *some* $l' \in ST$ and they are in *SK(l)* for all the other members $l \in ST$ with a minimum certainty τ_C, chosen as:

$$\tau_C = C_{PROMOTE} - C_{ENTRY}/2$$

Since the *l'* students above will support the *l ones*, they have to be sufficiently many in the group (according to teacher's advice), say one for each *g* members:

$$\mathrm{APS(ST)} = \{s \in \cup_{l \in ST} APS(l) \mid$$
$$\forall l \in ST(<s,c> \in SK(l) \wedge c \geq \tau_C) \wedge Card(\{l' \in ST \mid$$
$$<s,c> \in SK(l') \wedge c = C_{PROMOTE}\}) \geq Card(ST)/g$$

We use a reverse strategy and define implicitly the group *ZPD*, through criteria of *admissibility* of activities. Two conditions are defined, by working on the *APS(l)*s, the *ZPD(l)*s, and the SKs of the group members. As for the first one, given a group of students ST and a learning path LP, 1) *the group members must share a common portion of APS, and 2) each activity prerequisites is firmly possessed by at least one of the members*:

$$\cap_{l \in ST} APS(l) \neq \emptyset \quad \wedge \quad LP.P \subseteq \cup_{l \in ST} APS(l)$$

The second condition states that students in a group ST must share some common proximal development, and that an activity $la \in LP$ is admissible for ST iff, though possibly being off the *ZPD*s of some members, it is *not too distant* from them, and it is comprised in the *ZPD* of at least a number of members sufficient to support the others - τ is a threshold to establish admissibility, for learner *l*, of an la not in *ZPD(l)*:

$$\cap_{l \in ST} ZPD(l) \neq \emptyset \quad \wedge \quad \forall la \in LP\ \forall s \in la.A\ \forall l \in ST \quad \mathrm{D}(\mathrm{s}, \mathrm{ZPD}(l), \mathrm{LP}) < \tau$$
$$\wedge \quad \forall la \in LP\ Card(\{l \in ST \mid la.A \subseteq ZPD(l)\}) \geq Card(ST/g)$$

As above, *g* represents the number of students which can be supported by a peer. Being τ a threshold beyond individual ZPDs, i.e. beyond the daring zone for the individual learner, we set it as the minimum daring threshold for the skills in *la.A* [14]:

$$\tau = \min_{l \in ST} p(\textstyle\sum_{s \in la.A} AvgCertainty(Support(\mathrm{s}, \mathrm{SK}(l), \mathrm{LP}) / Card(la.A))$$
$$ZPD(ST) = \{s \in KD(\underline{LP}) \mid D(s, GK(ST), LP) \leq \tau$$

3 Assessment of Group Composition

In Sec. 2, we assumed to already have groups, and to identify their *ZPD*s in order to deploy appropriate paths. Here we tackle the symmetric problem. We assume that we have a group activity (*la*) to submit to the class. This translates in assuming that the *la.A* is within the reach of each one of the group (i.e. in their *ZPD*s). Whatever is the chosen strategy to create groups, it is to consider that two different and complementary aspects play a role in their assessment: the different groups may have different degrees of *intra-group compatibility* and/or *inter-group balance*.

Compatibility among the members inside a group (intra-group) could be considered with respect to several aspects. Here we consider only the aspects related to the knowledge possessed by the students, i.e., the information stored and updated in the personal *SK()* of the group members. In other words, we account for the degree of sharing of cognitive resources and potentialities, which can affect the way the members interact and help each other, and acquire new (firm) knowledge as a results of such collaboration. On the other hand, inter-group balance is the most difficult to achieve, since it entails an attempt for global optimization. Grouping the "smartest" and the "weakest" students separately should be avoided, despite the obviousness of this criterion with respect to intra-group compatibility. It is trivial to consider that the most preforming students may take the greatest learning advantage from interacting with each other. However, if we consider the performance of the overall class, this "segregating" choice may result in leaving behind the students with greatest difficulties. Moreover, supporting and being supported is part of a global social training and improves meta-cognitive abilities. It is often observed that the best way to check one's knowledge about a topic is to let her/him try to explain the core concepts to another person. As it often happens, in the attempt to formalize the activity of group creation we realize how much it is difficult, and how valuable is the experience of a teacher able to do this according to a deep pedagogical sensitiveness. We want to devise viable strategies to identify (a-priori, very hard) or assess (a-posteriori) the intra-group compatibility and inter-group balance.

As above mentioned the most trivial approach would be to create homogeneous groups, but this would exclude weakest students. Therefore the idea is to get the maximum possible *ZPD* for each group, yes minimizing the variance of *ZPD* extension among different groups. Though attractive, this solution is not feasible, since the cognitive span of *ZPD* is not a scalar value. In other words it is not reliable to measure *ZPD* by, e.g., the number of included activities, since this would disregard the most important qualitative element which is the actual content of the activities. Moreover computing a variance would imply to figure out how to compute the difference between a given *ZPD* and an "average" one. Last but not least, computing the complete individual as well as group *ZPD*s constitutes a very demanding task. As a matter of fact we have so far bypassed this problem, in favor of a plainer verification of the possible inclusion of single activities in the *ZPD*. Of course we must assume that it contains at least the skills in *la.A* (with *la* the activity assigned to the group). So we must investigate on a more feasible measure, capable to capture *ZPD* quantitative as well as qualitative span.

3.1 Preliminaries

First of all, given the activity *la* to perform by the groups, we compute for each learner in the class, and for each skill *s* in *la.A*, its inclusion in the individual *ZPD*, according to Sec.2. For some skill this inclusion might not be verified, yet, as discussed above, this can be balanced by the inclusion of the learner in a suitable group, where the condition holds for a sufficient number of members. From this computation, in particular, we retain the value of the *AvgEffort()* to endure reaching the skill. In the following we discuss two possible approaches to group assessment procedure. In both cases we assume a distribution of students in a class in groups $D=\{g_j\}$ *for i=1,…,n.*

3.2 First Approach

We firstly define the *TargetWorkload(l, la)* required to the members of $g \in D$ by the activity, as the sum of the *AvgEffort()* for each $s \in la.A$. Then we can consider:

- *IntraGroupTotEff(g)* as the sum of the *TargetWorkload(l, la)* of all members;
- *IntraGroupAvgEff(g)* as the average *TargetWorkload(l, la)* over all members;
- *InterGroupAvgEff(D)* as the average of the *IntraGroupAvgEff(g) for g in D*
- *InterGroupVarEff(D)* as the variance of the *IntraGroupAvgEff(g) for g in D*

A possible way to optimize the distribution of students in groups, is to minimize both *InterGroupAvgEff(D)* and *InterGroupVarEff(D)* at once.

Additionally, a possible measure of the quality of the distribution can be provided by the comparison between *InterGroupAvgEff(D)* and the sum of *TargetWorkload(l, la)* for all *l* in the class, divided by the *Card(D)*. This might be relevant when a better distribution among the possible ones is sought.

3.3 Second Approach

A second approach entails considering the daring threshold, *DTreshold(s,SK(l))*, used in the definition of the individual ZPD. We remind that it depends on the firmness of the skills owned by the learner *l* in the *Support* subset of the *SK(l)* (cfr Sec. 2.2). It is conceivable that for some learners in the group the distance *D(s, ZPD(l))* will be zero (i.e. $s \in ZPD(l)$), while for the others it will be a positive value. Notice that the distance between *ZPD(l)* and s can be measured in the same way as for *SK(l)*.

A quality of the group (meaning a characteristics helpful in order for the group members to reach *s* after a collaborative learning experience) is in the couple

$<m_g, \sigma_g>$, where

- m_g = average(*D(s, ZPD(l))*) for $l \in g$
- σ_g is the variance of the *D(s, ZPD(l))* for $l \in g$

A lower σ_g tells us that the group is homogeneous. A higher m_g might point out that the skill is hard to reach for the group members, or that there is a very limited subset of members that could pull the rest towards the skill.

Regarding the distribution, also in this approach a better quality is reached when m_g and σ_g are minimized at once throughout the groups.

3.4 Third Approach

A further approach takes into consideration the definition of individual *ZPD(l)*, given in Sec. 2, and in particular the daring factor *dF*, configured by the teacher. We define a partition of a group *g* as

- $F(g, s) = \{l \in g$ such that $s \notin ZPD(l)\}$
- $H(g, s) = \{l \in g$ such that $s \in ZPD(l)\}$

It is reasonable to think that by modifying dF we could enlarge or shrink a given *ZPD(l)*, so we consider the following measures:

- for each l in $F(g, s)$ the minimal value $\Delta^{+}dF(l)$ such that computing the $ZPD(l)$ using $(dF+\Delta^{+}dF(l))$ causes $s \in ZPD(l)$
- for each l in $H(g, s)$ the minimal value $\Delta^{-}dF$ such that computing the ZPD(l) using $(dF-\Delta^{-}dF)$ causes $s \notin ZPD(l)$

So a characteristics of the group is in the balance between the measures of affordability of the skill by the two group partitions, defined as the subtraction

$$Avg(\Delta dF(l)\ for\ l \in H(g, s)) - Avg(\Delta' dF(l)\ for\ l \in F(g, s))$$

The bigger this value, the less the difficulties of the learners whose *ZPD* has been stretched, because their difficulties can be eased by the potential support of the learners in *H(g, s)*. The lower this value, the closer (more homogeneous) are the learners in *H(g, s)* and *F(g, s)*.

Even in this case the variance of this value over the groups of a distribution gives an estimate of their inter-homogeneity.

4 Conclusions and Future Work

In this paper we described a stage in our effort to implement the concept of Zone of Proximal Development (*ZPD*), originated in the educational theories of L.V. Vygotskij, within a framework of web based technology enhanced learning. In particular our framework tries to join the more traditionally individualized activities of a system for personalized e-learning with the learning experience allowed in an environment supporting social and collaborative e-learning. This is done also by "using" the concept of *ZPD* to support an as free as possible navigation of the personal learning path, under the sole constraints given by the needs to take *ZPD* into consideration while navigating. The first requirement is a suitable definition of individual *ZPD* in terms of feasible activities. A further extension to groups can be a valuable help in the task of partitioning a class in a pedagogically effective way.

In our framework we attempted to define the *ZPD* in a reasonable yet challenging way, which may stimulate the student without causing frustration, and for this reason we defined it so that its radius is not the same for all students but depends on their present state of knowledge. As a matter of fact, it is often the case that the *ZPD* is merely considered as an additional bag of skills which is possible to acquire, disregarding its genuine interpretation as a region of cognitive *development*. On the contrary, we think that our definition of an individually tuned *ZPD* can support a pedagogically meaningful implementation of a zone of development, spurring the acquisition of techniques and knowledge skills from previously possessed ones.

We also attempted the definition of group *ZPD* starting from a given group and including from time to time “pedagogically sound” activities. Here we have extended our (still theoretical) work to support the symmetric operation, i.e. to start from a given activity and partition a class in the most effective set of groups. It was soon clear to us the difficulty of both automatically creating a class partition as well as of assessing the validity of a given one. Teachers are smart in this task, while automatic processing of students’ data requires a complex double optimization procedure. In fact, the final aim is to both maximize intra-group compatibility and/or inter-group balance. The first should ensure fair collaboration within the group, and the second should avoid creating “best” and “worst” groups by enforcing the sense of collaboration and of general belonging to a same super-group (the class). Given this double goal, the most obvious solutions soon appeared unfeasible. This work presented a first attempt to mark a line along which to continue investigating appropriate alternatives.

In the future, besides assessing starting group adequacy, it would be also beneficial to assess the positive/negative dynamics within the groups, according to the amount of growth of individual *APS*s and *ZPD*s.

References

1. Martens, A.: Modeling of Adaptive Tutoring Processes. In: Ma, Z. (ed.) Web-Based Intelligent e-Learning Systems, pp. 193–215. IGI-Global (2005)
2. Limongelli, C., Sciarrone, F., Vaste, G.: LS-PLAN: An effective combination of dynamic courseware generation and learning styles in web-based education. In: Nejdl, W., Kay, J., Pu, P., Herder, E. (eds.) AH 2008. LNCS, vol. 5149, pp. 133–142. Springer, Heidelberg (2008)
3. Limongelli, C., Sciarrone, F., Vaste, G.: Personalized e-learning in moodle: The moodle-LS system. J. of E-Learning and Knowledge Society 7(1), 49–58
4. Limongelli, C., Lombardi, M., Marani, A., Sciarrone, F.: A Teacher Model to Speed Up the Process of Building Courses. In: Kurosu, M. (ed.) HCII/HCI 2013, Part II. LNCS, vol. 8005, pp. 434–443. Springer, Heidelberg (2013)
5. Limongelli, C., Sciarrone, F., Temperini, M., Vaste, G.: The Lecomps5 Framework for Personalized Web-Based Learning: a Teacher’s Satisfaction Perspective. Computers in Human Behavior 27(4) (2011)
6. Ivanova, M., Popova, A.: Formal and Informal Learning Flows Cohesion in Web 2.0 Environment. Int. J. of Information Systems and Social Change, IJISSC 2(1), 1–15 (2011)
7. Sterbini, A., Temperini, M.: Learning from Peers: Motivating Students through Reputation Systems. In: Proc. Int. Symp. on Applications and the Internet, SAINT, pp. 305–308. IEEE (2008)
8. Cheng, Y., Ku, H.: An investigation of the effects of reciprocal peer tutoring. Computers in Human Behavior 25 (2009)

9. De Marsico, M., Sterbini, A., Temperini, M.: A strategy to join adaptive and reputation-based social-collaborative e-learning, through the Zone of Proximal Development. Int. Journal of Distance Education Technology, IJDET 11(3), 12–31 (2013)
10. De Marsico, M., Sterbini, A., Temperini, M.: The Definition of a Tunneling Strategy between Adaptive Learning and Reputation-based Group Activities. In: Proc. 11th IEEE Int. Conf. on Advanced Learning Technologies, ICALT, pp. 498–500 (2011)
11. Kreijns, K., Kirschner, P.A., Jochems, W.: Identifying the pitfalls for social interaction in computer supported collaborative learning environments: a review of the research. Computers in Human Behavior 19, 335–353 (2003)
12. Limongelli, C., Lombardi, M., Marani, A., Sciarrone, F.: A Teaching-Style Based Social Network for Didactic Building and Sharing. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 774–777. Springer, Heidelberg (2013)
13. Vygotskij, L.S.: The development of higher forms of attention in childhood. In: Wertsch, J.V. (ed.) The Concept of Activity in Soviet Psychology, Sharpe, Armonk (1981)
14. De Marsico, M., Sterbini, A., Temperini, M.: A Framework to Support Social-Collaborative Personalized e-Learning. In: Kurosu, M. (ed.) HCII/HCI 2013, Part II. LNCS, vol. 8005, pp. 351–360. Springer, Heidelberg (2013)
15. Dougiamas, M., Taylor, P.: Moodle: Using learning communities to create an open source course management system. In: Proc. World Conference on Educational Multimedia, Hypermedia and Telecommunications, vol. 1, pp. 171–178
16. De Marsico, M., Temperini, M.: Average effort and average mastery in the identification of the Zone of Proximal Development. In: Proc. 17th IEEE Int. Conf. on System Theory, Control and Computing, ICSTCC, 6th Int. Workshop on Social and Personal Computing for Web-Supported Learning Communities, SPeL, pp. 651–656 (2013)
17. Chaiklin, S.: The zone of proximal development in Vygotsky's analysis of learning and instruction. In: Kozulin, A., Gindis, B., Ageyev, V., Miller, S. (eds.) Vygotsky's Educational Theory in Cultural Context, pp. 39–64. Cambridge University Press (2003)