

# An Approach to Measuring the Difficulty of Learning Activities

Francisco J. Gallego-Durán<sup>(✉)</sup>, Rafael Molina-Carmona,  
and Faraón Llorens-Largo

Cátedra Santander-UA de Transformación Digital,  
Universidad de Alicante, Alicante, Spain  
{fgallego,rmolina,faraon}@dccia.ua.es

**Abstract.** In any learning environment, training activities are the basis for learning. Students need to practice to develop new skills and improve previously acquired abilities. Each student has specific needs based on their previous knowledge and personal skills. The allocation of a proper activity for a particular student consists in selecting a training activity that fits the skills and knowledge of the student. This allocation is particularly important since students who are assigned a too hard training activity will tend to leave it rather than making the necessary effort to complete it. Moreover, when the activity is too easy it does not represent a challenge for the student and the learning outcomes will tend to be very limited. An motivating activity, suitable for a given student, should be neither too easy nor too difficult. The problem arises when trying to measure or estimate the difficulty given any training activity. Our proposal is a definition of difficulty of a learning activity that can be used to measure the learning cost of a general learner. As a first step, the desirable features and the intrinsic limitations of a difficulty function are identified, so that a mathematical definition can be obtained quite straightforward. The result is an intuitive, understandable and objectively measurable way to determine the difficulty of a learning activity.

**Keywords:** Difficulty estimation · Learning activity

## 1 Introduction

Learning is a fascinating event that happens in nature as the maximum expression of adaptation. Having the ability to learn means being able to remember past events and associate them with present situations to take decisions. Although quite a few low-level mechanisms of learning remain unknown, life experience suggests that there is a strong correlation between the activities any individual performs and learning. This is the basis for training and education: creating experiences to make learners' brains adapt and remember. In fact, learning experiences designed by the teachers in order to generate concrete learning outcomes on learners are often referred to as Learning Activities.

The space of possible learning experiences is huge so, which experiences are more suitable for a learning objective and a particular learner? Among others, the concept of difficulty has a very important role to play in this context.

Difficulty is a widespread concept and it can be intuitively defined as the abilities and cost for a learner to successfully complete a given learning activity. In spite of being subjective, it is considered among the main factors determining learners' motivation [4, 5, 13, 26, 27]. In order for optimize learning, difficulty of any given exercise should match abilities of the learner [12, 14, 19, 24, 28]. Matching the abilities means being so difficult as to be an interesting challenge, at the same time as being so easy to be reachable with a limited amount of effort in time. Therefore, correctly estimating the difficulty of learning activities is the first step to be able to optimize the learning process.

The main objective of this research is creating a definition of difficulty of a learning activity that can be used to measure the learning cost for a general learner. A general learner would be any individual that performs learning activities and improves its results, as a consequence of learning.

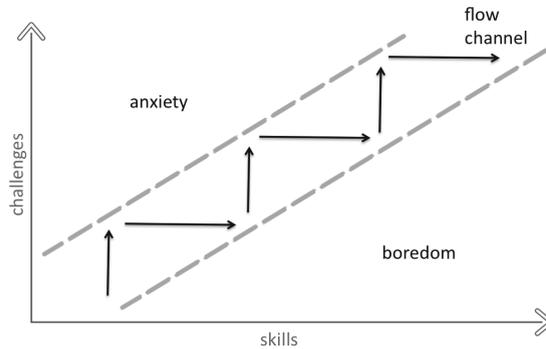
A brief background about the research context is presented in Sect. 2. Section 3 identifies the information sources available for measuring difficulty. The desired properties and the limitations of difficulty measures are presented in Sects. 4 and 5. Section 6 is devoted to design some mathematical functions that meet the properties and give useful outcomes as well as analysing their usefulness for learning activities. Finally, the conclusions and future work are presented in Sect. 7.

## 2 Background

To understand the relation between difficulty, motivation and learning, let us focus on the notion of *Flow Channel* (Fig. 1) [7, 25]. The Flow Channel represents the way difficulty and skills of the learner relate to each other:

- When difficulty is much higher than learners' skills, anxiety appears. This is psychologically explained by learners perceiving their skills as insufficient, thus getting demotivated. They normally feel that the activity requires too much effort compared to their perceived capabilities. This often leads to early abandon.
- On the contrary, if learners' skills already include what the activity provides as learning outcome, boredom shows up. Having to invest time and/or resources to get an already possessed outcome is interpreted as lost time. Interest vanishes, motivation decreases and boredom appears.
- When skills and difficulty are balanced, learners enter a state of Flow. In Schell words [25], Flow is sometimes defined as *a feeling of complete and energized focus in an activity, with a high level of enjoyment and fulfilment.*

This research assumes The Flow Channel theory as a key point for improving the design and selection of learning activities.



**Fig. 1.** Representation of the flow channel, by Csikszentmihalyi [3]

Some research work has been carried out calibrating difficulty by analysing student historical data [21], or using linear regression to estimate difficulty based on user data [2] or even on generating exercises automatically with a given established difficulty [20, 22]. But these studies are spread, discontinued and seem to be disconnected from each other. In general, the concept of difficulty within the academic world does not seem to capture too much attention.

More studies related to difficulty can be found changing the focus to the field of Computer Games. The parallelism with academic learning is complete: if a level of a game is too difficult or too easy, players tend to stop playing the game. Therefore, it is vital for a game to have a well designed progression of difficulty, if willing to catch the attention of the players. Most studies in this field try to develop methods to dynamically adjust difficulty to match player's skills [9, 10, 15, 16]. All these studies use existent levels of difficulty proposed in present Computer Games and focus on selecting the most appropriate for each player and game being played. Hunicke and Chapman [9, 10] take measures of performance of the player and tries to predict if the player is going to fail to anticipate and adjust the level of difficulty. The proposal is completely specific to First Person Shooter (FPS) games [23], as measures are defined for this specific type of gameplay. Mladenov and Missura [16] use data collected from previously played games to analyse a set of gameplay characteristics and input this data to a supervised Machine Learning algorithm. The goal is to have an offline prediction of the level of difficulty players are going to select in their next game. Missura and Gartner [15] take a different approach for automatically selecting difficulty for a given player among a set of finite difficulty levels. They divide the game into play-review cycles. They measure the performance of the player in the play cycles, and change difficulty level on review cycles accordingly to their estimations.

Herbrich et al. [8] present a very interesting work on measuring players' skills comparatively. Their system, called TrueSkill, is based on chess' Elo rating system [6]. Just like the Elo rating system, players have a 1-dimensional value ranking that predicts their probability of winning against other players by logistic comparison. Although this work is not directly based on difficulty, it is indirectly

valuing players' skill with similar intention: match players against those with similar abilities to foster balanced games.

Another interesting work is that proposed by Mourato and dos Santos [17]. Their goal is to procedurally generate content for Platform Games similar to Super Mario Bros [18]. The problem with this kind of content is how to classify the generated content with respect to difficulty. They propose a way to measure difficulty in Platform Games by measuring players' probability of failing after each individual obstacle in the game. The concepts are interesting but they lack a practical result with actual players and ready-to-be-played generated content.

Finally, Aponte et al. [1] present one of the most interesting reviewed works. In their work they state that their goal is "to evaluate a parameter or a set of parameters that can be considered as a measure of a game difficulty". They start by measuring the difficulty of a reduced Pacman game with 1 ghost. In their Pacman game, speed of the ghost is a configurable parameter to make the game more difficult at will. They measure the score of a synthetic player as number of eaten pellets and then show a graph with the evolution of this value depending on the speed of the ghost. This approach lets them show the progression of difficulty depending on the selected level (speed of the ghost). Based on that result, they define a set of properties that a general definition of difficulty should have, and propose a general theoretic definition of difficulty as the probability of losing at a given time  $t$ . They only propose this definition, but do not perform any kind of test or mathematical proof. It ends up as a simple proposition based on their arguments.

All these previous works demonstrate the incipient interest of the research community for measuring difficulty. This trend is confirmed by the growing focus on measuring learning in general. The NMC Horizon Report: 2016 Higher Education Edition [11] states that there is a renewed interest in assessment and the wide variety of methods and tools to evaluate and measure the elements related to the learning process.

### 3 Sources for Measuring Difficulty

Let us consider difficulty as a cost: in order to successfully finish an activity, any learner has to pay a cost in time and effort. Measuring time is trivial from a conceptual point of view. The problem comes from measuring effort. How can we measure effort? Do we have an objective definition of what effort is?

It will be considered that effort is indirectly related to progress. The more progress is achieved, the less effort is required to finish. Although this logic consideration is not a concrete definition of effort, it has many advantages:

- For many kinds of activity, progress is relatively easy to define and measure objectively.
- A measure for progress is also closely related to learning outcomes: most activities yield learning outcomes even when not fully completed. In fact, that learning outcomes become clear when success ratio increases out of repeating the activity.
- As progress to success is one of the key factors in motivation, measures taking progress into account also foster motivation.

Therefore, this research will consider “more difficult” an activity when less progress is done. In the sake of rigour, progress will be considered with respect to time: progress percentage per unit of time will be an inverse measure for difficulty. So, an activity being “more difficult” will imply that less progress is made per time unit. This will let us measure difficulty in an intuitive, understandable and objectively measurable way.

## 4 Desired Properties for Difficulty

There are several ways of defining difficulty as a relationship between time and progress. It is important to have guidance for selecting an appropriate measure from such a huge set of potential definitions. So, establishing a set of desired properties will ensure that the selected definition is useful under defined criteria. These desired properties will act as restrictions, reducing the search space.

Let us consider the next set of properties, having present that measuring and comparing learning activities is the final goal:

- Difficulty should always be positive. Progress and time are always positive or 0 values when measuring a learning activity. A negative difficulty coming out of these two values is impossible and would have no meaning.
- Difficulty should have a minimum value. A difficulty value of 0 would mean that no time/effort is required to finish a given activity. That would correspond to an activity that is already done.
- Difficulty should also have a maximum value. Making difficulty unbound would imply that any value could be “not so difficult” compared to infinite. Having a maximum value lets us fix impossible activities, which is desirable. An unbound upper limit that should be labelled as infinity makes formulation more complicated and has no advantage on comparisons.
- Fixing 1 as the maximum value for difficulty has advantageous properties. That bounds difficulty in the range  $[0, 1]$ , which lets us consider it as a probability. That makes sense and is compatible with previous considerations. Moreover, that enables the probability theory as a valid set of tools for working with difficulty, which is very desirable.
- Difficulty should not be a unique value but a function over time. While an activity is being done, difficulty keeps changing as progress is being made.
- Difficulty must be a continuous function over time. It makes no sense for a moment in time not to have a difficulty associated.
- Difficulty must be a non-strictly decreasing function. Every time a learner makes progress on a given learning activity, difficulty decreases by definition as less progress is required to meet success.

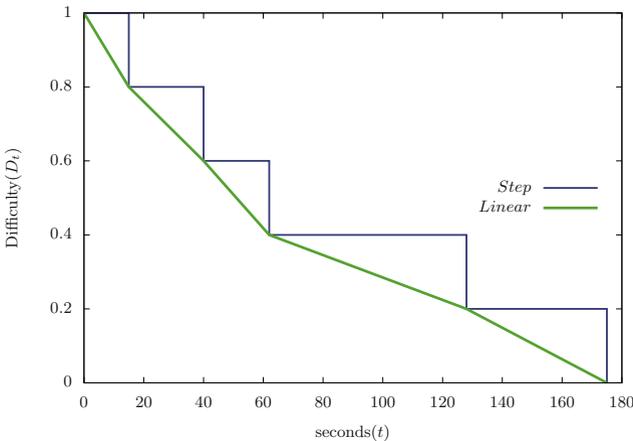
Let us consider an example of activity: “scoring five 3-point shots in a basketball court, in less than 5 min”. This is a training activity whose expected learning outcome is an improvement in shooting precision to basket<sup>1</sup>. This activity will take

<sup>1</sup> Although other learning outcomes can be considered from this activity, let us consider it just as a precision improvement exercise.

at most 5 min, and at least the time required to shot 5 times: time cost is straightforward. Regarding to effort, it will depend on previous conditions. A trained, muscular player may complete the activity fast, without much effort, whereas a weak novice could require many attempts to finish it successfully. Moreover, novice players may waste much more energy because they lack adequate technique. This could also be considered more effort.

The activity could be analysed many times and from different perspectives, and many definitions for “effort” could be found. Before entering an endless debate on what “effort” is or should be, let us consider a useful point of view with respect to our goal of measuring difficulty. An indirect measure for “effort” could be derived from the intrinsic failure/success measures of the activity. When 5 min are over, a player that scored 4 baskets is closer to success than other who only scored 1. It can be considered that having scored 4 baskets leaves out less progress to be done for succeeding than scoring just 1. Under this consideration, there is less effort pending to succeed when more percentage of the activity has been completed.

Let us compose a function with this properties, for the basketball example. Let us imagine a player that scores 5 baskets at times  $t_i \in \{15, 40, 62, 128, 175\}$ ,  $i \in \{1, 2, 3, 4, 5\}$  in seconds. Difficulty could be represented as shown in Fig. 2: whenever player scores baskets, difficulty decreases. Decreasing difficulty can be considered as a step function, maintaining its value except on scoring events. It also can be considered as a linear function, resulting on a much smooth shape. Moreover, a linear function seems to inform better about the pace of the player.



**Fig. 2.** Manually constructed difficulty function for basket example. Difficulty decreases as player progresses, scoring baskets in this example.

As it can be deduced from Fig. 2, these properties configure a very powerful definition of difficulty: it goes far beyond a simple scalar quantity, defining a representative function. This function represents progress of the player over time which

gives much more information about the process. This new information will also be useful for visual comparison of activity profiles as well as individual or group profiles.

## 5 Intrinsic Limitations

The selected properties limit the way activities should be defined. Not every possible activity will fit for this model. This is both a limitation and a design guide. Activities designed for this model of difficulty will have a set of properties:

- Activities require progress to be measurable (i.e. they should have a score). For instance, an activity defined as “selecting the proper answer from a set of 4” has no way of measuring progress. Although time to answer and success can be measured, there is no progress towards success. Resulting functions would represent either a full square or a line, depending on model selected.
- Score (i.e. progress) has to be non-strictly increasing function over time. As score is measuring progress to an end it does not make sense for it to decrease. General score measures having punishments or negative score events would not be appropriate. However, almost any score measure could be transformed in an equivalent non-strictly increasing measure for this purpose.
- Activities must have a measurable success status or, at least, a maximum score. This is required to define difficulty within its limits. Progress can be measured in unbounded activities, but cannot be scaled to a  $[0, 1]$  range.
- Activities must be considered over time. For instance, an activity about creating a program cannot be considered just as its final result. Having a single point of evaluation is similar to not being able to measure progress. It is also very important to measure time required to do the activity. If all the learners hand the result of an activity at the same time and no measures have been taken previously, no data will be available for the model.

These intrinsic limitations are part of the selected set of properties and shall be assumed. However, it can be seen positively. Having activities where progress is measurable over time and with well defined score limits or success status is very interesting for learners. Progress informs learners about the status of their evolution to success, and also remove their doubts about their skills being enough for the activity. Although these design impositions are not easy to achieve on every activity, they are definitely desirable from educational point of view.

## 6 Mathematically Defining Difficulty

With all desired properties and limitations clarified, a working mathematical definition of difficulty can be constructed. Let  $A$  be the set of all possible activities, and  $L$  the set of all possible learners. Let  $\alpha \in A$  be a concrete learning activity. As an activity,  $\alpha$  can be performed by any learner  $l \in L$ . Each  $l$  performs  $\alpha$  a number of times  $N_l \in \mathbb{N}$ . So let  $\alpha_l^i, l \in L, i \in \mathbb{N}, i \leq N_l$  represent the  $i$ -th realization of the activity  $\alpha$  by the learner  $l$ .

Each  $\alpha_l^i$  takes an amount of time  $t_l^i \in \mathbb{R}$ , measured in seconds. Let us consider, for simplicity, that each  $\alpha_l^i$  starts at time 0 and ends at  $t_l^i$ . Then, let  $S_t(\alpha_l^i) \in \mathbb{R}$  be a function that measures the score got by learner  $l$ , at time  $t$  on its  $i$ -th realization of  $\alpha$ . So,  $S_t(\alpha_l^i)$  is the function that measures the progress towards success of a learner that performs an activity.

The score function is expected to be explicitly defined for each activity. In fact, many different score functions can be defined for each activity. Therefore, let us assume that activities and their score functions are defined by activity designers. Also, for clarity reasons, let us assume that activities and score functions meet desired properties and limitations exposed on Sects. 4 and 5.

In previous sections, difficulty has been defined as the inverse of progress. However, this cannot be defined exactly this way. Difficulty must be defined in  $[0, 1]$  range, and the score function could have a much broader range. However, the score function should be non-strictly increasing, and should have an upper limit. Therefore, the score function could be safely assumed to start at 0, because the actual range of the function can always be moved to start at 0. Let  $S^*(\alpha)$  be the maximum score value for the activity  $\alpha$ ,

$$S^*(\alpha) \in \mathbb{R}, S^*(\alpha) \geq S_t(\alpha_l^i) \quad \forall l \in L, i \in N_l \quad (1)$$

This lets us define the “easiness function” as a scaled version of the score function over time in the  $[0, 1]$  range:

$$E_t(\alpha_l^i) = \frac{S_t(\alpha_l^i)}{S^*(\alpha)} \quad (2)$$

The function defined in Eq. 2 is called “easiness function” as it is exactly the inverse of the initial definition of difficulty. Therefore, the definition of difficulty follows:

$$D_t(\alpha_l^i) = 1 - E_t(\alpha_l^i) \quad (3)$$

This definition of difficulty is tied to the concept of progress. It represents an advantage over estimating difficulty with just a single scalar value: the resulting graph shows an evolution over time which informs of the whole realization of the activity. It also yields instant values for difficulty at any time of the realization. This values intrinsically represent the percentage of progress remaining to finish the activity. They could also be interpreted as the probability of failing the activity<sup>2</sup>.

However, these values are quite plain: they are instant values that do not capture information on the progress by themselves. The result is similar to considering any instant  $t$  to be independent from the others that compose the timeframe of the activity. For instance, this is like considering in the basketball example that scoring at first shot is equally probable to scoring after 4 baskets, or at a last attempt, when time is finishing. Nevertheless, a more accurate definition should consider that events occurring at time  $t$  are influenced by all events happened in the range  $[0, t]$ .

<sup>2</sup> This interpretation is bound to discussion about its real meaning as a probability.

Experience shows that influence of a timeframe over next time steps is strong on humans. It is convenient to consider how human factors relate over time: psychological status, strength, fatigue, motivation, etc. Time steps in the timeframe of any learning activity, performed by a human learner, are best considered to be strongly interdependent. Therefore, can be improved by making  $D_t$  depend on a function of all  $t' \in [0, t]$ , to make final values express this interdependency.

There are many approaches to make  $D_t$  dependent on the set of all past values of difficulty  $\{D_{t'}/t' \in [0, t]\}$ . Moreover, there is no theoretical way to determine the appropriate way to weight all the possible factors. What is more, different activities and learners will have different influence factors. This makes extremely difficult, if at all possible, to design a theoretical relation covering such a chaotic landscape. This suggests using an experimental approach instead. Therefore, this research starts modelling influence in a very simple way. This first model can be used as a benchmark to test other different approaches and experimentally determine better ways of defining difficulty.

Assuming that  $D_t, \forall t$  should depend on  $\{D_{t'}/t' \in [0, t]\}$  and  $0 \leq D_t \leq 1$ , let us define  $D_t$  as the area of the curve above  $E_t$  related to the maximum possible area up to the instant  $t$ ,

$$D_t(\alpha_i^l) = 1 - \frac{1}{t} \int_0^t E_t(\alpha_i^l) dt \tag{4}$$

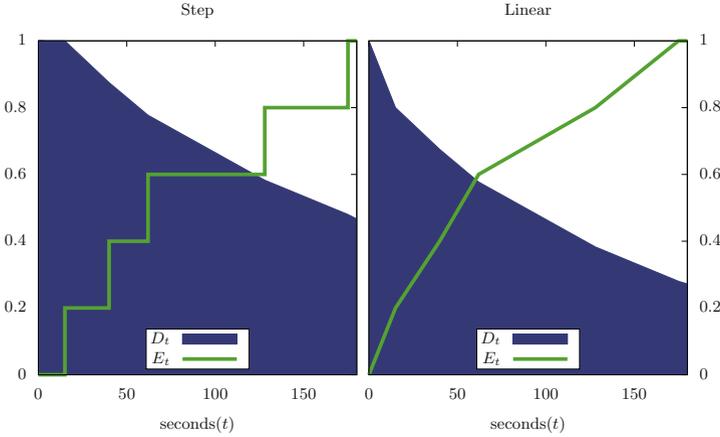
Equation 4 defines difficulty  $D_t$  as a value depending on all previous history of the  $i$ -th realization of an activity  $\alpha$  by a learner  $l$ . The dependency is made indirect, using the easiness function as a proxy for difficulty. This makes definition easier, eliminating recursive references and associated problems.

Using the new definition stated at Eq. 4 the graphical layout of  $D_t$  varies greatly, as Fig. 3 shows. Compared to Fig. 2, the new definition for  $D_t$  results in a function that responds much smoothly to score events. This new behaviour shows an interesting feature. Let us assume that  $t \in [0, t^*]$ . Using Eq. 4,  $D_{t^*}$  will directly depend on the performance shown by the learner during the realization of the activity (being  $D_{t^*} > 0^3$ ). In the basketball example, the faster baskets get scored, the lower  $D_{t^*}$  will be, and vice-versa. Therefore, after completing an activity, the lower the residual difficulty value  $D_{t^*}$ , the greater the performance shown by the learner.

The interesting property shown by  $D_{t^*}$  is a direct consequence of its cumulative definition. So, this property will be shown by  $D_{t'}, \forall t' \in [0, t^*]$ . Therefore,  $D_t$  can now be used as a performance measure with more information than  $E_t$ , as it integrates information about score and time/frequency in one single value. Careful analysis of  $D_t$  for different learners and realizations of the same activity could lead to establishing correlations with abilities learnt and degree of mastery.

---

<sup>3</sup> Unless  $D_0 = 0$ , which would only happen on activities completed at start time. That is a degenerate case with no interest in practice. Thus, it can be safely ignored.



**Fig. 3.** Behaviour of  $D_t$  using Eq. 4 with data from the basketball example from Sect. 4. Left, exact definition for  $E_t$  with step value changes. Right, linear interpolation for  $E_t$ .

## 7 Conclusions and Further Work

In this paper a general definition for difficulty has been presented. This new definition has been designed on the bases of a list of desired properties. By using this proposed definition, difficulty becomes measurable, can be compared and visualized, and is related to effort over time. Effort is modelled as required time to achieve a specific score value. Therefore, proposed definition of difficulty takes into account progress towards solving a learning activity, based on the score an agent achieves when performing the activity.

This proposed definition of difficulty has limitations in the sense that activities have to meet some requirements to be measurable under this definition. The activity should be performed over time and a score function to measure progress should be available. The score function should have upper and lower boundaries and be non-strictly increasing: there should be no possibility of losing score over time.

But this proposed definition has also many interesting advantages. Most of its advantages come from being drawable: this confers it the ability to show its progress over time, so that it transmits characteristics of the learner and the learning activity graphically. Different parts of the learning activity can be identified: for instance, most difficult parts will produce valleys in the graph that will permit not only their identification, but also their measurement. Activities can be compared using their difficulty graphs, yielding a much accurate knowledge about which ones require more effort, and the differences in the distribution of the effort over time. These advantages make the proposed definition of difficulty a powerful tool for analysing and comparing learning activities.

The formal definition of difficulty has other implications: it is possible to take it as a starting point to build adaptive learning systems that could automatically

assign the suitable learning activities to students. As a further step, we plan to analyze the behaviour of the difficulty definition in real learning activities with different features, as well as to make a deeper study about the meaning of the proposed charts. The final aim is maintaining learners in the state of flow or, in other words, motivating them.

## References

1. Aponte, M.-V., Levieux, G., Natkin, S.: Scaling the level of difficulty in single player video games. In: Natkin, S., Dupire, J. (eds.) ICEC 2009. LNCS, vol. 5709, pp. 24–35. Springer, Heidelberg (2009)
2. Cheng, I., Shen, R., Basu, A.: An algorithm for automatic difficulty level estimation of multimedia mathematical test items. In: Eighth IEEE International Conference on Advanced Learning Technologies, ICALT 2008, pp. 175–179, July 2008
3. Csikszentmihalyi, M.: *Flow: The Psychology of Optimal Experience*. Perennial Modern Classics. Harper Perennial Modern Classics, New York (1990)
4. D’Mello, S., Olney, A., Williams, C., Hays, P.: Gaze tutor: a gaze-reactive intelligent tutoring system. *Int. J. Hum.-Comput. Stud.* **70**(5), 377–398 (2012)
5. Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J.: Gamifying learning experiences: practical implications and outcomes. *Comput. Educ.* **63**, 380–392 (2013)
6. Elo, A.E.: *The Rating of Chess Players, Past and Present*. Arco Pub, New York (1978)
7. Getzels, J., Csikszentmihályi, M.: *The Creative Vision: A Longitudinal Study of Problem Finding in Art*. A Wiley-Interscience publication, Wiley (1976)
8. Herbrich, R., Minka, T., Graepel, T.: Trueskill<sup>TM</sup>: a bayesian skill rating system. In: Schölkopf, B., Platt, J., Hoffman, T. (eds.) *Advances in Neural Information Processing Systems*, vol. 19, pp. 569–576. MIT Press (2007)
9. Hunicke, R.: The case for dynamic difficulty adjustment in games. In: *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, ACE 2005, NY, USA*, pp. 429–433 (2005)
10. Hunicke, R., Chapman, V.: *AI for dynamic difficulty adjustment in games* (2004)
11. Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., Hall, C.: *NMC Horizon Report: 2016 Higher Education Edition*. New Media Consortium; EDUCAUSE Learning Initiative, Austin Texas; [S.l.] (2016)
12. Koster, R., Wright, W.: *A Theory of Fun for Game Design*. Paraglyph Press, Pittsburgh (2004)
13. Lee, J.J., Hammer, J.: Gamification in education: what, how, why bother? *Acad. Exch. Quart.* **15**(2), 2 (2011)
14. Ley, T., Kump, B.: Which user interactions predict levels of expertise in work-integrated learning? In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) *EC-TEL 2013. LNCS*, vol. 8095, pp. 178–190. Springer, Heidelberg (2013)
15. Missura, O., Gartner, T.: Predicting dynamic difficulty. In: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K. (eds.) *Advances in Neural Information Processing Systems*, vol. 24, pp. 2007–2015. Curran Associates, Inc. (2011)
16. Mladenov, M., Missura, O.: *Offline learning for online difficulty prediction* (2010)
17. Mourato, F.J., dos Santos, M.P.: Measuring difficulty in platform videogames. In: *4a Conferencia Nacional em Interacao Pessoa-Mquina*, pp. 173–180. Grupo Portugues de Computaao Grfica/Eurographics (2010)

18. Pedersen, C., Togelius, J., Yannakakis, G.N.: Modeling player experience in Super Mario Bros. In: Proceedings of the 5th International Conference on Computational Intelligence and Games, CIG 2009, Piscataway, NJ, USA, pp. 132–139. IEEE Press (2009)
19. Petkovic, D., Okada, K., Sosnick, M., Iyer, A., Zhu, S., Todtenhoefer, R., Huang, S.: Work in progress: a machine learning approach for assessment and prediction of teamwork effectiveness in software engineering education. In: Frontiers in Education Conference (FIE), pp. 1–3, October 2012
20. Radošević, D., Orehovački, T., Stapić, Z.: Automatic on-line generation of student's exercises in teaching programming. In: "Automatic On-line Generation of Students Exercises in Teaching Programming", Central European Conference on Information and Intelligent Systems, CECIIS (2010)
21. Ravi, G., Sosnovsky, S.: Exercise difficulty calibration based on student log mining. In: Mšdritscher, F., Luengo, V., Lai-Chong Law, E., Hoppe, U. (eds.) Proceedings of DAILE 2013: Workshop on Data Analysis and Interpretation for Learning Environments. Villard-de-Lans. France, January 2013
22. Sadigh, D., Seshia, S.A., Gupta, M.: Automating exercise generation: a step towards meeting the MOOC challenge for embedded systems. In: Proceedings of Workshop on Embedded Systems Education (WESE), October 2012
23. Saldana, J.M., Marfia, G., Rocchetti, M.: First person shooters on the road: leveraging on aps and vanets for a quality gaming experience. In: Wireless Days, pp. 1–6. IEEE (2012)
24. Schalk, P., Wick, D., Turner, P., Ramsdell, M.: Predictive assessment of student performance for early strategic guidance. In: Frontiers in Education Conference (FIE 2011), pp. S2H-1–S2H-5, October 2011
25. Schell, J.: The Art of Game Design: A Book of Lenses. Morgan Kaufmann Publishers Inc., San Francisco (2008)
26. Verdú, E., Regueras, L.M., Verdú, M.J., Leal, J.P., de Castro, J.P., Queirós, R.: A distributed system for learning programming on-line. *Comput. Educ.* **58**(1), 1–10 (2012)
27. Wang, A.Y., Newlin, M.H.: Characteristics of students who enroll and succeed in psychology web-based classes. *J. Educ. Psychol.* **92**(1), 137 (2000)
28. Yoo, J., Kim, J.: Can online discussion participation predict group project performance? investigating the roles of linguistic features and participation patterns. *Int. J. Artif. Intell. Educ.* **24**(1), 8–32 (2014)