International Journal of Artificial Intelligence in Education manuscript No.
(will be inserted by the editor)

Adaptive, Intelligent, and Personalized: Navigating the Terminological Maze Behind Educational Technology

Radek Pelánek

Received: date / Accepted: date

Abstract Educational technology terminology is messy. The same meaning is often expressed using several terms. More confusingly, some terms are used with several meanings. This state is unfortunate, as it makes both research and development more difficult. Terminology is particularly important in the case of personalization techniques, where the nuances of meaning are often crucial. We discuss the current state of the used terminology, highlighting specific cases of potential confusion. In the near future, any significant unification of terminology does not seem feasible. A realistic and still very useful step forward is to make the terminology used in individual research papers more explicit.

1 Introduction

Terminology in educational technology is far from clear and standardized. For example, Table 1 lists just some of the widely used general educational technology terms. These terms often have a specific meaning, but they also have significant overlaps. In this work, we focus on aspects of educational technology that are concerned with personalization, adaptation, or other kinds of intelligent behavior. This is a specific area where the terminological situation is confusing and particularly important since even small nuances of the meaning of terms can have significant consequences for understanding and reproducibility of described techniques. Confusing terminology is also a significant hurdle for the practical development of learning tools.

The messy state of terminology is not very surprising since educational technology is an interdisciplinary endeavor. Even when we restrict our attention to personalization and adaptation, there are several communities studying similar questions: educational technology, learning analytics, educational

Faculty of Informatics, Masaryk University, Brno, Czech Republic; xpelanek@fi.muni.cz, Phone +420-549 49 6991

Table 1 A sample of general educational technology terms (all these terms have their own Wikipedia page and Google Scholar returns for them over 10000 results)

Adaptive learning	Intelligent tutoring system
Computer-assisted language learning	Interactive learning
Computer-supported collaborative learning	Learning management system
Computerized adaptive testing	Massive open online course
Digital learning	Personalized learning
E-learning	Programmed learning
Educational game	Virtual learning environment
Educational software	
Educational technology	

data mining, artificial intelligence in education, learning engineering. Moreover, researchers often have a primary background in other disciplines, e.g., psychometrics, instructional design, cognitive science, the science of learning, statistics, machine learning, game design, or computer science. Each of these areas has slightly different terminology, customs, and implicit meanings that researchers often carry over.

The goal of this paper is not to clear the mess and provide the correct terminology that everybody should use. Such a goal is currently infeasible. The paper has more modest yet still useful goals.

Firstly, we provide discussions of terms used with similar meanings, e.g., item, problem, task, question; or skill, ability, knowledge. Such listings of related terms provide quick orientation for newcomers to the field. Even for senior researchers, they can be a useful tool when searching for related research. A readily available list of alternatives to choose from can also lead to a better choice of terms and a better description of terminology in papers—awareness of alternative terms naturally leads to the need to explain the choice of terms; it is also useful to mention related terms for readers accustomed to different terminology.

Secondly, we discuss terms with multiple meanings, e.g., skill, domain model, or stratification. For these terms, we discuss their different meanings and show how the current practice of term usage can easily lead to confusion and misunderstanding.

The overall aim of the paper is to raise awareness of terminological issues in the area of adaptive learning systems and to call for more explicit discussion of used terminology in research papers. This should lead to better intelligibility and reproducibility of research papers and, in the long run, to the development of better learning systems.

The selection of discussed terms, together with their grouping and organization, was done based on an analysis of the following sources: keywords and terms used in highly cited and recent papers at relevant journals and conferences (particularly AIED, EDM, LAK), glossaries of related fields (e.g., psychometrics, pedagogy, cognitive science), e-learning standards (e.g., xAPI, IMS Global learning tools interoperability). The final presentation is also based on the author's experience with the design and development of learning systems and consultations with colleagues. As the goal of the paper is raising awareness and starting a discussion, the focus of the text is not on completeness but rather on readability and clear illustration of potential terminological problems on selected examples. The paper covers a wide range of terms, and most of these terms are discussed in many research papers. In the choice of literature references, priority was given to recent overview papers and typical applications, not to the original usage of terms.

2 Multiple Terms with Similar Meanings

We start with the less serious terminological problem: cases where multiple terms are used with similar meanings. In some cases, the terms are nearly complete synonyms; in other cases, they may have different but overlapping meanings. The fact that there are many related terms is not a fundamental obstacle for understanding research reports. It can, however, be an unnecessary hurdle in communication. The awareness of various terms is definitely advantageous: it can help to clarify potentially important nuances of meaning, to notice similarities between strands of research that differ on a superficial level, or to find relevant related research.

2.1 System Description

Table 1 shows over 15 terms for describing educational technology, and these are just the most mainstream terms. Many others are used in research papers. These descriptions often fit the pattern shown in Figure 1: an adjective word (which describes the specific focus of the particular technology), a process word, and a technology word. The individual words can be combined in a nearly arbitrary way; many combinations are actually used by at least some authors. Some of the resulting phrases are quite standardized and have a clear meaning (e.g., *computerized adaptive practice* or *intelligent tutoring system*). Many terms are used, however, quite freely and with highly overlapping meanings.

2.2 People

Many people are involved in the learning process; Table 2 provides an overview of terms used to denote them. There are several basic types of roles, and each of them can be expressed by several terms. Each of them carries a specific meaning, but often they are used as synonyms, and their usage is given by customs of a specific research community. However, the choice between these terms is mostly stylistic and does not significantly influence the understanding of research.

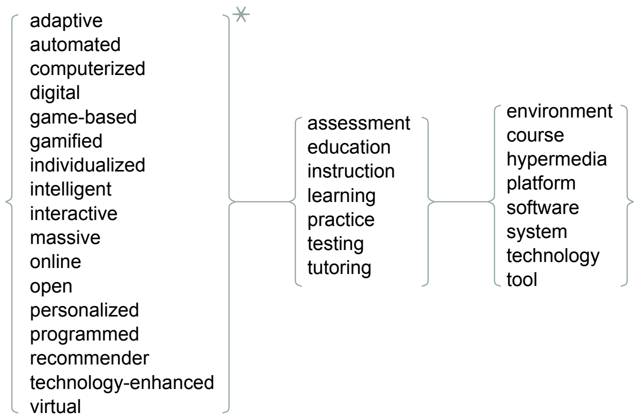

Fig. 1 Template for system description; the star over adjective words denotes that several of these words can be potentially combined

Table 2 Terms used to denote people involved in the learning process

somebody who is learning	student, learner, pupil, tutee, trainee, novice, beginner, user, participant
somebody who is guiding the learning	teacher, tutor, instructor, parent, teaching assistant, lecturer, mentor, coach, trainer, faculty member, supervisor, facilitator
somebody who is indirectly organizing the learning	administrator, manager, officer, staff member, policy-maker, stakeholder, observer, auditor, advisor
somebody who is creating learning materials	developer, designer, analyst, expert, content author, content writer, content developer, content creator
a group of people	team, group, crowd, cohort, class, grade, cluster

As a specific example, consider the key person in our area: somebody who is learning. This person is most commonly denoted as a *student* or a *learner*. In general usage, the term *student* tends to imply a formal educational setting, whereas the term *learner* is more general. In the context of computerized learning systems, most research is relevant in a general learning setting and does not necessarily require a formal educational setting. Therefore, strictly speaking, the term *learner* is often more appropriate. However, in practice, the usage of the term *student* is more common[1].

2.3 Learning Resources and Support

As a next step, let us consider the content of learning systems, i.e., the materials that the learner interacts with. On a general level, these materials may

[1] For example, according to Google Scholar search statistics, *student modeling* is twice more common phrase than *learner modeling*.

be called *learning objects* (Churchill, 2007) or *educational resources* (Hylén, 2006). These materials and related terms can be organized in many ways, e.g., Churchill (2007) provides one classification of learning objects. Table 3 gives an overview of terms with a focus on the manner of interaction.

Many of the terms are used as synonymous, often based on customs of a specific community, e.g., in psychometrics, a key tool is the item response theory, and thus the commonly used term is *item*. Particularly for the "resource that is studied" and "a group of resources," the choice of a specific term is mostly stylistic and reflects the specific educational setting and is not fundamental from the point of understanding the discussed ideas.

In the case of learning support, the terms often have distinct meanings and these differences can have important consequences for the design of learning environments. For example, the same type of explanatory message can be used either as a hint (shown during a solution process) or as a part of feedback (shown after an answer is submitted); hints are sometimes abused by students (Roll et al., 2014) and this should be taken into account in the design of the system; for feedback messages, this issue is not relevant. The timing and sequencing of learning support is often an important and nuanced issue, which depends on the exact meaning of terms; see, e.g., a discussion of feedback by Hattie and Timperley (2007). At the same time, the meaning of terms for learning support is not completely distinct; there are often significant overlaps among them. This part of the terminology would clearly benefit from a detailed clarification.

One important distinction among terms in Table 3 is that they are often used to refer to notions of different granularity. This distinction may be necessary for understanding, particularly when several of these terms are used together. For example, a *task* typically consists of several *steps*, but an *item* can be in certain contexts used as a synonym for both a *step* and a *task*. This type of relation between terms is often used implicitly by authors and may not be completely clear to readers with a different background. This is one of the cases where explicit clarification of terminology would be beneficial.

The structure of materials is probably terminologically most confusing. The same structure can be described by different authors as a *domain model*, *knowledge graph*, or *ontology*. More confusingly, terms like *domain model* can be used in several different meanings. We discuss this important topic in more detail in Section 3.1.

2.4 Types of Interaction

We now look in more detail at the category of "resource that is solved" since this category is a key one for adaptation. Table 3 provides an overview of terms that denote the complete resource. The interaction with the resource ("solving") can be realized in many ways, and this type of interaction is, in most cases, not implied by the term for the resource.

Table 3 Terms used to describe learning resources and their structure

resource that is solved	item, question, problem, exercise, activity, task, step, quiz, assessment event, flashcard, (serious) game, puzzle, homework, assessment, assignment, challenge
resource that is explored	simulation, model, microworld, multimedia, tutorial
resource that is studied	book, textbook, paper, instruction, lecture, lesson, text, audio, video, presentation, slides, animation, resource, material, media, chapter, document
learning support	scaffoldings, hints, help, feedback, cues, signalings, guidance, prompts, analogies, summaries, worked-out examples, step-by-step solutions, fading, faded examples, explanations, prompted self-explanations, (highlighted) text annotations
a group of resources	skill, knowledge component, concept, rule, schema, procedure, category, problem set, item set, level, topic, course, sequence, chunk, module, chapter, unit, battery, testlet, learning path, syllabus, collection, tag
structure of materials	domain model, skill model, knowledge structure, knowledge graph, curriculum, ontology, taxonomy, folksonomy

Table 4 Types of interaction for resources that are solved

user interface	menu, drag-and-drop, drop-down
selected response	multiple-choice questions, alternate-choice question, true/false, cloze (with menu)
multiple selections	matching, ordering, sequencing, categorization, tagging
constructed response	written answer, completion, fill-in-the-blank, open response, short answer, extended response, free-form question, open-ended question, essay
problem solving	interactive exercise, multi-step problem, performance test

Table 4 provides an overview of terms used to describe the type of interaction. The table uses categories based on the classification of exercise types (Pelánek, 2020a). However, for the displayed terms, the classification provides only a basic overview as the terms are used with different types of meanings. Terms like *drag-and-drop* focus on the user interface aspect of the interaction. The term *fill-in-the-blank*, on the other hand, describes the basic pedagogical principle (filling missing information into the given context) and can be realized in many ways: the basic one is written answer into the blank, but it can also be done by different forms of selected response from provided options. The relations between these terms are complex, including subsumption, synonyms, and partial overlap.

In this case, there is also an issue with multiple meanings per term, particularly concerning the distinction between "instance" and "type of" with respect to exercises, problems, or questions. Suppose that a model description states that "one input to the student model is exercise identification." What exactly does it mean? One possibility is that the model uses the identification of the type of exercise, e.g., whether the answer was to a multiple-choice question or constructed response question. Another possibility is that the model uses the identification of a specific instance, e.g., "3+8=__". There are also other possibilities in between these extremes, e.g., "constructed response for one-digit addition." These different interpretations have significant consequences for understanding and using student models.

2.5 Personalization Algorithms

One of the key goals of educational technology is to make learning personalized, i.e., to adapt the learning process to the needs and preferences of a particular learner. Several general terms are used for this type of personalization algorithms, e.g., *tutor model* (Sottilare et al., 2013), *instructional policy* (Käser et al., 2016), *pedagogical policy* (Iglesias et al., 2009), *recommendation system* (Manouselis et al., 2011), *adaptive navigation* (Brusilovsky and Pesin, 1998), or *mastery learning* (Pelánek and Řihák, 2018).

2.5.1 Level of Adaptivity

Adaptivity can be achieved at different granularities. Although there is a rather continuous spectrum of adaptivity approaches, there are three main granularity steps typically discussed in the literature:

– Adaptivity within solving one item, task, or problem, which can contain several steps. This typically involves personalizing various forms of learning support (hints, scaffoldings, feedback, explanations). This form of adaptive behavior can manifest several times per minute.
– Adaptation within larger steps, e.g., the choice or recommendation of exercises and topics to study. This type of adaptation typically manifests once every few minutes.
– Adapting the system, e.g., adding or removing items and knowledge components or changing the setting of used algorithms. This adaptation is based on data about student performance. It can be fully automated but often is supervised by a human. It takes place at the timeframe of days or months.

Table 5 lists the terms used for these types of adaptivity. The *inner loop, outer loop* terminology was introduced by Vanlehn (2006), who later proposed a more general framework of *regulative loops* (VanLehn, 2016). Aleven et al. (2016) provides a comprehensive discussion of this topic; they propose an "adaptivity grid" and use the terminology *step loop, task loop, design loop.* The terms *micro-adaptation* and *macro-adaptation* are used, for example, by Essa (2016).

The design loop adaptation is probably terminologically the least standardized. It is also denoted as *closing the loop* (Liu and Koedinger, 2017) or

Table 5 Terms used to describe different adaptivity levels

adaptation within one item	inner loop, micro-adaptation, step loop
adaptation within larger steps	outer loop, macro-adaptation, task loop
adapting the system by a human	design loop, closing-the-loop, human-in-the-loop

human-in-the-loop (Pelánek, 2017). It is often discussed in the literature as a research result without any specific title.

2.5.2 Spaced Repetition

One of the learning principles best supported by research evidence is the *spacing effect*, i.e., people remember better when using short study periods spread over time as opposed to the massed practice (Kang, 2016; Settles and Meeder, 2016).

This principle and its realization in educational technology are described in the literature using many terms with overlapping meaning:

- *spaced repetition* (Kang, 2016; Settles and Meeder, 2016),
- *distributed practice* (Rohrer, 2015),
- *spacing, practice, forgetting effects* (Pavlik Jr and Anderson, 2005),
- *optimal schedule of practice* (Pavlik and Anderson, 2008), *optimal gap* (Cepeda et al., 2008),
- *flashcards* (Kornell, 2009), *Leitner system* (Reddy et al., 2016; Settles and Meeder, 2016).

Moreover, the spacing effect has connections and interactions with other terms. It is a special case of *desirable difficulty*, i.e., a task that causes short-term difficulties but increases long-term performance (Bjork et al., 2011). The benefits of the spacing effect are typically combined with the *testing effect* and *retrieval practice*, i.e., trying to remember information instead of restudying (Roediger III and Butler, 2011). The spacing effect can also interact with *interleaved practice*, where the practice of different topics is done in an interleaved manner (Taylor and Rohrer, 2010).

2.5.3 Appropriate Challenge

A typical goal of personalization algorithms is to achieve an appropriate challenge for learners, i.e., to confront them with problems that are neither too easy nor too difficult. This approach is grounded in pedagogical and psychological concepts and theories, e.g., *the zone of proximal development*, *the concept of flow* (Nakamura and Csikszentmihalyi, 2014), *inverted-U hypothesis* (Lomas et al., 2013), or *instructional scaffolding* (Jumaat and Tasir, 2014).

One approach to achieve an appropriate challenge is to adaptively pick items from a large pool of options. This approach is used widely in adaptive systems but lacks a clear and distinct descriptive term. It is discussed under many different names, e.g., *adaptive choice of questions* (Pelánek et al., 2017) or *item selection* (Klinkenberg et al., 2011). These techniques mimic the use of techniques from computerized adaptive testing. In testing, a common approach is to use item response theory (De Ayala, 2008) to find questions where the predicted probability of correct answer is near 50% since these questions provide the most information about student state. In the case of adaptive practice, tools typically aim at a higher target success rate (e.g., 75%) in order to not frustrate learners.

In the knowledge space literature, the idea of an appropriate challenge is captured under the term *outer fringe*, which is the most advanced part of knowledge space for which the learner has sufficient knowledge (Falmagne et al., 2013; Doignon and Falmagne, 2016; Doble et al., 2019).

The aim of an appropriate challenge is widely used in computer games, and it is often achieved using techniques similar to these used in learning environments. The terms used to describe these techniques are, however, quite different, e.g., *dynamic difficulty adjustment* (Hunicke, 2005) or *difficulty curve* (Sarkar and Cooper, 2019).

2.6 Student State

To achieve personalization, a learning system typically uses some kind of student personal data and state estimates. Personal data typically consist of demographic information; in this case, the terminology is relatively standardized and unproblematic. Student states, properties, and behaviors are more terminologically challenging. As Table 6 shows, there are several basic dimensions of student state, and each of these can be expressed using different terms.

For the cognitive state, the listed terms are often used as synonyms based on customs in different communities. There are differences in the granularity of learning units with respect to which the terms are used. The term *skill* is typically used for fine-grained units (e.g., the addition of fractions), whereas *competency* is used for coarser units (e.g., using fractions to solve practical problems). Research and development in AIED typically deal with fine-grained units. The main terminological issues with terms for cognitive states is that each of them can be used with several different meanings; we discuss this topic in more detail in Section 3.3.

For affect, emotion, meta-cognition, and long-term traits, the landscape of terms is much broader and more complex. In addition to general terms used to describe them, there are many specific states and properties. In learning environments, a focus is mainly on emotions and affective states related to how interested, engaged, and concentrated the student is; this can be expressed using several terms with overlapping meaning.

Table 6 Terms used to describe student states, properties, and behaviors

cognitive state	skill, ability, knowledge, proficiency, competence
affect: general	affect, emotion, valence, arousal
affect: engagement	motivation, engagement, concentration, boredom, flow, engaged concentration, immersion
affect: other	confusion, frustration, confidence, fatigue, delight, joy
long term traits	grit, drive, growth mindset, goal orientation, self-efficacy, competence, connection, autonomy, agency, attribution, self-regulation, help-seeking, interests, attitudes
counterproductive behaviors	gaming the system, systematic guessing, hint abuse, help abuse, off-task behaviour, wheel-spinning, cheating, plagiarism, multiple-account cheating, procrastination, drop out

Student affect is interlinked with the occurrence of various forms of counterproductive behaviors. For these, there is again a wide range of specific terms. These often have some specific meaning explicitly described by authors, e.g., *gaming the system* (Baker et al., 2008), *off-task behavior* (Baker, 2007), *wheel-spinning* (Beck and Gong, 2013). However, the meaning of these terms often overlaps, e.g., systematic guessing may be seen as a form of gaming the system (but only when the learning environment does not penalize it).

2.7 Modeling

The adaptive behavior of educational technology is based on modeling, particularly modeling of the learning domain and student states. In this area, the terminology is important and unclear since terms often have several possible meanings. We discuss these topics in more detail in the next section. Here, we go briefly over cases where we have several modeling terms with similar meanings.

2.7.1 Student Performance Data

A key input to student models are data about student performance. Models also often predict future performance, and these predictions are used to evaluate and compare models. It is thus very important what aspects of student performance are incorporated into student modeling and what terms we use to describe them.

Table 7 lists some terms used for this purpose. These terms and their combinations typically carry specific meaning, which is, however, often assumed implicitly. For illustration, consider the phrases *unsolved task*, *unfinished attempt*, and *incorrect answer*. A typical meaning behind these phrases could be:

– *unsolved task* = student did not yet try to solve the task,

Table 7 Terms used to describe the evaluation of student performance

what is evaluated	answer, solution, step, task, attempt, submission, session
evaluation is called	result, score, grade, performance, classification
positive evaluation	correct, right, solved, finished, successful, completed, tackled
negative evaluation	incorrect, wrong, unsolved, unfinished, unsuccessful, failed, error, erroneous, mistake, misconception, missed, skipped

- *unfinished attempt* = student tried, did not yet succeed, but can possibly succeed on his own in future (the solution was not shown to him),
- *incorrect answer* = student did try, did not succeed and another chance in the future is not meaningful, since he was informed about the correct choice.

These meanings are definitely not fixed. There can be many other shades of their meaning, the boundary between them is quite fuzzy, and their usage by authors differs. However, a subtle difference in meaning can have quite a significant impact on student modeling. Should an unfinished attempt be used as evidence of weak cognitive skill or rather low engagement? That depends on the exact meaning of the term.

2.7.2 Difficulty and Complexity

One common goal of adaptive educational technology is to appropriately tune the difficulty of learning. In order to do so, we need to quantify the difficulty of learning resources and capture it in models. Here the primary term is *difficulty.* Alternatively, some authors use *easiness*, which is usually just an inverse of difficulty. A closely related but distinct term is *discrimination*, which is a measure of how well an item distinguishes between learners of different abilities (De Ayala, 2008). Additional terms describe specific difficulty measures or specific aspects of difficulty, e.g., *success rate*, *failure rate*, *response time*, *time intensity*, *cognitive load*, *workload.*

A more nuanced distinction is between *difficulty* and *complexity.* These terms are sometimes used in similar meaning, but authors that explicitly discuss them mostly agree on their distinct meaning (Liu and Li, 2012; Beckmann et al., 2017): complexity is an intrinsic property of a task and is given by its internal structure, whereas difficulty is related to student-task interaction (performance of students).

2.7.3 Modeling Terms

Table 8 lists several other cases where we have multiple modeling terms with similar or overlapping meanings.

The terms used to describe model elements are mostly synonymous; they are just used by different communities, e.g., *feature* in machine learning, *covariate* in statistics, *construct* in psychology.

Table 8 Modeling terms

model element	parameter, feature, construct, factor, attribute, covariate, variable, predictor
setting model parameters	parameter fitting, parameter estimation, learning, calibration, normalization, standardization, equating, scaling
visualization of model elements	dashboard, open learner model, activity visualization, progress bar, skillometer, badge, leaderboard, achievement, homework statistics
desirable model properties	accuracy, reliability, validity, generalizability, portability, robustness, resolution, interpretability, explainability, identifiability, group invariance

The model parameters, whatever they are called, need to be set somehow, i.e., we need some procedure to find good values of parameters. This is a central step in model building and there are multiple terms to describe such a procedure. Some of these terms have very close meaning (e.g., *parameter estimation* and *parameter fitting*), some are used with rather specific and distinct meaning (e.g., *equating* in item response theory).

Once we have the model parameters, we often want to display some of them to users in the form of a suitable visualization. This type of visualization is again described by a variety of terms, each with a specific meaning but with significant overlaps.

Models have many desirable properties that we would like to achieve. Most of the terms used to describe these properties have a quite clear and distinct technical meaning. Nevertheless, there are many nuances concerning their exact meaning and relations, and these nuances can be quite important since they have a significant impact on the way we perform evaluation and comparison of models, which subsequently influence the behavior of learning environments. In fact, the exact meaning of these terms is often a topic of a separate discussion, e.g., see Cook and Beckman (2006) for discussion of *reliability* and *validity* or Rudin (2019) for discussion of *interpretability* and *explainability*.

3 Multiple Meanings for the Same Term

Now we discuss cases where the same term is used with multiple meanings. This is a more serious terminological problem since it can be the cause of confusion and misunderstanding.

In current educational technology, one common source of multiple meanings for a single term is the usage of the same terms for describing human learning and machine learning, e.g., *transfer*, *transfer learning*, *active learning*, *supervised learning*, *feedback*, *bias*, *long term memory*. These terms often have quite specific, technical meaning in machine learning literature, whereas in pedagogy, they are used with a meaning that is quite different and more general. However, the basic context of their usage is typically quite clear, and thus in practice, they do not cause significant problems for understanding.

In our discussion, we focus on terms that, even within a very specific context, can have several meanings, and the differences between these meanings may be quite fundamental for understanding. The discussed terms are related to modeling and evaluation.

3.1 Domain Model

Any educational technology needs to have some representation of the domain that it is trying to teach its users. This representation is often called *domain model*, although other terms are also used, e.g., *skill model* or *knowledge structure*. Although the domain model is a key component of any educational technology and its development and improvement is the subject of many research papers, it is seldomly explicitly defined or even described[2].

Unfortunately, the term can have many specific meanings, and the absence of a clear description can lead to confusion. The domain model can be concerned with the following aspects:

- definition of knowledge components, the choice of their granularity,
- mapping of items to knowledge components,
- relations among knowledge components, particularly prerequisite relations and subsumption relations, but potentially even more complex relations captured by a general ontology,
- description of cognitive processes, e.g., using production rules or constraints, possibly including also misconceptions ("buggy rules").

We discuss individual types of domain models in more detail. To make the discussion clearer, we illustrate all types of domain models on the example of fractions, specifically using the addition of fractions as the main illustration (Fig. 2).

Note that below we discuss only different conceptual views of domain modeling. In addition to these, the term domain modeling is also relevant from the software engineering perspective (Evans, 2004). The developers who implement learning systems often use the term domain modeling with a specific, technical focus, e.g., to describe the representation of relations in a database. This is an additional source of confusion.

3.1.1 Mapping Items to Knowledge Components

One approach to domain modeling is to focus on the definition of knowledge components and the mapping of items to knowledge components. This

[2] One example of an explicit definition is by Sottilare et al. (2016): "The domain model contains the set of skills, knowledge, and strategies/tactics of the topic being tutored. It normally contains the ideal expert knowledge and also the bugs, mal-rules, and misconceptions that students periodically exhibit."

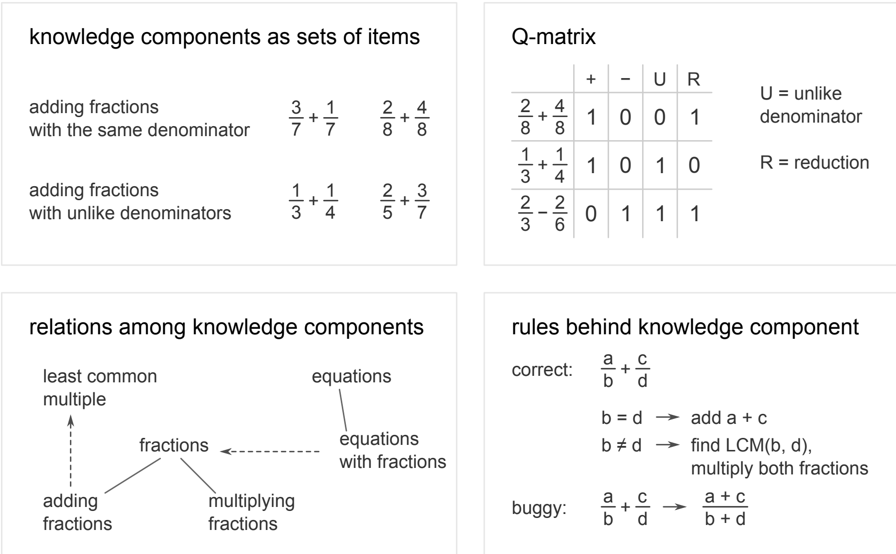

Fig. 2 Illustration of different approaches to domain modeling on the example of fractions

mapping can take the form of disjoint sets, or it can be a more general N:M mapping, which is commonly called *Q-matrix* (Tatsuoka, 1983; Barnes, 2005). Fig. 2 gives examples for the case of fractions.

An important part of this approach to domain modeling is the definition of suitable knowledge components and the choice of their granularity (Koedinger et al., 2012; Pelánek, 2020b). In the case of fractions, this leads to questions like: Should adding fractions with the same and unlike denominators be two separate knowledge components? Should the addition and subtraction of fractions be two separate knowledge components?

This approach to domain modeling is used, for example, by John et al. (2015): "a domain model captures relationships between the learning objects and the knowledge components or skills they exercise."

3.1.2 Relations among Knowledge Components

Another possible focus of domain modeling is on the relations among knowledge components. Two major relations are subsumption (more/less general) and prerequisite relations. Fig. 2 provides an illustration of these kinds of relations for several knowledge components involving fractions.

This approach to domain modeling is closely related to taxonomies and ontologies and to the terminology used in these fields[3]. It is also used in knowledge space theory (Falmagne et al., 2013; Doignon and Falmagne, 2016), which focuses particularly on the prerequisite relation.

[3] Wikipedia page for *domain model* has the definition "In ontology engineering, a domain model is a formal representation of a knowledge domain with concepts, roles, datatypes, individuals, and rules, typically grounded in a description logic."

3.1.3 Rules, Constraints, Principles

The previous approaches to domain modeling treat individual knowledge components basically as elementary units (represented by atomic identificators)—they describe their relations, but not any principles behind them. A significantly different understanding of domain modeling is to focus on just these internal principles.

These principles can be described as rules for solving items belonging to a knowledge component (as illustrated in Fig. 2). With this approach, it is often useful to describe not just the correct rules but also "buggy" rules and common misconceptions. This rule-based approach to domain modeling (sometimes called *cognitive modeling*) is often used in *intelligent tutoring systems* (specifically *cognitive tutors*) and allows them to provide personalized hints and feedback during the problem-solving process (Aleven, 2010).

Instead of rules, the principles behind a domain can also be described using constraints. In this case, a domain model is a set of constraints (Martin and Mitrovic, 2003). For specific examples, including fractions, see Mitrovic (2010).

A less typical example of this approach to domain modeling is provided by Wang et al. (2005), who study creative problem solving and describe a domain model as "simply a special case of user model summarized from domain experts with a different building process."

3.2 Student Modeling

In general, the goal of student modeling is to estimate the state of a student. However, there are many aspects of student state that can be modeled, e.g., cognitive, affective, or motivational (Pelánek, 2017; Chrysafiadi and Virvou, 2013). Very different models can thus be described by the same label *student model*.

A specific term that is used very often in student modeling and that can be potentially confusing is *knowledge tracing*. This term is used with two different meanings:

1. as a general term for any model that tries to estimate dynamically changing knowledge of students, i.e., tracing knowledge as it changes through time,
2. as a term denoting a very specific type of model (also called Bayesian knowledge tracing, BKT), which makes very specific assumptions about learning, e.g., that the latent knowledge state is binary (van De Sande, 2013; Pelánek, 2017).

3.2.1 Student and Domain Model

In the case of modeling knowledge (cognitive state), student modeling is inherently interconnected with domain modeling—the knowledge is always with respect to some subset of the domain. A *student model* thus must be based in some way on a *domain model*. However, it is often not clear where one ends, and the other begins. For example, knowledge tracing models contain parameters like learning rate and slip rate, which are specific for each knowledge component. In the commonly used version of the model, these parameters are not individualized, i.e., they are parameters of knowledge components and model some aspect of a domain, so they should be part of a domain model. At the same time, they may be influenced by a specific student population and are used to estimate student knowledge; in practice, they are typically included in the student model. Sometimes, the student and domain models may not be explicitly distinguished, and domain modeling is hidden under the term student modeling.

3.2.2 Different Perspectives on the Nature of the Model

The term *student model*, even when referring to one precisely defined type of model (like Bayesian knowledge tracing), can be used with several different meanings depending on the perspective in which it is used:

- A teacher's perspective: The reality that we are modeling are the real skills of students. The model is a simplification of this reality, i.e., the skill estimates for individual students. This perspective is not much concerned with how these estimates are obtained.
- A researcher's perspective: The reality that we are modeling is the learning process in a particular domain. The model is a simplification of this reality, i.e., the assumptions that the model makes and its basic functional form. This perspective is not much concerned with specific parameter values or a parameter fitting procedure.
- A developer's perspective: The model is what needs to be implemented, i.e., the specific data attributes, equations, and parameter fitting procedures.

These perspectives are often implicit in discussions of models; making them explicit would often help to simplify communication and clarify expectations. In the current state, developers may find many research papers disappointing since their description of student models is far from sufficient from the developer's perspective.

Let us consider a specific aspect of student models that is unclear and, at the same time, can be important for understanding and reproducing research findings: Is parameter fitting considered to be part of a model? In some cases, it is not. For example, it makes sense to consider Bayesian knowledge tracing as the same model of learning regardless of whether we use expectation maximization or brute force to fit its parameters. In other cases, the parameter fitting is (implicitly) part of the model. For example, many logistic models take a very similar functional form (Rasch, Elo, AFM, PFA), and the main difference is sometimes in the way parameters are computed. A specific example is the Rasch model and the Elo rating system (Pelánek, 2016).

3.2.3 Types of Parameters and Parameter Fitting

Another important and terminologically unclear aspect of student models are parameters. There are different types of parameters and different ways to obtain the values of these parameters. It is useful to distinguish at least three types of parameters:

– *Online parameters.* These parameters are updated after each new observation (datapoint). In most student models, student skill is this type of parameter. The computation of these parameters is usually done using a computationally efficient approach (often using some update equation) and is not called parameter fitting. These update equations are specific for each type of model.
– *Offline parameters.* These parameters are updated automatically, but not frequently, using a large dataset of observations. In student models, typical examples of this type of parameter are knowledge component parameters like students' learning rate or guess rate. The values of these parameters are found using some kind of computationally demanding search, which is called *parameter fitting* or *parameter estimation.* The parameter fitting procedures are often general, e.g., gradient descent, expectation-maximization algorithm, or grid search.
– *Hyperparameters.* These parameters are not part of a model itself (they do not model any aspect of reality); they determine the behavior of a parameter fitting procedure. Typical examples are regularization parameters or the number of iterations. These parameters are typically set manually by researchers based on intuition, experience, or experimentation.

The boundary between these classes of parameters is blurry. As an example, consider a model with parameters for the mean and standard deviation of the prior distribution of student skills. These parameters are on the boundary between offline parameters and hyperparameters; their treatment depends on the details of the model and its usage.

The above-given labels for different types of parameters are not standard—there is currently no clear terminology for distinguishing types of parameters. This is unfortunate since the distinction between them is quite important both for practical applications (it has significant consequences for computational complexity) and for understandability and replicability of the evaluation of models. For example, item difficulty may be treated both as online and offline parameters.

3.3 Skill

A term with a particular predisposition to cause confusion is *skill.* This term is used with many different meanings. The two basic ones are as a synonym to ability, i.e., denoting students' skill, and as a knowledge component, i.e., denoting a part of a domain model (Koedinger et al., 2012). These two meanings are significantly different. Even within these basic meanings, we can find important distinctions. For our discussion, we group different meanings into four groups.

3.3.1 Skill as a Latent Construct

One way to understand the term *skill* is to take the perspective of a teacher or a researcher in cognitive science. With this perspective, skill is "what is inside learners head," i.e., the neural representation of the skill from a biological point of view or multifaceted cognitive construct from a psychological point of view. This is the latent construct that we really care about but that we cannot easily grasp or measure.

This understanding of skill is clearly quite complex even for fine-grained skills like the addition of fractions. For example, when we consider the teacher's perspective of skill, it may be something like "Peter understands the basic rules for adding fractions, but he is not fluent yet and often makes careless errors, particularly in the case of unlike denominators." This complex verbal formulation is still a simplified abstraction of the real underlying skill.

In this context, the term *skill* has a similar meaning as *ability*, *proficiency*, or *knowledge*. However, in certain situations, there may be important nuances hidden behind the usage of these terms (e.g., *skill* may be used to denote procedural knowledge, but not declarative knowledge).

3.3.2 Skill as Model Estimate

In learning systems, we cannot work with the latent construct itself. Thus we work with some simplified representation of the hidden inner state of students. This simplification is typically a one-dimensional numerical representation, i.e., Peter's addition of fractions skill becomes something like 0.6.

Even when considering the skill as a student parameter expressed by a number, it can have different meanings:

1. The number (skill) expresses the uncertainty of the estimate of the underlying concept. Since the skill used in a model is based on limited observation of the student, it is just a statistical estimate with some uncertainty. With this understanding, the interpretation of "Peter's skill is 0.6" is "based on the observations, there is 60% chance that Peter has already mastered addition of fractions."
2. The number (skill) expresses the degree of knowledge, i.e., how large part of the knowledge component the student already mastered. With this understanding, the interpretation of "Peter's skill is 0.6" is "Peter can solve 60% easiest items concerning the addition of fractions."

The confusion between these two meanings can lead to problems with setting criteria for mastery learning (Pelánek, 2018a).

In this context, the term *skill* can often be interchanged with *ability*, *proficiency*, or *knowledge*; the usage is often given by customs of the specific research area. For example, Bayesian knowledge tracing research typically uses the term *skill*, whereas Item response theory research uses the term *ability*; both models can be used to denote the model estimate of the same (or very similar) latent construct.

3.3.3 Skill as a Set of Items

The term *skill* is also used in a domain model perspective where it is independent of individual students. Primary usage in this context is as a synonym for *knowledge component* or *concept*, i.e., corresponding to a set of items or to a column in a Q-matrix (see illustrations in Fig. 2). Note that in this context, as opposed to the previous one, it is not possible to interchange the term *skill* with *ability*, *proficiency*, or *knowledge*.

3.3.4 Skill as a Rule

Correspondingly to multiple usages of domain modeling, the term skill can also have several meanings from the domain modeling perspective. Another view is of skill as a set of rules or constraints, which can be either formal or informal. Fig. 2 shows basic rules for the addition of fractions; these can be elaborated into a full-fledged formal model, and the term *skill* can be used to describe this model. In other cases, it may be unrealistic to fully formally describe all rules. As an illustration, consider the distinction between continuous and simple present tense in English. Yet, we may still refer to these implicit rules as a *skill*.

3.4 Evaluation

Evaluation is a key part of the development and research in educational technology. This area is also full of terminological pitfalls. In this case, the issues are mostly not specific to educational technology—they are often encountered also in other applications of machine learning or statistics. We thus cover only briefly some of the important terms with several potential meanings.

Accuracy can be used to talk about general predictive properties of models or as one specific technique for evaluation of binary classification (the proportion of correct prediction to all predictions). Methods for measuring predictive accuracy are usually called *metrics* or *measures*. The term *metric* has in mathematics a very specific, technical meaning (a distance function satisfying several requirements). In the context of the evaluation of predictive accuracy, however, it is mostly used in a sense of “any function that is used to make comparisons.”

A common approach to assessing the generalizability of models is to use *cross-validation*. This typically involves splitting data into several sets and using separate data for training and evaluating models. The specific meaning of *cross-validation* can, however, significantly differ. Specifically, there is confusion in the usage of terms *testing set* and *validation set*, which have a standard meaning but are often used in reverse (Ripley, 2007). A division of data into individual cross-validation sets is often done in some specific way. To describe this process, researchers sometimes use the term *stratification*. Unfortunately, the term is used in at least two very different meanings: to ensure that the class distribution in each cross-validation set is approximately the same as in the initial dataset or to ensure that all data for a single student (or item) are in a single set (Pelánek, 2018b).

Several other evaluation terms are common sources of problems, not just in the evaluation of educational technology. A typical example is the term *significance*. Some occurrences of the term in research papers are with the common-sense meaning (subjective significance). A standard research meaning is a statistical significance, often with an implicitly implied p-value level 0.05. Recently, however, there has been a backlash against the (over)usage of statistical significance (McShane et al., 2019). Specifically, in experiments with educational technology, we often have very large sample sizes, and with large data, we can often achieve statistical significance without practical significance. One possible approach to measuring practical significance is the quantification of *effect size*, which can be computed in many different ways (Fritz et al., 2012). Another related, useful, but overloaded term is *error bar*. Error bars are used to depict variability of data in graphs; they can, however, be computed in several different ways, e.g., standard deviation, interquartile range, confidence intervals computed by a formula based on specific assumptions or by bootstrapping (Cumming et al., 2007). The exact meaning of terms like *effect size* or *error bar* may be important for understanding and interpreting research results.

4 Conclusions

Our discussion of educational technology terminology is definitely not complete; completeness is not the aim of the work. The main point is to highlight specific cases where unclear terminology can lead to confusion. It turns out that there are many such cases. For example, phrases like “a technique for improving domain model” or “an algorithm for estimating skill” have several significantly different meanings.

The inconsistent state of terminology has several negative consequences that hinder the progress of the field. A multitude of terms with similar meanings complicates finding existing research and can lead to reinventing the wheel. Multiple meanings of a single term make reading and understanding research papers more difficult. Unclear terminology is also one of the obstacles to the reproducibility of research. For example, the details of cross-validation methodology are seldom described in sufficient detail and with sufficient clarity of used terms to allow replication of experiments.

Although terminology may seem like an academic topic, it is also very important for the practical development of learning environments. Clear terminology within the development team is very useful: it simplifies communication, prevents bugs in the code, and leads to an easier application of research results. In fact, the original impulse for this article came from addressing terminological issues in the practical development of a learning environment.

The current state of educational technology terminology has deep roots, particularly the inherently interdisciplinary nature of the field. It is thus not realistic to expect that the terminological maze will disappear. Nevertheless, we can surely do better at the navigation of the maze. The basic step is to discuss terminology more explicitly in research papers, particularly to specifically describe the used meaning for overloaded terms like *skill*, *domain model*, or *accuracy*. In cases of multiple terms with similar meanings (e.g., *item*, *problem*, *question*, *task*), it is useful to mention alternative terms and describe reasons for the particular choice.

For specific cases where the state of terminology is particularly confusing, it may be fruitful to perform an in-depth analysis of the current state of term usage or to collect and compare opinions of AIED practitioners with different backgrounds. Examples of such cases are:

1. Domain modeling, where it would be useful to have further clarification and clear terminology for different approaches to domain modeling as outlined in Section 3.1.
2. Student modeling, where it would be particularly useful to further clarify different meanings of the term *skill* and relations to related terms like *ability*, *knowledge component*, or *concept*.
3. Learning support, as expressed by terms like *feedback*, *hint*, *explanation*, *scaffolding*, *fading*, *worked-out examples*. Terms in this area typically have distinct but overlapping meanings. Clarification of these terms can be, among others, very useful for the application of research results into practical applications.

Acknowledgment

The author thanks Ryan Baker, Michael Yudelson, Ken Koedinger, Tomáš Effenberger, Jaroslav Čechák, and Valdemar Švábenský for their comments, ideas, and suggestions.

References

Aleven V (2010) Rule-based cognitive modeling for intelligent tutoring systems. In: Advances in intelligent tutoring systems, Springer, pp 33–62

Aleven V, McLaughlin EA, Glenn RA, Koedinger KR (2016) Instruction based on adaptive learning technologies. Handbook of research on learning and instruction pp 522–560

Baker R, Walonoski J, Heffernan N, Roll I, Corbett A, Koedinger K (2008) Why students engage in "gaming the system" behavior in interactive learning environments. Journal of Interactive Learning Research 19(2):185–224

Baker RS (2007) Modeling and understanding students' off-task behavior in intelligent tutoring systems. In: Proceedings of SIGCHI conference on Human factors in computing systems, pp 1059–1068

Barnes T (2005) The q-matrix method: Mining student response data for knowledge. In: American Association for Artificial Intelligence 2005 Educational Data Mining Workshop, pp 1–8

Beck JE, Gong Y (2013) Wheel-spinning: Students who fail to master a skill. In: Proceedings of Artificial Intelligence in Education, Springer, pp 431–440

Beckmann JF, Birney DP, Goode N (2017) Beyond psychometrics: the difference between difficult problem solving and complex problem solving. Frontiers in psychology 8:1739

Bjork EL, Bjork RA, et al. (2011) Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. Psychology and the real world: Essays illustrating fundamental contributions to society 2(59-68)

Brusilovsky P, Pesin L (1998) Adaptive navigation support in educational hypermedia: An evaluation of the isis-tutor. Journal of computing and Information Technology 6(1):27–38

Cepeda NJ, Vul E, Rohrer D, Wixted JT, Pashler H (2008) Spacing effects in learning: A temporal ridgeline of optimal retention. Psychological science 19(11):1095–1102

Chrysafiadi K, Virvou M (2013) Student modeling approaches: A literature review for the last decade. Expert Systems with Applications 40(11):4715–4729

Churchill D (2007) Towards a useful classification of learning objects. Educational Technology Research and Development 55(5):479–497

Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. The American journal of medicine 119(2):166–e7

Cumming G, Fidler F, Vaux DL (2007) Error bars in experimental biology. The Journal of cell biology 177(1):7–11

De Ayala R (2008) The theory and practice of item response theory. The Guilford Press

van De Sande B (2013) Properties of the bayesian knowledge tracing model. Journal of Educational Data Mining 5(2):1–10

Doble C, Matayoshi J, Cosyn E, Uzun H, Karami A (2019) A data-based simulation study of reliability for an adaptive assessment based on knowledge space theory. International Journal of Artificial Intelligence in Education 29(2):258–282

Doignon JP, Falmagne JC (2016) Knowledge spaces and learning spaces. New handbook of mathematical psychology pp 274–321

Essa A (2016) A possible future for next generation adaptive learning systems. Smart Learning Environments 3(1):16

Evans E (2004) Domain-driven design: tackling complexity in the heart of software. Addison-Wesley Professional

Falmagne JC, Albert D, Doble C, Eppstein D, Hu X (2013) Knowledge spaces: Applications in education. Springer Science & Business Media
Fritz CO, Morris PE, Richler JJ (2012) Effect size estimates: current use, calculations, and interpretation. Journal of experimental psychology: General 141(1):2
Hattie J, Timperley H (2007) The power of feedback. Review of educational research 77(1):81–112
Hunicke R (2005) The case for dynamic difficulty adjustment in games. In: Proceedings of Advances in computer entertainment technology, pp 429–433
Hylén J (2006) Open educational resources: Opportunities and challenges. Proceedings of Open education 4963
Iglesias A, Martínez P, Aler R, Fernández F (2009) Reinforcement learning of pedagogical policies in adaptive and intelligent educational systems. Knowledge-Based Systems 22(4):266–270
John RJL, McTavish TS, Passonneau RJ (2015) Semantic graphs for mathematics word problems based on mathematics terminology. In: EDM (Workshops)
Jumaat NF, Tasir Z (2014) Instructional scaffolding in online learning environment: A meta-analysis. In: 2014 International Conference on Teaching and Learning in Computing and Engineering, IEEE, pp 74–77
Kang SH (2016) Spaced repetition promotes efficient and effective learning: Policy implications for instruction. Policy Insights from the Behavioral and Brain Sciences 3(1):12–19
Käser T, Klingler S, Gross M (2016) When to stop? towards universal instructional policies. In: Proceedings of Learning analytics & knowledge, pp 289–298
Klinkenberg S, Straatemeier M, van der Maas HL (2011) Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. Computers & Education 57(2):1813–1824
Koedinger KR, Corbett AT, Perfetti C (2012) The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. Cognitive science 36(5):757–798
Kornell N (2009) Optimising learning using flashcards: Spacing is more effective than cramming. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition 23(9):1297–1317
Liu P, Li Z (2012) Task complexity: A review and conceptualization framework. International Journal of Industrial Ergonomics 42(6):553–568
Liu R, Koedinger KR (2017) Closing the loop: Automated data-driven cognitive model discoveries lead to improved instruction and learning gains. Journal of Educational Data Mining 9(1):25–41
Lomas D, Patel K, Forlizzi JL, Koedinger KR (2013) Optimizing challenge in an educational game using large-scale design experiments. In: Proceedings SIGCHI Conference on Human Factors in Computing Systems, pp 89–98
Manouselis N, Drachsler H, Vuorikari R, Hummel H, Koper R (2011) Recommender systems in technology enhanced learning. In: Recommender systems

handbook, Springer, pp 387–415
Martin B, Mitrovic A (2003) Its domain modelling: art or science. In: Proceedings of Artificial Intelligence in Education, pp 183–190
McShane BB, Gal D, Gelman A, Robert C, Tackett JL (2019) Abandon statistical significance. The American Statistician 73(sup1):235–245
Mitrovic A (2010) Modeling domains and students with constraint-based modeling. In: Advances in intelligent tutoring systems, Springer, pp 63–80
Nakamura J, Csikszentmihalyi M (2014) The concept of flow. In: Flow and the foundations of positive psychology, Springer, pp 239–263
Pavlik PI, Anderson JR (2008) Using a model to compute the optimal schedule of practice. Journal of Experimental Psychology: Applied 14(2):101
Pavlik Jr PI, Anderson JR (2005) Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. Cognitive Science 29(4):559–586
Pelánek R (2016) Applications of the elo rating system in adaptive educational systems. Computers & Education 98:169–179
Pelánek R (2017) Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. User Modeling and User-Adapted Interaction 27(3):313–350
Pelánek R (2018a) Conceptual issues in mastery criteria: Differentiating uncertainty and degrees of knowledge. In: Proceedings of Artificial Intelligence in Education, Springer, pp 450–461
Pelánek R (2018b) The details matter: methodological nuances in the evaluation of student models. User Modeling and User-Adapted Interaction 28:207–235
Pelánek R (2020a) A classification framework for practice exercises in adaptive learning systems. IEEE Transactions on Learning Technologies
Pelánek R (2020b) Managing items and knowledge components: domain modeling in practice. Educational Technology Research and Development 68(1):529–550
Pelánek R, Řihák J (2018) Analysis and design of mastery learning criteria. New Review of Hypermedia and Multimedia 24(3):133–159
Pelánek R, Papoušek J, Řihák J, Stanislav V, Nižnan J (2017) Elo-based learner modeling for the adaptive practice of facts. User Modeling and User-Adapted Interaction 27(1):89–118
Reddy S, Labutov I, Banerjee S, Joachims T (2016) Unbounded human learning: Optimal scheduling for spaced repetition. In: Proceedings of Knowledge discovery and data mining, pp 1815–1824
Ripley BD (2007) Pattern recognition and neural networks. Cambridge university press
Roediger III HL, Butler AC (2011) The critical role of retrieval practice in long-term retention. Trends in cognitive sciences 15(1):20–27
Rohrer D (2015) Student instruction should be distributed over long time periods. Educational Psychology Review 27(4):635–643
Roll I, Baker RSd, Aleven V, Koedinger KR (2014) On the benefits of seeking (and avoiding) help in online problem-solving environments. Journal of the

Learning Sciences 23(4):537–560
Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1(5):206–215
Sarkar A, Cooper S (2019) Transforming game difficulty curves using function composition. In: Proceedings of CHI Conference on Human Factors in Computing Systems, pp 1–7
Settles B, Meeder B (2016) A trainable spaced repetition model for language learning. In: Proceedings of Annual meeting of the association for computational linguistics (volume 1: long papers), pp 1848–1858
Sottilare RA, Graesser A, Hu X, Holden H (2013) Design recommendations for intelligent tutoring systems: Volume 1-Learner modeling, vol 1. US Army Research Laboratory
Sottilare RA, Graesser AC, Hu X, Olney A, Nye B, Sinatra AM (2016) Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling, vol 4. US Army Research Laboratory
Tatsuoka KK (1983) Rule space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement 20(4):345–354
Taylor K, Rohrer D (2010) The effects of interleaved practice. Applied Cognitive Psychology 24(6):837–848
Vanlehn K (2006) The behavior of tutoring systems. International Journal of Artificial Intelligence in Education 16(3):227–265
VanLehn K (2016) Regulative loops, step loops and task loops. International Journal of Artificial Intelligence in Education 26(1):107–112
Wang HC, Li TY, Chang CY, et al. (2005) A user modeling framework for exploring creative problem-solving ability. In: Proceedings of Artificial Intelligence in Education, pp 941–943