Cognitive Interaction Technologies for Education White Paper

Basil Ell, Philipp Cimiano, Stefan Kopp, and Mario Botsch

Cognitive Interaction Technology, Center of Excellence, Bielefeld University
Bielefeld, Germany

Abstract. The world is changing at an ever growing pace. New technologies and a globalized economy require new knowledge and skills as new individual and societal challenges arise. Learning in school for life becomes increasingly inadequate. Instead, life-long and life-wide learning becomes more and more important. In this white paper, we, that is, the Center of Excellence Cognitive Interaction Technology (CITEC), present our vision of Education 4.0, where cognitive interaction technologies play a key role in addressing some of the key challenges in life-long learning.

1 Introduction

Education, *"the act or process of imparting or acquiring general knowledge, developing the powers of reasoning and judgment"*,[1] plays a key role in our society. Learning contributes to an individual's ability to meet the challenges of life and work and to participate in a competitive economy by means of a satisfying employment situation. From a societal perspective, educated individuals contribute to a society's endurance in international competition and to a society's ability to address the key societal challenges of our time, such as public health, climate change, demographic change, hunger, and unemployment, to name just a few.

Where education used to be mainly a process of preparing oneself or others intellectually for mature life, since skills and knowledge required change quickly and at an unprecedented rate, education nowadays needs to happen continuously, life-long, and life-wide.

Education 4.0 technology, such as digital learning assistants, which we discuss in more detail in this paper, could in the future gain further importance and enable a larger number of users from diverse economic backgrounds and diverse skill levels to acquire knowledge and skills in a personalized manner, in the most effective way, anytime, anywhere, life-long, life-wide, and targeted to the requirements of the job market.

Developing technical systems that are intuitive and easy to operate for human users is the objective of the Cluster of Excellence "Cognitive Interaction Technology" (CITEC) at Bielefeld University. About 250 researchers are working at the Cluster. As part of the Excellence Initiative of the German Research

[1] Source: `http://www.dictionary.com/browse/education` (Acc. November 2016)

Foundation (Deutsche Forschungsgemeinschaft, DFG), CITEC is funded by the state and federal governments.

We are convinced that we at CITEC have a crucial set of skills at our disposal to coin the term and advance the state of the art in the area of Education 4.0 due to our diverse experiences with cognitive technology, user modeling, semantic knowledge representation, natural language processing, dialogue technology, assistive technology, and simulation.

2 What is Education 4.0

Education 4.0 is a relatively new term that only recently began to appear in the scientific literature, is rarely used, and underspecified. Currently, there are two understandings where the first one is inspired by the term "Industry 4.0" and the other one having a blog post [6] by Moravec from February 2008 as a foundation.

Industry 4.0 is characterized by the disruptive change taking place in the manufacturing industry through the pervasive use of information and communication technology. Other fields were affected in a similar way, thus the terms Work 4.0 or Healthcare 4.0 were coined. Industry 1.0 was characterized by mechanization, water power, and steam power. Industry 2.0 was characterized by mass production, electricity, and the assembly line. Industry 3.0 was characterized by computer and automation, and finally, Industry 4.0 is characterized by cyber-physical systems.

Moravec provides a table where he compares Education 1.0 and Education 2.0 and explores what Education 3.0 could mean, along several dimensions such as the role of technology, how meaning is constructed, and how teaching is done. According to his view of Education 3.0, meaning is socially constructed and contextually reinvented, technology is used everywhere, students are taught by teachers, students teach students, students teach teachers, humans teach technology, and technology teaches humans. We omit other dimensions here. Others build a definition of Education 4.0 on top of Education 1.0–3.0 as defined by Moravec. For example, according to Costa et al. [1], in Education 4.0 meaning is societally constructed, contextually reinvented and experimental. Furthermore, Education 1.0–4.0 are characterized as follows: Education 1.0 is characterized by a unidirectional communication between teachers and students, Education 2.0 becomes a creative activity involving open-source applications. In Education 3.0, non-human agents can provide recommendations regarding educational content based on users' preferences or settings. Educational technologies 4.0 will have artificial intelligence in all applications. Some visions incorporate: the blend of human and non-human brain recognition, which will "download" skills and knowledge, total ubiquitous capabilities (knowledge repository of human civilization), or an aware and cognoscenti multidimensional network from which meta-knowledge (complex decision-making) will arise. Harkins [3] also extends Moravec's definitions: memorization (Education 1.0), Internet-enabled learning (Education 2.0), students are empowered to produce, not merely to consume knowledge (Education 3.0), students are empowered to produce innovations, the follow-on substantiations of knowledge production (Education 4.0). This perspective is also followed by Knight [5], where Education 4.0 encourages students to create and produce innovations by networking in learning communities.

In this white paper we do not provide a clear definition of Education 4.0. Instead, we present a set of visions, which are more closely related to the first strand of interpretation, that is, based on the analogy to Industry 4.0

3 Visions of Education 4.0

We illustrate our visions of future learning by providing several illustrative scenarios. The key aspects and challenges are discussed in the subsequent section.

3.1 Blended Learning and Acquisition of Theoretical Skills and Knowledge by Experiments

Martin is on his way to work in a self-driving car. During the self-driving phase he usually has thirty minutes of spare time without being interrupted, e.g., by phone calls. Martin lives in Germany and is an IT consultant in a large international IT consulting company. For his next project, that will start in three weeks, he will be consulting a high-tech agricultural project on vertical farming in Tokyo, Japan. To be prepared for the project as well as for his stay in Japan, he wishes to learn about the crops that will be cultivated, Japanese etiquette, and places of interest and activities in the vicinity.

Regarding the Japanese etiquette, he had heard that it is worth to be prepared for cultural differences. He asks his digital learning assistant about learning units that are relevant in the context of his project. The assistant inquires some details about the project such as the company and the location as well as how much time he has today and in total. He allows the assistant to access the project-related documents. Furthermore, the assistant inquires his learning goals. He specifies that he is interested in the top 10 sights, that it is sufficient to be informed about Japanese etiquette 10 percent above average, and that he needs to learn the basic agricultural facts about the crops.

Martin is presented with a 3D-generated social situation where he has to decide how to react – e.g., whom to greet first when entering a meeting room, or how to criticize a strategy – by selecting among a given set of options. At that time he has little to no understanding of the concept of face,[2] so he selects the wrong option. The scenario continues and the reaction of the people involved are unexpected by Martin. He does not even notice subtle changes in their posture. The mission fails – something had gone wrong. He is now aware that there is something that he does not understand and he needs to investigate what that

[2] According to Huang [4], *Face is a sense of worth that comes from knowing one's status and reflecting concern with the congruency between one's performance or appearance and one's real worth.*

is. He provides the system with some hypotheses what had gone wrong, but all of them did not uncover the problem. Therefore, the system informs him about the concept of face. Then, he is presented with the same mission where he failed again, since he still did not grasp that concept. He is therefore presented a more detailed explanation. This time, he fulfills the mission. To check whether he understood the concept in sufficient depth the system generates further missions where the same concept is relevant and it increases the level of difficulty step by step. Moreover, it introduces other concepts that are relevant to understand the reactions of the people involved. The missions he is presented take place in or around places of interest in Tokyo and the persons introduced to him in the missions are the people he will most likely interact with during his project.

What Martin is not aware of is where the system retrieves the learning material from and how these are tailored towards his specific requirements. Since the system has access to project-related documents, information about key personnel, their roles in the project and their relative position to Martin in the company's hierarchy are automatically extracted from these documents.

The assistant retrieves descriptions of learning units from several market platforms for learning units as well as from repositories of freely available resources. There are resources that explain how to behave when invited by His Majesty the Emperor, but given the information the assistant has about Martin and the project, this resource is ignored. Resources about more likely situations such as *project meeting*, *dinner with colleagues*, and *invited at home by a colleague* are given higher relevance scores.

3.2 Theoretical Skill and Knowledge Acquisition within an Group of Learners

Nora studies Computer Science and attends a course on databases. It is an online course, taken by thousands of students from all over the world. The lecture consists of topic-specific videos, where experts explain relevant concepts, and of exercises where students can train their problem-solving capabilities and evaluate their proficiency. Some students may need to train using more exercises than others. However, the digital learning assistant provided together with the course can generate an large number of exercises according to various sets of skills possessed by a learner and skills to be acquired/trained by the learner. The current chapter of the course is about ontological modeling. Nora does not yet fully understand when something should be modeled as a concept or rather as an individual. In a series of tasks, which allow her to make wrong decisions in a kind of virtual laboratory, at a later point she experiences the drawbacks of her decisions, so that she can learn from experience.

As a participant of the online course, interacting with the digital learning assistant only would constitute a missed opportunity to connect with other participants. The learning assistant brings together participants with similar or complementary skill sets, so that participants can converse about the lecture's topic. Certain exercises need to be worked on in groups and the learning assistant may propose how to form those groups.

Years later, Nora is now working as a software developer, databases still play an important role in her job. She is still connected with her learning group. Some participants have dropped out of that group, since the topic of databases is not relevant for them anymore. For the other participants, however, new requirements in industry led to relevant new developments in the field of databases. The members discuss these developments in the learning group, which turned into a community of practice.

3.3 Basic Math – Theoretical Skill Acquisition

Understanding and calculating with fractions is a typical task taught in basic mathematics education. The mathematical aspects can easily be defined given formal definitions of the operations of addition, subtraction, multiplication, and division, as well as the necessary methods of making the denominators of two fractions equivalent and how to reduce a fraction. However, learning about fractions goes beyond an understanding of the formal definitions or understanding the formal definitions may not even be the main goal. In some cases, an intuition of how to perform basic operations and a feeling about quantities needs to evolve, e.g., what does it mean if the numerator is greater than the denominator or vice versa? What if they are equal? Is $1/3$ greater than $2/4$? What if one of them is not known (e.g., $a/4$)? What happens if a number is multiplied with a positive fraction that is less than one?

Beyond the mathematical aspects, an understanding of why fractions and the operations based on fractions are relevant can be the goal of a learning unit. Why do we need fractions at all? An example could be children sharing a cake. Is a cake that is cut into twelve pieces the same as before it was cut? Wouldn't it in certain situations be fair to create weighted partitions, for example when water is to be shared among animals? Shouldn't the horse receive more water than the dog? All these examples are applications of the concept of fractions. Exercises can even become more complex given questions such as: How to cut a cake into seven equivalent quantities or why can't it be done by subsequent bisection. Here, a proof needs to be provided.

Given a formal description of how to handle fractions, e.g., in the form of exercise templates where numbers need to be inserted, a learning environment can generate an infinite number of exercises and can automatically check whether the student provided the correct answer. However, the task can be modeled with a deeper understanding of the domain and the necessary skills, such as: which are the subproblems that need to be solved by the learner and how to generate exercises where certain subproblems do not occur. For example, adding two fractions that have a common denominator is easier than if they have different denominators. Having two fractions with different denominators where one fraction can be reduced so that the denominators become identical can be taught in an exercise. Having a fine-grained description of the domain enables to create a diverse set of exercises where multiple techniques can be taught and the user can be evaluated against the set of necessary skills.

The learning assistant could employ a set of typical misconceptions. For example, it can be aware of the misconception that two fractions with unequal denominator can be added simply by building the sum of the numerators divided by the sum of the denominators (example taken from Dräger and Müller-Eiselt, [2]). If the user enters a wrong result that can be derived via this wrong method of adding fractions, then the learning assistant can display a message that explains what the learner is doing wrong and how the sum of fractions with different denominators ought to be calculated. Having a cognitive model of the task as well as of the user's capabilities and misconceptions, the learning assistant can create a learning path tailored toward the user.

3.4 Professional Cooking – Manual Skill Acquisition

Professional cooking is an activity that requires a set of logical as well as manual skills. Some of the logical skills are organizational skills, i.e., when to do what. This can be challenging when several meals are prepared at the same time and resources, such as a stove, are shared with others. Another skill deals with the adaptation of plans or recipes. What if an ingredient needs to be replaced due to the fact that it is unavailable or due to the fact that the ingredient is inappropriate given the guest's allergy? Furthermore, given a set of ingredients, how to combine them to create a meal? How to develop a repertoire of tastes that complement each other? How to come up with a food course for a special occasion?

Other skills cover: know how a fried egg looks like just before it is ready (visual sensation), how to beat mayonnaise or how to knead dough (haptic sensation, tactile tasks, motor skill)? How does it sound when the oil in the pan is too hot (auditory sensation)? How does it smell if a certain procedure goes right/wrong (olfactory sensation)? How to know how long to cook a big sweet potato? How to know that a leg of lamb is medium? How does a ripe tomato feel (haptic sensation)? Here, simulation environments or serious games may help to acquire the necessary knowledge and skills. However, only a subset of senses are addressed.

The logical/theoretical tasks regarding planning/adaptation etc. won't be discussed in more detail here, since these skills can be acquired in a similar fashion as discussed in the other visions. The focus here lies in the visual, auditory, haptic, and olfactory sensations and skills.

Technology cannot yet satisfactorily replicate the feel of an (un)ripe tomato or the smell of a ripe mango. However, audio and video can be recorded, e.g., the picture of a ripe tomato or the process of cutting a tomato. And it is not necessary in the context of teaching about professional cooking that a system generates output for all human senses. Instead, having the cooking domain modeled in terms of which tasks are necessary, which skills are involved, what is easy and what is difficult, and what can go wrong, it can recommend learning objectives and provide learning material and exercises. Moreover, it may connect users. Given that there is a user lacking the skill of dissecting a tuna but is mastering the skill of beating Aioli mayonnaise and another user is lacking the skill of beating Aioli mayonnaise and instead is mastering the skill of dissecting a tuna, these two users can be connected so that they can exchange their knowledge and show each other how to perform the task. Moreover, one user could have difficulties performing fine filleting, but another user has recently acquired that skill. One user could then teach the other user. The learning assistant therefore needs to build user models that express which skills a user possesses since when, which skills the user does not possess, and which skills the user needs to acquire according to the user's learning goals. Connecting users is especially important for two reasons: one the one hand it can be very motivating to interact with people sharing the same interests, and on the other hand, certain tasks and skills can be very difficult to represent formally in a machine-understandable manner. Moreover, for certain tasks no learning material might exists but users can still acquire the necessary skills by directly exchanging knowledge with others and acquiring the necessary skills.

Especially interesting in the cooking scenario is the use of augmented reality and object recognition. For example, these two techniques could be applied to teach different methods to cut vegetables, such as the Japanese vegetable cutting techniques Hyoshigiri (cutting in bars) and Kakugiri (cutting in cubes). Here, a cutting grid and other auxiliary means could be superimposed over real vegetable. Via object recognition it can be detected whether the right type of knife is used and how close the shapes of the resulting pieces are compared to the intended shape.

Cooking is selected here as an example since, beyond logical thinking, a set of other skills are necessary. This blend of skills can also be found in other domains, to which a similar learning assistant could be applied, such as, how to repair a car or TV, how to learn to dance Bhangra, how to paint, how to sing, how to play the violin, how to take care of children, how to treat an illness, how to harvest crops, how to deal with snake bites, how to survive in the wild, how to fight, how to meditate, or how to mediate.

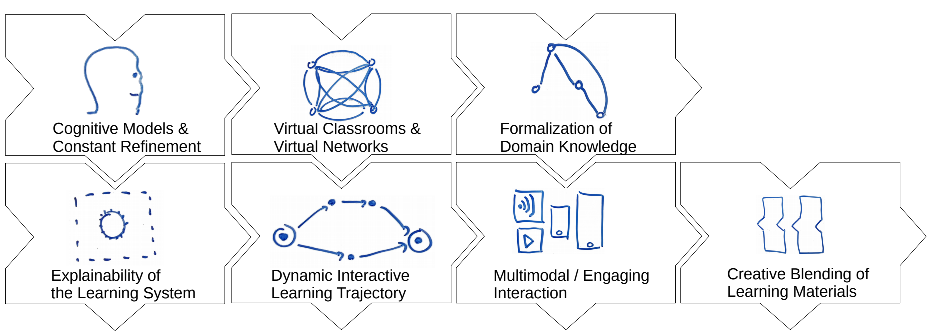

Fig. 1. The core aspects of Education 4.0.

4 Aspects and Challenges of Education 4.0

Figure 1 depicts the seven core aspects and challenges of Education 4.0. Note that the aspects in the upper row have a more static character whereas the aspects in the lower row have a more procedural character. However, this separation is not very strict. The following sections discuss each of the building blocks.

4.1 Cognitive Modeling and Constant Refinement

Cognitive modeling can play a role to model several aspects in a life-long learning scenario. Aspects that can be created by modeling are the *user model*, the *task model*, and the *skill model*. Thereby, the user's current state of knowledge acquisition and skill acquisition, the learning goal, the skill to be acquired, the knowledge to be acquired, outcomes when the skill is correctly acquired, and outcomes when the skill is not correctly acquired can be represented. These cognitive models are not independent and they are refined over time the more a user interacts with technology. By regarding these models, a learning assistant can retrieve learning objects according to the user's current level of knowledge acquisition and skill level and the user's correct assumptions and false beliefs. Beyond retrieval of learning objects, learning objects can be tailored, parameterized, or generated and learning trajectories can be planned or adapted. Depending on knowledge about the user, the system can select learning objects that support the pedagogical strategy that is most effective for the user.

4.2 Virtual Classrooms and Virtual Networks

While learning can happen by interacting with a learning assistant, it can for several reasons for some tasks be crucial that a user interacts with other users. Of course, it can be motivating to interact with others, to share knowledge with others or to just inquire information about a learning unit. Competition with real people can be beneficial for some users. However, for some domains the knowledge may not be formalized sufficiently such that a system can generate exercises or that the necessary skills have been defined by domain experts and learning experts. It may also be the case that the learning domain itself cannot be clearly defined yet, for example, if the domain is constantly evolving, such as in an area of active research.

Groups of learners with a similar skill level, i.e., where users can help each other, can be built. However, the groups should not be isolated. Communication vertically to the skill level and membership in multiple levels should be supported. Ideally, information flow is supported from primary knowledge creation communities, e.g., groups of researchers, toward application-oriented communities and the interested general public.

Knowledge about a domain is subject to change. Therefore, learning infrastructures should support long-term interactions between participants. Users may subscribe to updates, so that they can be informed about new learning materials, new concepts, and new conceptions within their domains of interest. Groups can grow together, collectively moving toward more expertise, by being informed about new knowledge, by working on exercises, and by supporting groups with a lower skill level. Note that this implies that there is no necessary or strict separation between learners and teachers in this skill/knowledge acquisition and distribution ecosystem. Here, learning assistants can accompany users and groups of users in long-term interactions and can enhance the interaction among users by recommending which groups to join, whom to ask, and whom to help.

4.3 Creative Blending of Learning Materials

In the first scenario, Martin has the desire to learn about Japanese etiquette, places of interest, and key personnel at the same time. Besides gathering the required information, which is a challenge already, another challenge is to combine several learning items into combined learning units, or to parameterize learning units with other information. Possibly, the project-related learning goals can be combined with the user's interests, so that the learning tasks can be more engaging. For example: the hard goal is to learn about etiquette. Moreover, the user is interested in food and is curious in leaning about Asian cuisine, which the user either specified explicitly or the system might derive that interest from observing and analyzing his food pictures on social media picture sharing websites. The underlying hypothesis is that this blending is beneficial from a pedagogical perspective: learning becomes more engaging and retaining knowledge becomes more effective.

4.4 Multimodal / Engaging Interaction

Education 4.0 will not be restricted to book and board. Instead, a plurality of media, modalities, and communication channels will co-exist, with the purpose of selecting the best media/modality/channel given a user model, a learning goal, and a skill model. For example, there might be a situation where some content is best explained via video for some users, whereas others prefer obtaining knowledge by learning from examples or by making experiences in a virtual laboratory. For the purpose of obtaining a basic intuition about a mathematical concept, the formal mathematical foundations do not need to be learned in depth. Some users understand a concept by reading about it, others prefer a step-by-step animation. Some users work best in groups and are motivated by the interaction, some users only find time to learn while on the night train, so that they can only asynchronously interact with their co-learners.

4.5 Formalization of Domain Knowledge

A knowledge domain can be modeled in a way such that exercises can be generated automatically according to a user's learning goals, according to the skills already possessed by the user, and according to the user's current misconceptions. This requires a fine-grained model of the knowledge domain, of the tasks a user needs to understand, the skills that are required, and how skills are related to each other. The domain model expresses which tasks are necessary to solve a problem, which sub-tasks they involve, which tasks are more challenging than others, and which exercise, modality, and pedagogical strategy can be used to train or assess a certain skill. It allows the learning assistant to generate more exercises of a certain task, so that the user can train his or her skills until the task is sufficiently acquired. The domain model allows to generate exercises in various forms, depending on the modality most effective for the user. Moreover, skills, tasks, and exercises are related to explanation material, such as textual explanations of how to approach or solve a certain task, to audio content, video content, or interactive simulations.

4.6 Dynamic Interactive Learning Trajectory

An important concept is that the digital learning assistant has the function of a tutor that has an understanding of a user's current skill level (and about the user's misconceptions). Therefore, the learning assistant can plan a learning trajectory specifically tailored toward the user. More exercises of a kind can be generated until a satisfactory skill level is obtained. Learners usually have different skill sets and previous knowledge. Given a common domain of knowledge and skills to be acquired, the learning assistant can guide each user on another learning trajectory through this domain.

Besides planning how to traverse the knowledge domain, temporal constraints may exist and need to be regarded. Martin has three weeks, that means, a certain number of minutes of learning: in the car, in the office, at home, and maybe via other channels, too. Maybe he watches a video that is not too demanding in the evening or he listens to a lecture while in the gym. Or, he might even receive some test questions per email after lunch, or the assistant may propose him to have lunch with his Japanese colleague. The learning assistant needs to store in Martin's user profile his learning habits, that means when he has time, which device he is using when, and what kind of task is best given a device/time/etc – thereby constantly refining the user model and eventually dynamically adapting the learning trajectory. The learning path the system plans needs to take all these factors into consideration. Furthermore, cognitive load needs to be taken into consideration when planning a learning trajectory. Given a goal that is related to a certain skill and given a hierarchical model of skills and sub-skills, a path via skill-related learning resources needs to be found that does not require too many new skills at a time in order not to overwhelm the user, unless the user is most motivated when challenged by complex tasks.

4.7 Explainability of the Learning System

In the scenario about Martin, given that, e.g., a mission failed, or, in the scenario of basic mathematical skill acquisition, if the result is incorrect, the user can ask: *What went wrong?* Since the system has beliefs about what the user understood (since certain missions require the understanding of certain concepts and the proficiency of certain skills and given that the user passed those missions successfully) and what the user probably did not understand, or even the system may have beliefs about how the user misunderstood a concept or situation, the system can generate an answer that does not contain what the user already knows, thus is tailored to the expected level of understanding of the user and therefore can be maximally effective. Besides directly answering the question "*What went wrong*", the system might also engage in a clarifying dialog, by asking: "*What do you think went wrong?*"

Explainability is orthogonal to other aspects, since explainability is enabled by the aspects of cognitive modeling and domain knowledge.

5 Conclusions

In this white paper we presented our vision of Education 4.0 given a diverse set of scenarios that tackle blended learning, the acquisition of theoretical skills and knowledge, the acquisition of manual skills, learning by obtaining experience from experiments, and the support of groups of learners. Crucial aspects are the *cognitive modeling* of users, tasks, and skills, where the user model can constantly be refined while a user interacts with a digital learning assistant. *Virtual classrooms and virtual networks* enable the exchange within an heterogeneous set of users to enable knowledge exchange and long-term interaction and to support the emergence of communities of practice. *Creative blending* of multiple learning tasks may lead to an engaging learning experience and may have a positive effect on the effectiveness of knowledge and skill acquisition. A further aspect of engagement is the aspect of *multimodality*, where a plethora of channels, devices, and pedagogical strategies can be applied for the purpose of supporting learners in the most effective way. The *formalization of domain knowledge* enables exercises to be generated automatically and specifically tailored toward a user and the user's current skill level and taking into account the user's misconceptions. This formal knowledge, together with the cognitive models, also allow the planning of *dynamic interactive learning trajectories* so that for each learner it can be planned how to reach a set of learning goals and how to re-plan, given the measured performance of the user. Finally, we discussed the aspect of *explainability* as an important property of a learning assistant. This property necessarily requires other aspects such as the cognitive models and the formalization of domain knowledge.

At CITEC we are well positioned in several fields crucial to tackle the challenges discussed above, which renders us well equipped to address these challenges and to strongly contribute to the emerging field of Education 4.0. These fields are Semantic Technologies, Social Cognition, Speech Processing and Speech Understanding, Machine Learning, Modeling of Cognitive Capabilities, Social Cognition and Interaction, Knowledge Representation, and Semantic Modeling.

References

1. G. J. M. Costa, N. S. A. Silva, and T. F. R. Fonseca. Moral reasoning in e-learning generations: from 1.0 to 4.0. *ETHICOMP 2013*, pages 110–117, 2013.
2. J. Dräger and R. Müller-Eiselt. *Die digitale Bildungsrevolution.* Deutsche Verlags-Anstalt, 2015.
3. A. M. Harkins. Leapfrog principles and practices: Core components of education 3.0 and 4.0. *Futures Research Quarterly*, 24(1):19–31, 2008.
4. S. Huang. Two studies of prototype semantics: Xiao 'filial piety' and mei mianzi 'loss of face'. *Journal of Chinese linguistics*, pages 55–89, 1987.
5. B. A. Knight. Teachers' use of textbooks in the digital age. *Cogent Education*, 2(1):1015812, 2015.
6. J. Moravec. Moving beyond Education 2.0. `https://educationfutures.com/blog/2008/02/moving-beyond-education-20/`, 2008.