Cognitive Load Levels While Learning with or Without a Pedagogical Agent

Madlen Müller-Wuttke[1(✉)] and Nicholas H. Müller[2]

[1] Technische Universität Chemnitz, Thüringer Weg 11,
09126 Chemnitz, Germany
madlen.mueller-wuttke@phil.tu-chemnitz.de
[2] Socio-Informatics and Societal Aspects of Digitalization,
Faculty of Computer Science and Business Information Systems,
University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany
nicholas.mueller@fhws.de

Abstract. An eye-tracking study examining the benefits of a proactive human-computer-interaction has been conducted. Regarding the beneficial aspects of knowledge transfer, the data shows a clear benefit once the system has been enhanced by the so called electronic educational instance which tracks a user's gaze and thereby infers whether or not someone is actually looking at a screen while the e-learning software is conveying its knowledge. To further show how a physiological input might be used as a form of input, current literature of smoothing pupillary response data is discussed and a preliminary tool is presented. Due to this, pupillary data can be used to indicate cognitive load levels while learning and would therefore allow a proactive system to change the e-learning program accordingly.

Keywords: Pedagogical agent · E-learning · Proactive system · Cognitive load

1 Introduction

Forms of interaction with a learning system are still mostly limited to traditional forms of human computer interaction, like a keyboard and a mouse. Although there are a number of new technologies available to interact with a computer system and they are getting more affordable day by day, equipment like VR goggles, haptic interfaces or gestures in midair are still confined to specific niches. And while there are numerous advances in regards to how a topic is framed or in what way it is presented to the user, the main areas of research activity in electronic learning [1] and the validation of an e-learning application is most often limited to an analysis about the framing of given information, the appearance of the learning system or certain elements inside the environment (e.g. [2–4]).

But according to the media equation theory [11], humans have a tendency to actually use and behave in front of and towards a computer system, like within a human-to-human-interaction. Reeves and Nass tested this assumption by having participants work with a computer and then asked them to rate the system, once on the same system as they have previously worked with and another group on a different

computer. This lead to the result of people being more 'nice' in their answers as long as they had to give their rating on the same as they worked on.

As our previous conceptual research has shown [8], human-computer-interaction should be able to profit from a non-verbal backchannel to the learning environment. This would enable a learning application to take environmental information into account. If a user is distracted or the environment is too noisy, this should be taken into account by the learning system – leading to pausing the conveyance of information or by offering to repeat a section, which might not have been sufficiently understood due to outside disturbances.

With this in mind we developed an electronic educational instance [9, 10] which is working as a plug-in component for already established applications. Due to this, the learning application is capable of checking for the gaze of a user and determine the focus of attention regarding the screen. Furthermore, by checking microphone levels, the application is able to analyze the noise level of the surroundings and decide whether or not it would be enough to just dial up the volume levels of an explanatory video- or audio-stream or if it would be beneficial for the learning success to pause the learning session, until the noise-level decreases.

Therefore, the system possesses an environmental-feedback-channel, leading to the inclusion of non-verbal human-computer-interaction possibilities and thereby leading to an interaction experience which is more closely related to a human-to-human-interaction, as stated in the media equation theory by Reeves and Nass [11]. The system would then be capable to analyze a situation as any real-world teacher would do and react to user specific deviations from the learning session. This is realized by using a common webcam and microphone, as is most of the time already built into current notebooks and tablets. For the presented study, the camera checks for the presence of two eyes and interprets this as being attentive towards the screen while the microphone filters out the audio from the learning application and is focused on background noises. As soon as one of the two criteria for pausing the application occurs, gaze away from the screen or noise levels too high, the content on screen is paused. In order to check for the persona effect, as discussed by Lester et al. [12], we used an SMI eye-tracker to reliably check for the specific areas of interest by the users, specifically our own pedagogical agent. The use of an eye-tracker in order to gain insights into user behavior during a learning session have often been used to validate the position of certain elements of learning applications.

One of the most commonly recorded metrics while recording eye-tracking-data is that of a pupil dilation. As Rosch and Vogel-Walcutt states [5], there appears to be a link between the size of a pupil and the current cognitive load level of a subject. Based on the research of Mathot et al. [6, 7], their algorithm for the extrapolation of cognitive load based pupil dilation was used to re-analyze the data of our study with 139 participants. Building on our earlier publications [8–10] we are going to report the results of the cognitive load levels while learning with our enhanced learning system. In theory, once being able to apply this calculation in real-time, which it is currently not, this could function as a third non-verbal feedback channel, enabling the system to check in real-time whether or not the current form of presentation is suitable for the individual learner. So, in theory, this would allow for the system, for example, to

switch to another, more time consuming and more detail-oriented form of explanatory knowledge conveyance.

2   Eye-Tracking Study Regarding Pedagogical Agents

In order to test for this possibility of a cognitive load feedback channel, we re-examined data from a previous experiment about knowledge of the Dreamweaver software, during which we recorded the eye tracking data to check for the 'did-they-or-did-they-not' focus attention towards the pedagogical agent. In addition, we used this to test for the persona effect [12] by Lester et al., which should lead to a higher learning success as soon as there is a depicted agent visible on the screen. Four groups were tested (see Table 1 and Fig. 2).

Table 1.   Experimental groups

Group 1	Depicted agent, with proactivity
Group 2	Depicted agent, without proactivity
Group 3	Audio-only agent, with proactivity
Group 4	Audio-only agent, without proactivity

The data is based on N = 74 undergraduate students from a study conducted in 2014 [13]. Participants were asked to take part in a second-task wizard-of-oz-experiment during which the proactive system of the EEI (see Fig. 1) would stop the e-learning software whenever the study participants were distracted by the second task. Once the training was completed, the volunteers had to apply their gained knowledge during a practical task session.
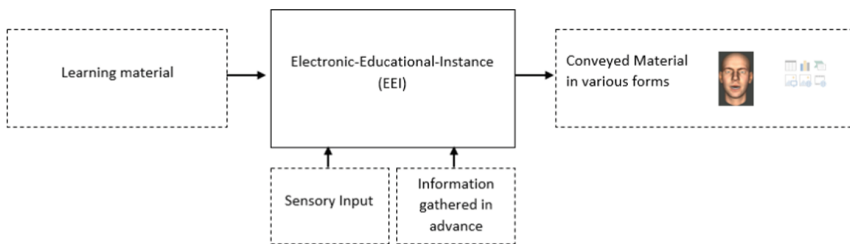

Fig. 1.   The Electronic Educational Instance [9, 10] (EEI)

During the learning part of the study, we recorded the eye tracking data and during the apply knowledge phase of the study, we recorded whether or not the participants chose were able to apply their gained knowledge correctly or incorrectly and we recorded the mouse-track-distance in pixels, number of mouse clicks as well as the time until the subjects ended the experiment. In addition, participants took part in a multiple-
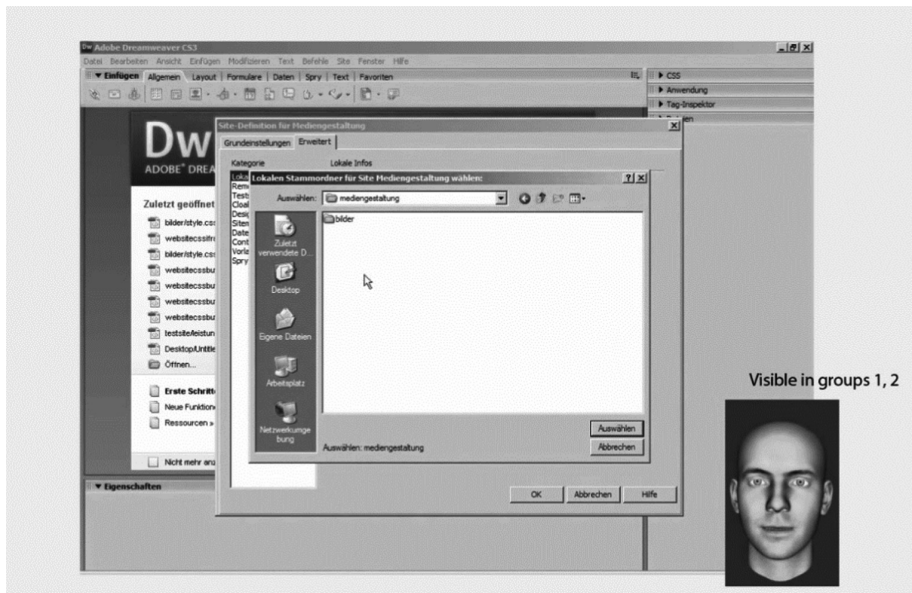
Fig. 2. Learning material and the group variations regarding the pedagogical agent [14]

choice questionnaire regarding their previous knowledge about the Dreamweaver software as well as a learning success test once the experiment was completed.

When checked with a Kruskal-Wallis test, the data shows a significant group difference regarding the AOI (Area-of-Interest) of the pedagogical agent ($H_{Agent}(3) = 41.74$, $p < .001$), but not regarding both the other, much bigger, AOIs regarding the learning material or the white space, meaning everything else on the screen ($H_{LerningMaterial}(3) = 3.46$; $H_{WhiteSpace}(3) = 2.83$). In addition, we checked with a Mann-Whitney test whether or not the two groups with a depicted agent showed a difference regarding their fixations towards the pedagogical agent, but they showed none ($U = 93$, $r = -.31$). Meaning, both groups with a depicted pedagogical agent did actually look at the agent while the groups with a blank space at the position of the agent did not.

Interestingly enough, the results (see Table 2) showed a significant group difference regarding the correct solution of the applied knowledge task ($H_{Solution}(3) = 9.55$, $p < .05$) but not regarding the mouse track or the time until completion ($H_{MousePath}(3) = 1.17$; $H_{Time}(3) = 5.44$; $H_{Clicks}(3) = .77$).

Table 2. Overview regarding mousepath, time, clicks and correct/incorrect solution [13]

	Mousepath	Time	Clicks	Correct	Wrong
Group 1	93917,34	00:07:32	120,43	16	5
Group 2	103404,83	00:08:54	116,7	9	11
Group 3	99358,96	00:06:28	116,7	18	2
Group 4	101946,50	00:07:34	124,8	11	9

It can be deduced that the two groups 1 and 3 (group 1 with the depicted agent and proactivity components activated, group 3 without a depicted agent, audio only, but proactive behavior as well) showed the best learning performance with the proactive behavior. In a comparison of these two groups, group 3, with no depicted agent, but with proactive behavior, appears at first sight to be the best group result. However, a Mann-Whitney test shows no significant differences between the two groups 1 (depicted and proactive) and 3 (audio only and proactive).

Therefore, it can be confirmed that the proactive system component has had an essential influence on the learning success and the ability to apply knowledge in practice. Also, as shown for example by Louwerse et al. [15], it could be confirmed that subjects with a visualized pedagogical agent also fixated them in their field of vision.

The results concerning the mouse distance, time and the number of clicks seem to be interesting in that they do not statistically show a group difference. A purely descriptive analysis of the data (see Table 2) shows that the respective peer groups were faster in terms of proactivity (1, agent with proactivity, and 2, agent without proactivity, as well as 3, audio only with proactivity, and 4, audio only without proactivity), insofar as the system acted proactively. However, this was not done to a statistically significant extent. Accordingly, statistically speaking, all groups had an equal period of time, which is why a similar pixel-track was searched for clues to the solution and a comparable number of clicks were used to bring about a solution. In the case of non-proactive groups, however, this was more likely to be unsuccessful, although the groups without a proactive component also had correct solutions to the practical task.

An in-depth examination of the knowledge difference in the pre- and post-test with regard to the depiction of the agent reveals no significant group difference in the previous knowledge test when distinguishing between group 1 and group 2 ($U_{PreTestKnowledge} = 183$, p = .668) but during the post knowledge test ($U_{PostTestKnowledge} = 127$, p < .05). Meaning, since both groups have a depicted agent, but group 2 does not have the proactive component, this shows that it is not the agent that is responsible for the knowledge gain but rather an information transfer based on the proactive component. Accordingly, there should be no significant differences between the two proactivity groups 1 and 3, confirmed by the following results ($U_{PreTestKnowledge} = 145$, p = .082; $U_{PostTestKnowledge} = 174$, p = .254).

As a result of the study's knowledge test before the experiment, it can be concluded that the basic knowledge of HTML code, its components and the knowledge about the function and operation of the Dreamweaver software, can be considered negligible, since currently web content is mainly used by content management systems or as service software offered by aggregators. Knowing the background of CSS layouts and tags is no longer necessary for the active distribution of content online, which is why it could be expected that this knowledge was not to be found during the pre-experiment knowledge test. Having it established at the end of the experiment speaks to the effectiveness of the original e-learning tool (see Fig. 3).
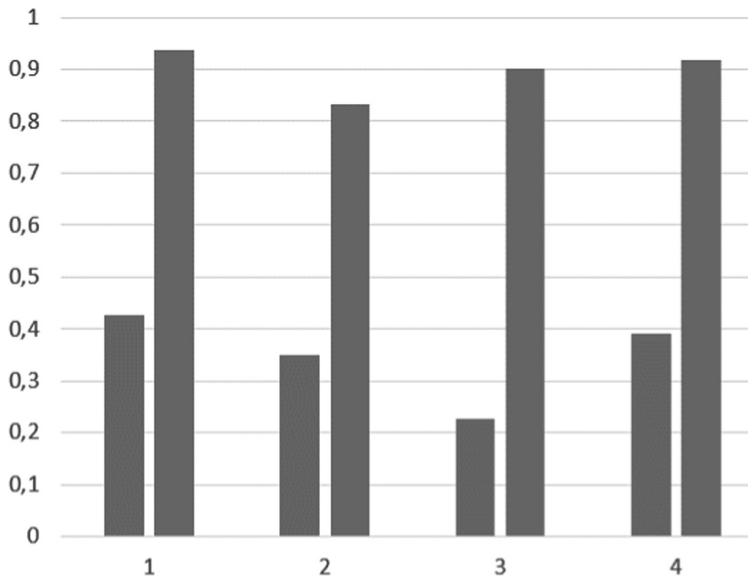
Fig. 3. Knowledge before (left bars) and after the experiment (right bars) of the four groups

3   The Cognitive Load Analysis

In the context of information acquisition, cognitive capacity is of essential importance and also of interest in the context of pedagogical agent research [25]. Starting point in the considerations on the cognitive load theory [23] is that the working memory, which plays an essential role in the acquisition of knowledge, is limited. In order to avoid cognitive overload when learning or linking new information, it is necessary that learning materials are prepared by an appropriate design. Sweller [23] identifies in this context three different forms of cognitive load in the learning process, which can be added to a continuum of the available cognitive capacity resources.

The intrinsic cognitive load is conditioned by the learning material or the complexity/difficulty of the matter to be taught. This can be broken down into suitable information units by segmentation or by a suitable form of instruction design [26]. However, this only influences the amount of the knowledge itself, as each learning material has an intrinsic complexity that can not be further reduced. Insofar as it is accordingly necessary that different areas of knowledge must be understood simultaneously in order to enable further conclusions, this can only be achieved by segmenting the necessary elements of prior knowledge.

The extraneous cognitive load is considered as the part necessary for the understanding of the presentation or design of the learning material. The previous section on intrinsic cognitive load addressed the possibility of segmentation. For example, if such a didactic presentation of the subject matter is missing, or if crucial foundations have been disregarded in the design of learning environments [26], the learner needs a great deal of effort to acquire the knowledge. In this context, it could be assumed that the

presentation of a pedagogical agent and its activities could place an additional burden on the cognitive system. The study by Schroeder et al. [27] suggests, however, that even in this case, the learning support of the agent outweighs.

The germane cognitive load results from the two aforementioned load categories. Sweller [23] describes this as the total available cognitive capacities that can be used to understand the subject matter, for processing, schema formation or automation.

Therefore, the cognitive load is a multi-dimensional construct that represents the activity of the cognitive system in performing a task. With regard to the automatic analysis, this system is differentiated by [16] into three parts: mental load, mental effort and performance. Following [17] mental load is the performance required to solve a given task while mental effort defines the actually invested power. The performance, after completion of a task, describes error rate and speed. This theory is based on the fact that there are only limited resources available for the execution of a task and that these are used up to the degree of difficulty [18]. This utilization of cognitive resources can be measured by various available techniques.

In order to be able to evaluate the pupillary reflex, there are methods to measure the size of the pupil, which are summarized under the term pupillometry. Thereby the size of the pupil is measured by means of imaging techniques. Usually the eye is illuminated by IR light and the movement is recorded by measuring the angular difference between the center of the pupil and the IR reflection spot on the eye. So, by measuring the eye-movements, the pupil size is a by-product.

For this reason, a contact-free gaze-measuring device is used in the present experimental setup, which uses infrared lamps to image a so-called Purkinje reflection [19] on the cornea. This measurement method is suitable for media psychology research practice through the non-contact configuration [20].

Therefore, we used an infrared-based eye movement apparatus from SMI for the implementation of the study [21]. The infrared illumination and the camera for eye movement detection are mounted by a bracket below a 21" wide-screen display and we measured the infrared reflection on the cornea and the distance to the pupil center (see Fig. 4). Although regarding this specific article, we only focused on the dilation measurements of the pupil during the learning phase of the experiment, regardless of the fixation or saccade of the eye itself.

Possibly the main challenge by analyzing pupillary reactions in order to infer cognitive load, is the problem of rapid pupillary dilations, blinks and other artifacts from differences in lighting. Therefore, pupillary response data has to be smoothed over time in order to be able to find relevant differences and to be able to identify actual increases or decreases in cognitive load. We chose the procedure described in [6, 7] in order to smoothen the data-set and implemented them into a prototypical software. By using maximum and minimum pupil-dilation diameters, we normalized the single-user data into a spectrum of 0 to 1. Future iterations of the software will also include a z-transformation to allow for a comparison of different user data.
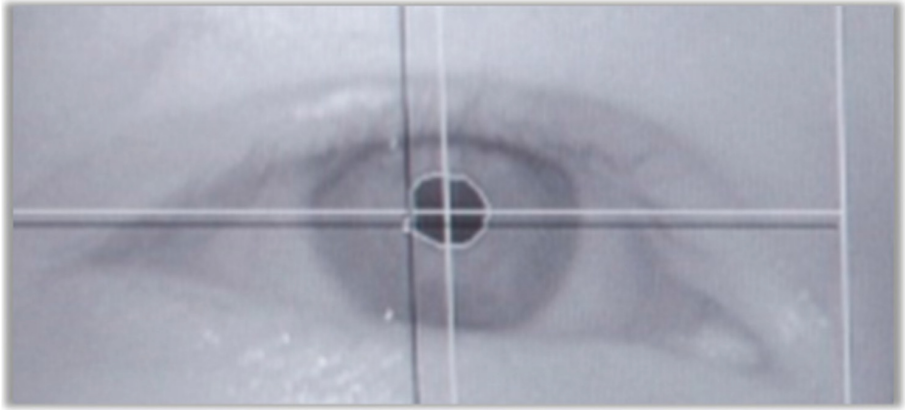
Fig. 4. IR reflection on an eye with the measured pupil-diameter

4 Cognitive Load Data

Using the aforementioned method, we smoothed the pupillary responses of our subjects. Looking at the data, there are still numerous variances visible in the data stream (see Fig. 5).
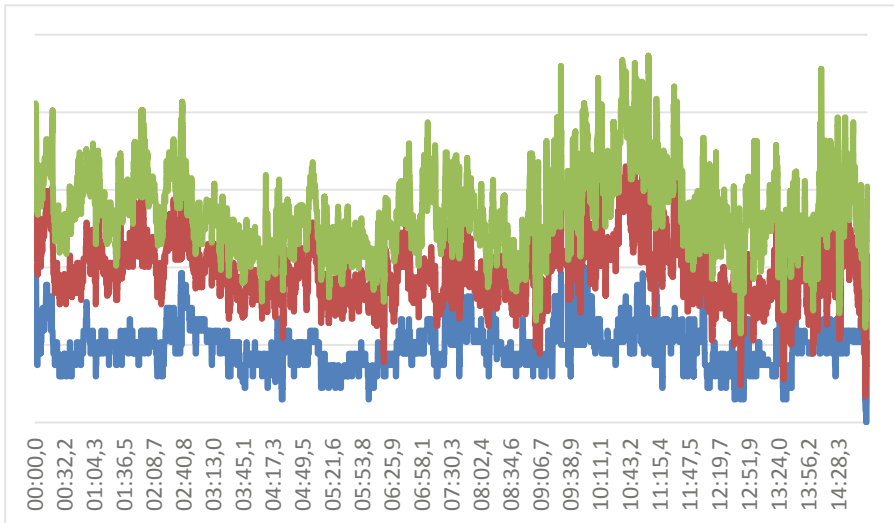

Fig. 5. Three random participants pupil dilations after smoothing of the data

Interestingly enough, the first two streams seem to be running pretty much in unison, while the third shows variations. Nevertheless, since the lighting was kept constant during the trial, the smoothed data is free of artifacts like blinks or lost

pupillary tracking, Therefore, we are now able to gain an insight into the cognitive load while users are interacting with the learning system.

Our next steps will be to overlay the data stream onto the recorded videos of the experiment and to check where during the knowledge conveyance peaks in the pupillary response are visible. In addition, we are going to collate the area of interests (pedagogical agent, learning material and white-space) with the cognitive load data in order to get a better understanding of how different aspects of the learning environment might influence the cognitive load in general. Maybe this could lead to a better understanding of the relationship between intrinsic, extraneous and germane cognitive load [22–24].

The proactive EEI (see Fig. 1) could now use this data in real-time to analyze the cognitive load levels and change the mode of knowledge presentation, if and when it should be necessary. For example, when looking at the data, there seems to be a very low level of pupillary dilation at around the 4 min mark of the e-learning program. While on the other hand there appears to be a very demanding task at around the 10 min mark.

5   Conclusion

We presented a study utilizing eye-tracking data while our participants learned with a tried and tested e-learning software. Our goal was to enhance the knowledge conveyance by adding an Electronic Educational Instance (EEI) which enables the traditional e-learning software to behave proactively to changes of the user's behavior in front of the screen or the environmental surroundings. While cognitive load was never the focus of the study itself, we utilized the eye tracking data regarding the recorded pupillary responses during the experiment. Following the state of the art, we implemented an algorithm to smooth the data stream which effectively eliminated noise, artifacts and especially blinks from the recording.

As the study was able to show, a proactive e-learning system is capable of ensuring a successful knowledge transfer. But while this establishes a reaction by the e-learning software to stop and pause while either attendance is not focused on the screen or noise levels in the surrounding environment were too loud, the cognitive load data allows for much more varied actions while learning.

Future iterations of the proactive e-learning might now benefit from an implemented real-time cognitive load capable EEI by:

- Suggesting a pause by the learner after a phase of high cognitive load
- Choosing another form of knowledge conveyance in relation to the load level, e.g. a schematic picture for beginners while a textual description could be sufficient for experts
- Explanation of a certain topic with multiple examples
- Adapting the learning-success-test to the obviously harder to learn topics.

The next iteration of our cognitive load module for the EEI will be tested regarding these and other aspects of possible beneficial actions while learning in conjunction with the cognitive load level.

References

1. Heidig, S., Clarebout, G.: Do pedagogical agents make a difference to student motivation and learning? Educ. Res. Rev. **6**, 27–54 (2011)
2. Lusk, M.M., Atkinson, R.K.: Animated pedagogical agents: does their degree of embodiment impact learning from static or animated worked examples? Appl. Cogn. Psychol. **21**, 747–764 (2007)
3. Wang, N., Johnson, W.L., Mayer, R.E., Risso, P., Shaw, E., Collins, H.: The politeness effect: pedagogical agents and learning outcomes. Int. J. Hum.-Comput. Stud. **66**, 98–112 (2008)
4. Veletsianos, G.: How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences. Comput. Hum. Behav. **28**, 275–283 (2012)
5. Rosch, J.L., Vogel-Walcutt, J.J.: A review of eye-tracking applications as tools for training. Cogn. Technol. Work **15**(3), 313–327 (2012)
6. Mathot, S., et al.: Safe and sensible preprocessing and baseline correction of pupil-size data. Behav. Res. Methods **50**(1), 94–106 (2018). https://doi.org/10.3758/s13428-017-1007-2
7. Mathot, S.: A simple way to reconstruct pupil size during eye blinks (2013). https://doi.org/10.6084/m9.figshare.688001.v1
8. Wuttke, M., Heidt, M., Rosenthal, P., Ohler, P., Wuttke, M., Müller, N.H.: Proactive functions of a pedagogical agent – steps for implementing a social catalyst function. In: Zaphiris, P., Ioannou, A. (eds.) LCT 2016. LNCS, vol. 9753, pp. 573–580. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39483-1_52
9. Wuttke, M.: Pro-active pedagogical agents. In: Fakultät für Informatik (ed.) Proceedings of International Summer Workshop Computer Science, pp. 59–62, July 2013
10. Wuttke, M., Heidt, M.: Beyond presentation - employing proactive intelligent agents as social catalysts. In: Kurosu, M., Ioannou, A. (eds.) HCI 2014. LNCS, vol. 8511, pp. 182–190. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-07230-2_18
11. Reeves, B., Nass, C.: The Media Equation. How People Treat Computers, Televisions, and New Media Like Real People and Places. Cambridge University Press, New York (1996)
12. Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhogal, R.S.: The persona effect: affective impact of animated pedagogical agents. In: Pemberton, S. (ed.) Human Factors in Computing Systems: CHI 1997 Conference Proceedings, pp. 359–366. ACM Press, New York (1997)
13. Wuttke, M.: Proaktive Agenten im Lernkontext. Die Auswirkungen neuer Inputkanäle in der lernstoffvermittelnden Mensch-Computer-Interaktion. Dissertationsschrift. Universitätsbibliothek Technische Universität Chemnitz, Chemnitz (2018)
14. Wuttke, M., Völkel, S., Ohler, P., Müller, N.H.: Analytical steps for the validation of a natural user interface. In: Zaphiris, P., Ioannou, A. (eds.) LCT 2017. LNCS, vol. 10295, pp. 55–63. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-58509-3_6
15. Louwerse, M.M., Graesser, A.C., McNamara, D.S., Lu, S.: Embodied conversational agents as conversational partners. Appl. Cogn. Psychol. **23**(9), 1244–1255 (2008)
16. Yeh, Y.Y., Wickens, C.D.: Dissociation of performance and subjective measures of workload. Hum. Factors: J. Hum. Factors Ergon. Soc. **30**(1), 111–120 (1988)
17. Paas, F.G.W.C., Merrienboer, V., Jeroen, J.G.: Instructional control of cognitive load in the training of complex cognitive tasks. Educ. Psychol. Rev. **6**(4), 351–371 (1994)
18. Just, M.A., Carpenter, P.A., Miyake, A.: Neuroindices of cognitive workload: neuroimaging, pupillometric and event-related potential studies of brain work. Theor. Issues Ergon. Sci. **4**(1-2), 56–88 (2003)

19. Duchowski, A.T.: Eye Tracking Methodology Theory and Practice. Springer, London (2007). https://doi.org/10.1007/978-1-84628-609-4
20. Bente, G.: Erfassung und Analyse des Blickverhaltens. In: Mangold, R., Vorderer, P., Bente, G. (eds.) Lehrbuch der Medienpsychologie, pp. 297–324. Hogrefe-Verlag, Göttingen (2004)
21. SMI - SensoMotoric Instruments: SMI - SensoMotoric Instruments (2008). RED/RED250: http://www.smivision.com/en/gaze-eye-tracking-systems/products/red-red250.html. Accessed 01 May 2017
22. Sweller, J., Chandler, P.: Why some material is difficult to learn. Cogn. Instr. **12**(3), 185–233 (1994)
23. Sweller, J.: Cognitive load theory, learning difficulty, and instructional design. Learn. Instr. **4**(4), 295–312 (1994)
24. Sweller, J.: Element interactivity and intrinsic, extraneous, and germane cognitive load. Educ. Psychol. Rev. **22**(2), 123–138 (2010). https://doi.org/10.1007/s10648-010-9128-5
25. Schroeder, N.L.: The influence of a pedagogical agent on learners cognitive load. Educ. Technol. Soc. **20**(4), 138–147 (2017)
26. Niegemann, H.M., Domagk, S., Hessel, S., Hein, A., Hupfer, M., Zobel, A.: Kompendium multimediales Lernen. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-37226-4
27. Schroeder, N.L., Adesope, O.O., Gilbert, R.B.: How effective are pedagogical agents for learning? A meta-analytic review. J. Educ. Comput. Res. **49**(1), 1–39 (2013)