Computers in Human Behavior 68 (2017) 170–179

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures – An eye-tracking study

Marta Koć-Januchta [a, *], Tim Höffler [a], Gun-Brit Thoma [a], Helmut Prechtl [b], Detlev Leutner [c]

[a] IPN – Leibniz-Institute for Science and Mathematics Education at Kiel University, Olshausenstr. 62, D-24118, Kiel, Germany
[b] University of Potsdam, Institute of Biochemistry and Biology, Didactic of Biology, Karl-Liebknecht-Str. 24-25, House 29, 14476, Potsdam-Golm, Germany
[c] Duisburg-Essen University, Faculty of Education Sciences, Instructional Psychology, P.O. Box, 45117, Essen, Germany

ARTICLE INFO

Article history:
Received 14 March 2016
Received in revised form
31 October 2016
Accepted 17 November 2016

Keywords:
Cognitive style
Verbalizer
Visualizer
Eye-tracking
Multimedia learning

ABSTRACT

This study was conducted in order to examine the differences between visualizers and verbalizers in the way they gaze at pictures and texts while learning. Using a collection of questionnaires, college students were classified according to their visual or verbal cognitive style and were asked to learn about two different, in terms of subject and type of knowledge, topics by means of text-picture combinations. Eye-tracking was used to investigate their gaze behavior. The results show that visualizers spent significantly more time inspecting pictures than verbalizers, while verbalizers spent more time inspecting texts. Results also suggest that both visualizers' and verbalizers' way of learning is active but mostly within areas providing the source of information in line with their cognitive style (pictures or text). Verbalizers tended to enter non-informative, irrelevant areas of pictures sooner than visualizers. The comparison of learning outcomes showed that the group of visualizers achieved better results than the group of verbalizers on a comprehension test.

Issues of cognitive style and learning preferences have been an underlying topic of educational and psychological discussions for years (e.g., Mayer & Massa, 2003; Riding, 1997; Witkin, 1973). Sometimes the whole concept is disputed (e.g., Kirschner & van Merriënboer, 2013), sometimes endorsed (e.g., Cassidy, 2004). Regarding visual-verbal cognitive style and its influence on learning from text-picture combinations, relatively few studies have been conducted (e.g., Höffler, Prechtl, & Nerdel, 2010; Riding & Douglas, 1993). There are even fewer studies which try to examine actual differences between visualizers and verbalizers via a direct observational method like, for example, eye-tracking (e.g., Mehigan, Barry, Kehoe, & Pitt, 2011; Tsianos, Germanakos, Lekkas, Mourlas, & Samaras, 2009). Our study is therefore an attempt to directly examine verbal and visual learners' eye-movements in the context of multimedia learning. Some eye-tracking studies already indicated that visualizers and verbalizers might differ in the way they view pictorial and textual stimuli (Mehigan et al., 2011; Tsianos et al., 2009). Thus, when learning with texts and pictures, learners' visualizer-verbalizer cognitive style might have a direct influence on learning behavior and preferences. Such a finding would help to provide evidence for the existence or non-existence of different cognitive styles and their influence on learning behavior and, furthermore, learning outcome.

1. Theoretical background

1.1. Cognitive style, learning style, or learning preferences?

According to Messick (1984), cognitive style can be defined as an individual difference in the way of organizing and processing information. Sternberg and Grigorenko (1997) described cognitive style as a platform placed between cognition and personality. Often, studies on cognitive style focus on the visualizer-verbalizer dimension, which originally derives from dual-coding theory (Paivio, 1986). According to this theory, incoming information is processed and mentally represented in two ways: verbally and visually. Hence connecting these two mental representations should improve learning outcomes (e.g., Mayer, 2014). Although there is much evidence that some people tend to think in words

* Corresponding author.
E-mail addresses: marta.koc-januchta@uni-due.de (M. Koć-Januchta), hoeffler@ipn.uni-kiel.de (T. Höffler), thoma@ipn.uni-kiel.de (G.-B. Thoma), prechtl@uni-potsdam.de (H. Prechtl), detlev.leutner@uni-due.de (D. Leutner).

http://dx.doi.org/10.1016/j.chb.2016.11.028

and others in pictures (e.g., Mayer & Massa, 2003), there is some controversy as to the impact of this distinction on learning behavior and learning outcome (Kirschner & van Merriënboer, 2013; Massa & Mayer, 2006).

Furthermore, there is great inconsistency in the literature on how to refer to the distinction of visualizers and verbalizers: Some researchers refer to the term cognitive style (e.g., Richardson, 1977), others to learning style (e.g., Kirby, Moore, & Schofield, 1988), or learning preferences (e.g., Plass, Chun, Mayer, & Leutner, 1998). As a result of a factor analysis, Mayer and Massa (2003) identified cognitive style, learning preferences, and spatial ability as three different factors. They distinguished between these three constructs, defining spatial ability as a specific type of cognitive ability, visualizer-verbalizer cognitive style as thinking in pictures or words, and learning preferences as preferences in choosing graphics or text in instructional materials. Based on this distinction, the current study's focus is on cognitive style. We focus on differences between learners who think either more in pictures (visualizers) or in words (verbalizers). Learning preferences, as well as the correlated construct learning style, we understand as a predilection for specific kinds of learning materials (verbal, visual), that can be, but not necessarily has to be a consequence of cognitive style.

Research results are also inconsistent in terms of the structure of the visualizer-verbalizer distinction. Some studies describe this distinction as a one-scale dimension, which two endings correspond to either verbal or visual cognitive style (Mayer & Massa, 2003), others as two different scales (e.g., Paivio & Harshman, 1983). Kozhevnikov, Kosslyn, and Shephard (2005) even subdivided the visual scale into two subscales: Object and spatial. Object visualizers score poorly on spatial imagery tasks, whereas spatial visualizers score highly. The authors reported that many scientists and engineers seem to be spatial visualizers, while visual artists are usually rather be categorized as object visualizers. As the question on the number of scales does not seem to be fully answered yet, we used a large number of different established scales in our study to be able to satisfyingly characterize visualizers and verbalizers. Furthermore, we studied the learning behavior of visualizers and verbalizers in learning tasks which consist of visual (that is, pictorial) and verbal representations.

1.2. Learning with text and pictures

Many studies (Clark & Paivio, 1991; Mayer, 2014; Wittrock, 1989) show that a combination of text and pictures supports learning and deepens understanding and problem-solving processes. For example, in a study conducted by Plass et al. (1998) on visualizer and verbalizer learning preferences, a combination of text and pictures or text and animations led to better learning outcomes than text alone. However, simply combining text and pictures does not always lead to improvements of learning results. The effectiveness of the combination is highly dependent on such aspects as the form of visualization, the type of learning task, the number of referential connections between text and pictures, and personal characteristics of the learner (e.g., Mayer, 2014; Schnotz & Bannert, 2003). Thus, learning achievements differ with respect to individual differences, such as, for example, prior knowledge (e.g., Kalyuga, 2007), spatial ability (e.g., Hegarty, 2005; Höffler & Leutner, 2011; Höffler, 2010), or cognitive style (Höffler et al., 2010).

According to the cognitive theory of multimedia learning individuals process information using two channels: verbal for verbal or auditory representations and visual for visual or pictorial representations (Mayer & Moreno, 2003; Paivio, 1986). Verbal and visual processing is also reflected in the structure of working memory postulated by Baddeley (1998). The capacities of visual and verbal components of working memory (phonological loop and visuospatial sketchpad) are limited (Baddeley, 1998; Chandler & Sweller, 1991), differ strongly depending on individual differences such as intelligence (Baddeley, 2003), and are deeply connected with cognitive load experienced by an individual (cognitive load theory; Sweller, 1994). The more difficult the learning material, the higher the perception of intrinsic load (Plass, Moreno, & Brünken, 2010). Some studies show that working memory capacity and cognitive style (in this case, so called field dependence/independence cognitive style) are correlated (Mousavi, Radmehr, & Alamolhodaei, 2012). Referring to these findings, we make assumptions regarding the way in which visualizers and verbalizers might process information in multimedia learning differently. We assume that limited capacities of working memory's components and individual differences regarding cognitive style can result in favoring either the verbal or visual channel while processing information in multimedia learning (Mayer & Massa, 2003).

Visualizer-verbalizer cognitive style seems to have an impact on the learning process. Visualizers achieve better when learning from pictures and text and profit more from pictorial information, while verbalizers rely more on text (e.g., Plass et al., 1998). Additionally, Riding and Douglas (1993) showed that text-picture combinations are more beneficial to visualizers, whereas conditions providing only textual information result in better results for verbalizers. These findings can support our assumptions and also suggest that visualizers might be better in integrating information represented in both channels described in the cognitive theory of multimedia learning. Moreover, Plass et al. (1998) showed that the absence of the preferred mode of information presentation (e.g., pictorial for visualizers) resulted in poorer learning. On the other hand, Massa and Mayer (2006) could not replicate such an effect. The discrepancy between these findings might be a result of differences in defining visualizers and verbalizers in both studies, though. Massa and Mayer measured visual-verbal cognitive style as well as learning preference, while Plass et al. concentrated on learning preferences.

The inconsistencies of research results regarding advantages of instructional text and pictures for the learning of visualizers and verbalizers – and the predicted differences in processing information and learning outcomes between these groups – encouraged us to examine how visualizers and verbalizers learn from two different, in terms of topic and type of knowledge, combinations of pictures and texts.

1.3. Eye-tracking and learning

Eye-Tracking research revealed that people differ in their patterns of reading a text. Generally, the most effective strategy is to pay special attention to topic sentences and topic-relevant information in the text (Hyönä, Lorch, & Kaakinen, 2002). While dealing with stimuli containing text and pictures, research showed that learning is heavily driven by text (Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010), and that learners tend to spend more time looking at the text than at the pictures (Rayner, Rotello, Stewart, Keir, & Duffy, 2001). However, best learning outcomes can generally be achieved when information from pictures and texts is integrated.

The way of looking at a stimulus depends on its construction. Some studies showed, in line with the *spatial contiguity principle* (Mayer, 2014) that shorter physical distance between textual and pictorial information facilitates the integration of information from these two sources by finding correspondences between them (e.g., Holsanova, Holmberg, & Holmqvist, 2008). Especially a serial layout of the information material, which organizes the material sequentially, enhances the integration (Holsanova et al., 2008). The presence of more abstract illustrations, like decontextualized charts or diagrams, rather than concrete, more realistic ones in the learning material, is less demanding for working memory and therefore helps to process the text more efficiently and also promotes integration of information (Mason, Pluchino, Tornatora, & Ariasi, 2013). On the other hand, when readers interpret pictures and text as separate and self-contained, they tend to choose only one of them and ignore the other one (Holsanova et al., 2008; Sweller, van Merriënboer, & Paas, 1998).

Previous studies showed also that the way of gazing at the presented stimuli is strongly related to individual differences such as prior knowledge and intelligence. For example, Hannus and Hyönä (1999) showed that high-ability students spent more time gazing at relevant areas of stimuli than did low-ability students. They also returned to relevant areas of stimuli and switched between relevant parts of text and pictures more often than low-ability students. The latter spent more time gazing at irrelevant, blank spaces between and around texts and pictures (Hannus & Hyönä, 1999). The same effect was found in a study conducted by Jarodzka, Scheiter, Gerjets, and Van Gog (2010) in which experts and novices described fish' locomotion. The former spent more time gazing at important parts of presented videos than the latter.

Several studies also suggested that cognitive style can have an influence on the way of looking at the stimuli. Mehigan et al. (2011) claim that it is possible to use eye-tracking technology for identifying visual-verbal learners. In their study, they referred to the visual-verbal dimension of the Felder-Silverman Learner Style Model (FSLSM), using for participants' selection the Felder-Solomon Index of Learning Styles (FSILS) questionnaire. Visual learners and verbal learners performed different patterns of gazing at pictorial and textual information while undertaking an e-learning task: Visual learners outperformed verbal learners in focusing on pictorial learning objects, while verbal learners spent more time on textual content than visual learners. The stronger the visual style of learning was, the more time the learners spent on pictorial content. Similar results were shown by Tsianos et al. (2009), who found that visualizers (or imagers, as the study referred to the verbal-imager axis of Riding and Cheema's Cognitive Style Analysis; Riding & Cheema, 1991) concentrated mostly on pictorial content, verbalizers on text, while intermediates were placed in between.

1.4. Objectives of the study

All in all, there is strong indication that the visualizer-verbalizer cognitive style has an influence on learning behavior. The question remains, however, if and how exactly verbal and visual learners differ in their way of learning from texts and pictures, and if these differences result in different learning outcomes. The main objective of the present study was therefore to further examine the differences between visualizers and verbalizers regarding their gaze behavior and their learning outcome when learning with text-picture combinations.

Our hypotheses are as follows:

(1) In line with findings of Mehigan et al. (2011) and Tsianos et al. (2009), visualizers will generally spend more time focusing on pictures than verbalizers, and verbalizers will generally spend more time focusing on texts than visualizers.

(2) Assuming that visualizers are – in a way – experts in using pictures, and referring to studies conducted by Hannus and Hyönä (1999) and Jarodzka et al. (2010), visualizers, similar to high-ability students, will be better at identifying relevant areas in pictures than verbalizers, which means that visualizers will enter relevant areas of pictures sooner than verbalizers. Verbalizers, similar to low-ability students, will enter the irrelevant areas of pictures sooner than visualizers.

(3) Assuming that both groups will use the best integrative way of learning (Mason, Tornatora, & Pluchino, 2013), but *within* those sources of information that correspond to their cognitive style, visualizers will shift their point of focus across pictures more frequently than verbalizers, and verbalizers will shift their point of focus across texts more frequently than visualizers.

(4) Visualizers profit from text-picture combinations more than verbalizers (Höffler et al., 2010; Riding & Douglas, 1993), and might be better in integrating information represented in both the verbal and visual channel (as described in the cognitive theory of multimedia learning), so they will achieve better learning outcomes on comprehension scales than verbalizers.

2. Method

2.1. Participants

University students of all majors between 20 and 29 years of age were invited to participate in the study – with the exception of physics, biology and psychology students, because of their too broad pre-knowledge on the presented topics of learning. About 90 prospective candidates were pre-classified as either verbalizers or visualizers on the basis of a telephone interview consisting of 14 yes/no questions based on the Verbal-Visual Learning Style Rating questionnaire (VVLSR; Mayer & Massa, 2003), the Individual Differences Questionnaire (Paivio & Harshman, 1983), and the Santa Barbara Learning Style Questionnaire (SBCSQ; Mayer & Massa, 2003). Those individuals whose answers to the interview questionnaire did not allow us to clearly distinguish them as either visualizer or verbalizer were not invited to the study. 48 students with relatively clear visual or verbal orientations were selected for further individual testing. From these 48 participants, we had to exclude 16 participants from further analyses because of calibration problems and/or deficient quality of the eye-tracking data. That left us with 32 participants (68.8% female, age $M = 24.63$; $SD = 2.31$ years).

2.2. Instruments and procedure

First, the participants filled in a set of questionnaires that included demographic questions and six questionnaires regarding the visualizer-verbalizer cognitive style the Verbal-Visual Learning Style Rating, VVLSR, Mayer & Massa, 2003, one single item; the Individual Differences Questionnaire, visual scale, Paivio & Harshman, 1983, $\alpha = 0.93$; the Santa Barbara Learning Style Questionnaire, SBCSQ, Mayer & Massa, 2003, $\alpha = 0.92$; the Vividness of Visual Imagery Questionnaire, VVIQ, Marks, 1973, $\alpha = 0.94$; the Verbalizer-Visualizer Questionnaire, VVQ, Richardson, 1977, $\alpha = 0.77$; and the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ), shortened version (Blazhenkova & Kozhevnikov, 2009), with three scales spatial, $\alpha = 0.86$, verbal, $\alpha = 0.79$, and object, $\alpha = 0.93$. This first part of the study lasted for about 25 min.

Second, participants answered questions measuring their prior knowledge of two topics to be presented, which were functioning of a toilet cistern and learned helplessness. These two contrasting topics were chosen to identify gaze patterns for different types of knowledge. The toilet cistern topic is an example of knowledge regarding functioning of mechanical systems. The learned helplessness topic illustrated the Seligman and Maier experiment on dogs (1967) and is an example of conceptual knowledge. For both topics, we expected participants to have low prior knowledge. To assess their prior knowledge we asked participants to answer open questions regarding each of these two topics. Interrater-agreements were 91.5% (4 items) for the toilet cistern set and 97.3% (4 items) for the learned helplessness set.

Third, for each of the two topics, participants were shown, on a 22 inch computer screen, a set of learning materials being composed of pictures with accompanying texts. While studying the learning materials, eye movements were measured with a SMI RED 120 Hz Eye Tracking system, offering the possibility of free head movements (40 cm $\times$ 20 cm at 70 cm distance), accuracy of 0.4°, spatial resolution (RMS) of 0.03°, and sampling rate of 60 Hz and 120 Hz.

Fourth, a posttest including 12 yes/no and 2 open questions was administered. An example of a yes/no question was: “The first dog continued to stay helpless because it was conditioned to bear the electrical hits. True or false?”. An example of an open question was: “Please explain the meaning of the term ‘learned helplessness’!”.

Reliabilities (Cronbach's alpha) for the yes/no questions were $\alpha = 0.75$ (6 items) for the toilet cistern set and $\alpha = 0.60$ (6 items) for the learned helplessness set. Interrater-agreements for open questions were 89.4% (7 items) for the toilet cistern set and 83.5% (6 items) for the learned helplessness set. The competent judges rated in open questions the level of comprehension, while the yes/no questions indicated the level of retention.

The whole procedure lasted for about one and a half hours. In order to assure high quality of the data, the participants were asked to watch the stimuli the whole time with as little movements of their head as possible. Moreover, participants were informed that the examination will last longer than it actually did to avoid time stress. All examinations were conducted during the day between 10 a.m. and 6 p.m.

2.3. Learning materials and areas of interest (AOIs)

While the learning material was self-developed, the toilet cistern topic was inspired by the work of Hegarty, Kriz, and Cate (2003). We decided to present each topic to be learned in a series of three pictures and three boxes of related text (Fig. 1).

Both sorts of information were placed near to each other (cf., Holsanova et al., 2008), and the pictures were designed in a rather abstract manner (Mason, Pluchino et al., 2013; Mason, Tornatora et al., 2013) in order to facilitate the integration of knowledge. Both texts and pictures contained comparable information (i.e., were self-contained). The equivalence of text and picture contents was checked beforehand with a small sample of seven participants who were shown pictures or texts separately and asked to explain what they had learned from them. Furthermore, each set was designed in a comparable way (pictures above, texts below). The sets had different levels of difficulty. The learned-helplessness topic was consentaneously ranked as easier than the toilet-cistern set by seven participants of the pilot study. Therefore, the toilet-cistern set was presented for 2.5 min, and the much easier to understand learned-helplessness set was presented for only 1.5 min.

In order to get the most accurate information about the gaze behaviors regarding specific parts of the display, each of the two sets of text-picture combinations was analyzed with respect to *areas of interest* (AOIs), that is, regions in the stimuli we were especially interested in. At first, for each set, we created AOIs representing texts and pictures, that is, three AOIs for the three texts, three AOIs for the three pictures, and six to eight AOIs for empty space around texts and pictures. For more detailed analyses, we created AOIs inside pictures, separating regions containing relevant areas from irrelevant ones. The relevant areas were defined with help of three experienced judges as those parts of pictures which were essential for understanding the depicted information and could not be omitted in order to understand it. The irrelevant areas were understood as those parts of pictures that provided no such information such as decorations or trimmings. For the textual parts, we did not define specific AOIs, as the eye-tracker was unfortunately not precise enough for such a detailed analysis within texts.

In analyzing the eye-tracking data, we focused on basic AOI events (according to Holmqvist et al., 2011), that means on such parameters as *entry time* (duration from start of the trial to the first hit of the AOI in ms), *dwell time* (sum of durations from all fixations and saccades that hit the AOI in ms), and *transitions* (movements from one AOI to another; see Holmqvist et al., 2011).

3. Results

As expected, participants, both verbalizers and visualizers, showed low prior knowledge on both learning topics and did not differ significantly (Table 1).

As all hypotheses are in regard of the statistical interaction of cognitive style (representing a between-subjects factor) and specific aspects of learning behavior and learning outcome (representing within-subjects factors), repeated-measures analyses of variance (RM-ANOVA) were computed, using a multivariate approach. In all analyses it was tested whether the statistical

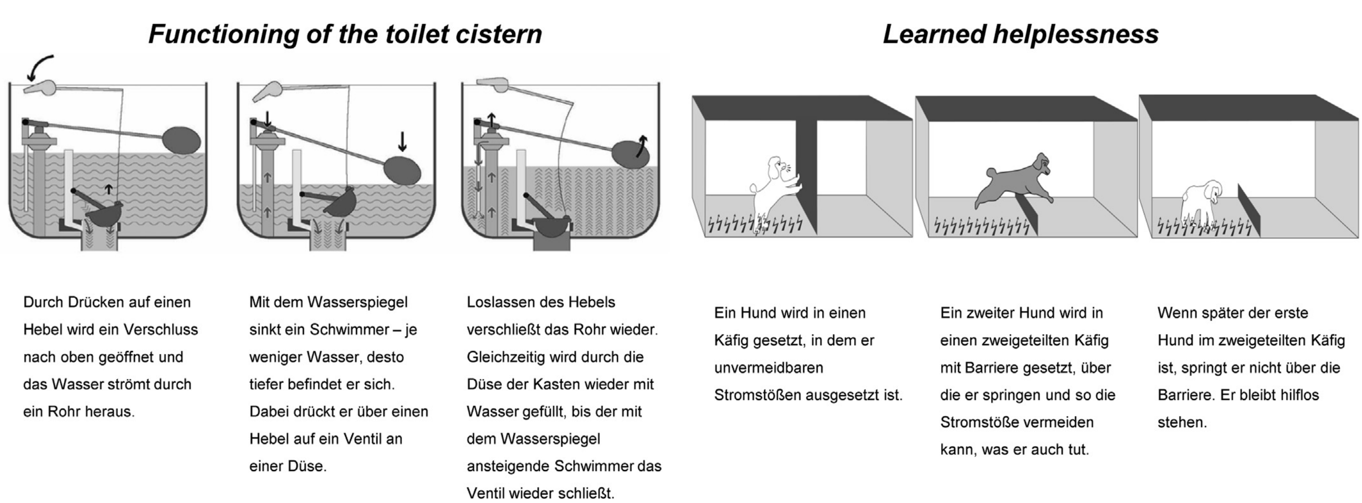

Fig. 1. The two sets of learning materials (stimuli) used in the study.

Table 1
Prior knowledge on the study topics.

Topic	Visualizers	Verbalizers	t	df	p	Range[a]
	M (SD)	M (SD)				
Toilet cistern	0.31 (0.28)	0.20 (0.23)	1.21	30	0.236[ns]	0–1
Learned helplessness	0.06 (0.17)	0.02 (0.06)	1.03	18.95	0.316[ns]	0–1

[ns]Non-significant.
[a] The range of points the participants could possibly receive on the scale.

interaction of cognitive style and learning behavior (or learning outcome) would differ with respect to the learning topic (toilet cistern versus learned helplessness) by calculating the triple interaction of cognitive style, learning behavior (or learning outcome), and learning topic. Note that we were not interested in determining any main effects of the learning topics as the two learning sets differed with regard to pre-defined learning time and instrumentation of learning outcome variables. Therefore, before calculating RM-ANOVAs, all dependent learning behavior variables (or learning outcome variables) representing levels of within-subjects factors in our analyses were linearly transformed to standardized (z) scores with $M = 0$ and $SD = 1$. The means and standard deviations of all transformed variables are given in the Appendix. After computing RM-ANOVAs we calculated simple main effects. As we were mainly interested in differences between visualizers and verbalizers, in the following, we only report simple main effects describing comparisons between visualizers and verbalizers.

3.1. Hypothesis 1: time focusing on pictures and texts

For illustrative purposes, two exemplary heat maps for the toilet cistern topic show the viewing patterns (in terms of the amount of fixations participants made in certain areas of the stimulus) of a typical verbalizer (Fig. 2) and a typical visualizer (Fig. 3). Red and yellow colors indicate many fixations, green and blue colors fewer fixations. It can easily be seen that there are huge differences between the two viewing patterns with the visualizer focusing on pictures and the verbalizer focusing on texts.

The first hypothesis states that (a) visualizers will generally spend more time focusing on pictures than verbalizers and that (b) verbalizers will generally spend more time focusing on texts than visualizers. In order to test this hypothesized interaction, we calculated a RM-ANOVA on dwell time with cognitive style (visualizer, verbalizer) as the between-subjects factor and area of interest (AOI: text, picture) and learning topic (toilet cistern, learned helplessness) as two completely crossed within-subjects factors. Results indicated the expected interaction of cognitive style and AOI, $F(1,30) = 19.45$, $p < 0.001$, $\eta^2 = 0.393$. All other effects in the linear model, including the triple interaction of cognitive style, AOI, and learning topic, were not statistically significant, all $F < 1$.

Fig. 4 displays the interaction of cognitive style and AOI. As expected, visualizers ($M = 0.56$, $SD = 0.58$) spent more time focusing on pictures than verbalizers ($M = -0.56$, $SD = 0.83$), simple main effect $t(30) = 4.47$, $p < 0.001$, $d = 1.63$, whereas verbalizers ($M = 0.56$, $SD = 0.88$) spent more time focusing on texts than visualizers ($M = -0.56$, $SD = 0.56$), simple main effect $t(30) = 4.29$, $p < 0.001$, $d = 1.57$. The missing triple interaction indicates that this interaction pattern does not differ for the two learning sets.

3.2. Hypothesis 2: identifying relevant areas in pictures

The second hypothesis states that (a) visualizers will enter relevant areas in pictures sooner than verbalizers, and that (b) verbalizers will enter the irrelevant areas of pictures sooner than visualizers. In order to test this hypothesized interaction, we calculated a RM-ANOVA on entry time with cognitive style (visualizer, verbalizer) as the between-subjects factor and AOI (irrelevant, relevant) and learning topic (toilet cistern, learned helplessness) as two completely crossed within-subjects factors. Results indicated the expected interaction of cognitive style and AOI, $F(1,30) = 7.26$, $p = 0.011$, $\eta^2 = 0.195$. All other effects in the linear model, including the triple interaction of cognitive style, AOI, and learning topic, were not statistically significant (all $p > 0.10$).

Fig. 5 displays the interaction of cognitive style and AOI. In contradiction to our expectations, visualizers ($M = -0.10$, $SD = 0.68$) did not enter the relevant areas of pictures sooner than verbalizers ($M = 0.10$, $SD = 0.77$), the simple main effect was not significant, $t(30) = 0.75$, $p = 0.462$, $d = 0.27$. On the other hand, in line with our expectations, verbalizers ($M = -0.34$, $SD = 0.80$) entered the irrelevant areas of pictures sooner than visualizers ($M = 0.34$, $SD = 0.52$), simple main effect $t(30) = 2.83$, $p = 0.008$, $d = 1.03$. The missing triple interaction indicates that this interaction pattern does not differ for the two learning sets.

3.3. Hypothesis 3: shifting the point of focus

The third hypothesis states that (a) visualizers will shift their point of focus from picture to picture more frequently than verbalizers, and that (b) verbalizers will shift their point of focus from text to text more frequently than visualizers.

In order to test this hypothesized interaction, we calculated a RM-ANOVA on number of transitions with cognitive style (visualizer, verbalizer) as the between-subjects factor, type of transitions (from picture to picture, from text to text), and learning topic (toilet cistern, learned helplessness) as two completely crossed within-subjects factors. Results indicated the expected interaction of cognitive style and AOI, $F(1,30) = 9.94$, $p = 0.004$, $\eta^2 = 0.249$. All other effects in the linear model, including the triple interaction of cognitive style, type of transitions, and learning topic, were not statistically significant (all $p > 0.20$).

Fig. 6 displays the interaction of cognitive style and AOI. As expected, visualizers ($M = 0.36$, $SD = 0.83$) shifted their point of focus from picture to picture more frequently than verbalizers ($M = -0.36$, $SD = 0.68$), simple main effect $t(30) = 2.66$, $p = 0.012$, $d = 0.97$, whereas verbalizers ($M = 0.38$, $SD = 1.11$) shifted their point of focus from text to text more frequently than visualizers ($M = -0.38$, $SD = 0.44$), simple main effect $t(19.57) = 2.54$, $p = 0.020$, $d = 0.93$. The missing triple interaction indicates that this interaction pattern does not differ for the two learning sets.

Additionally, we tested whether there is a difference between visualizers and verbalizers concerning shifting the point of focus from picture to text or vice versa. This difference, however, was not significant, $t(30) < 1$.

3.4. Hypothesis 4: learning outcome

The fourth hypothesis states that visualizers will achieve better learning outcomes on comprehension scales than verbalizers. In order to test this hypothesized interaction, we calculated a RM-ANOVA on learning outcome with cognitive style (visualizer, verbalizer) as the between-subjects factor and type of learning outcome (retention, comprehension) and learning topic (toilet cistern, learned helplessness) as two completely crossed within-subjects factors. Results indicated the expected interaction of cognitive style and type of learning outcome, $F(1,30) = 4.71$, $p = 0.038$, $\eta^2 = 0.136$. All other effects in the linear model, including

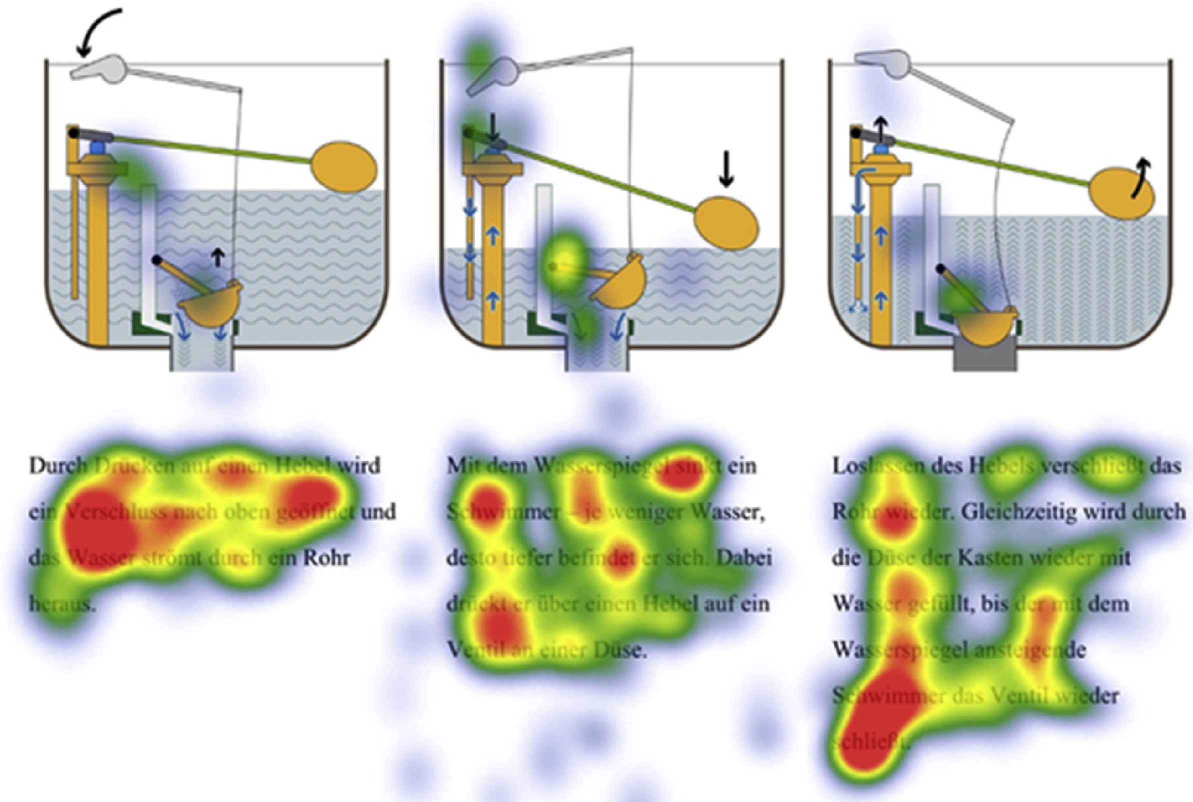

Fig. 2. A heat map of a verbalizer.

Fig. 3. A heat map of a visualizer.

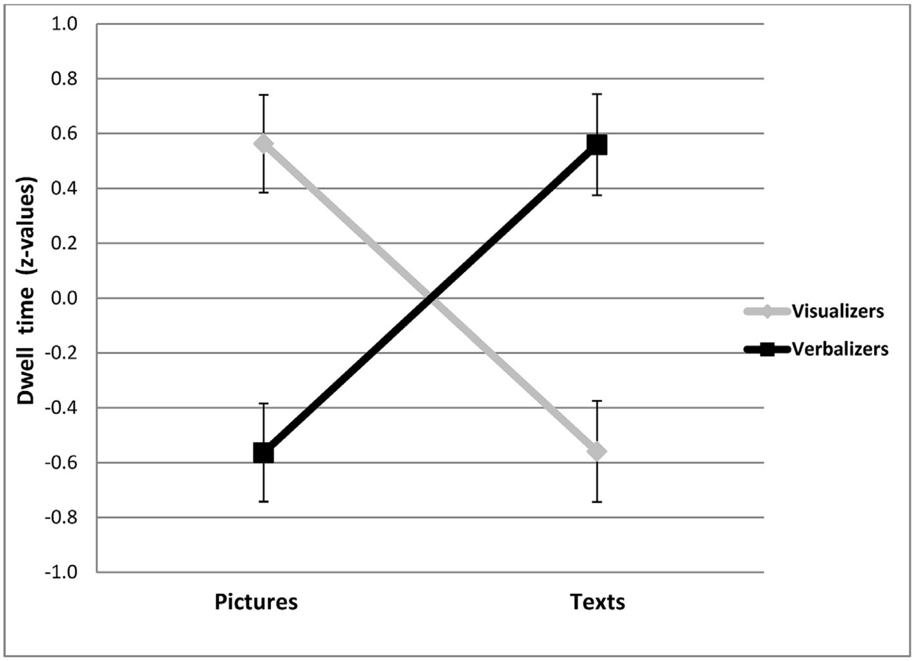

Fig. 4. Dwell time (z-values) as a function of cognitive style.

the triple interaction of cognitive style, AOI, and learning topic, were not statistically significant (all $p > 0.05$).

Fig. 7 displays the interaction of cognitive style and type of learning outcome. As expected, visualizers ($M = 0.32$, $SD = 0.64$) outperformed verbalizers ($M = -0.32$, $SD = 0.65$) on comprehension, simple main effect $t(30) = 2.85$, $p = 0.008$, $d = 1.04$. However, there was no significant difference between visualizers and verbalizers on retention, simple main effect $t(30) = 0.49$, $p = 0.626$, $d = 0.18$. The missing triple interaction indicates that this interaction pattern does not differ for the two learning sets.

4. Discussion

In the present study, we provided more evidence for the assumption that people, classified according to their visualizer-verbalizer cognitive style, differ in their learning behavior in terms of using pictorial and verbal information while learning. When confronted with information that is comparable in terms of content and presented to them in texts and in pictures, verbalizers tend to rely on verbal information and visualizers tend to rely on pictorial information (see also Mehigan et al., 2011; similarly;

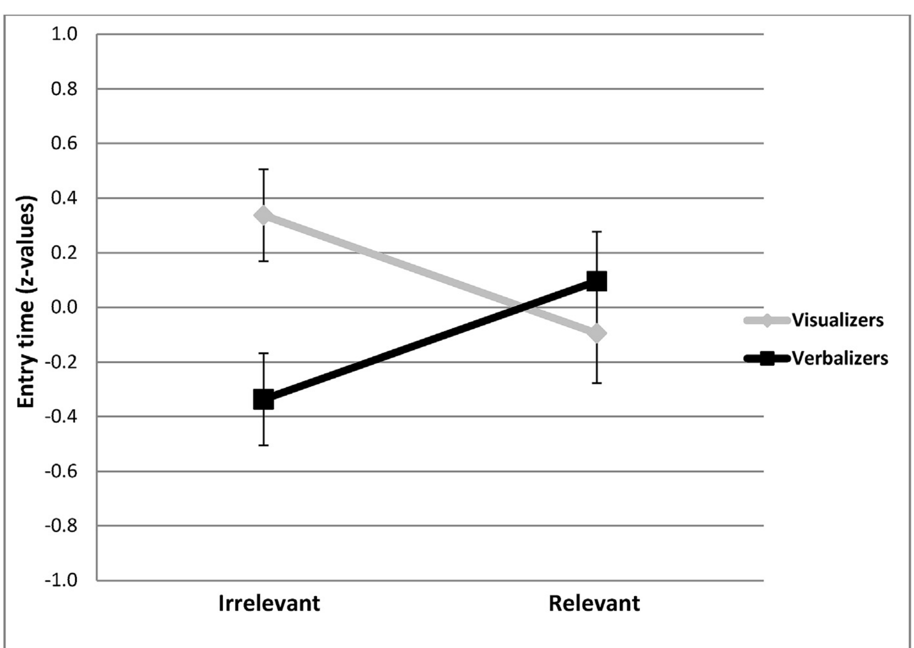

Fig. 5. Entry time (z-values) as a function of cognitive style.

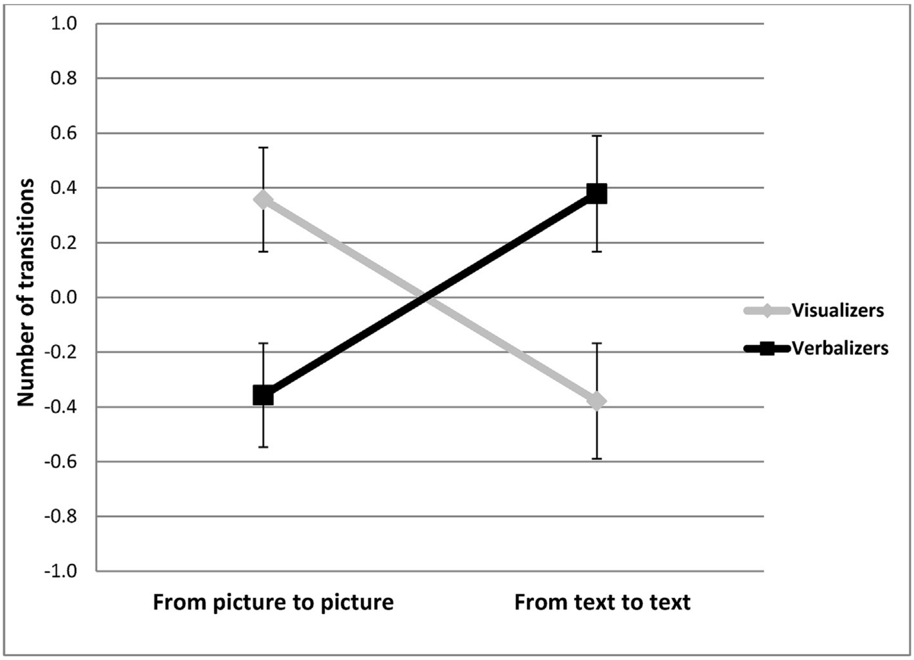

Fig. 6. Number of transitions (z-values) as a function of cognitive style.

Tsianos et al., 2009), regardless of the type of knowledge the topic provides (e.g., conceptual knowledge vs. knowledge regarding functioning of a mechanical system) and its level of difficulty. Our results support not only the existence of the visual-verbal cognitive style but also its influence on learning behavior.

Having in mind that many studies showed learning to be heavily driven by text (Hannus & Hyönä, 1999; Schmidt-Weigand et al., 2010) and that students generally prefer to look at text rather than at pictures while learning (Rayner et al., 2001), our results indicate that visualizers seem to contradict this typical text-oriented way of learning: Concerning the patterns of eye movements in our study, visualizers showed rather clearly a strong picture-oriented way of learning. Verbalizers, on the other hand, not only spent less time on the pictures than visualizers but also tended to enter irrelevant parts of pictures sooner than verbalizers. Visualizers tended to switch to irrelevant parts of stimuli later. These observations partly support our assumption that visualizers are kind of experts on pictures (at the same time, verbalizers seem to be experts on texts) which is in line with findings of Hannus and Hyönä (1999) with high-ability and low-ability students, and those of Jarodzka et al. (2010) with experts and novices, in that visualizers, comparably to high-ability students and experts, seem to concentrate on the relevant areas of pictures while learning (even though they did not have a higher prior knowledge).

The ability to use pictures in a more efficient way can have some advantages. In our study, learning outcomes regarding

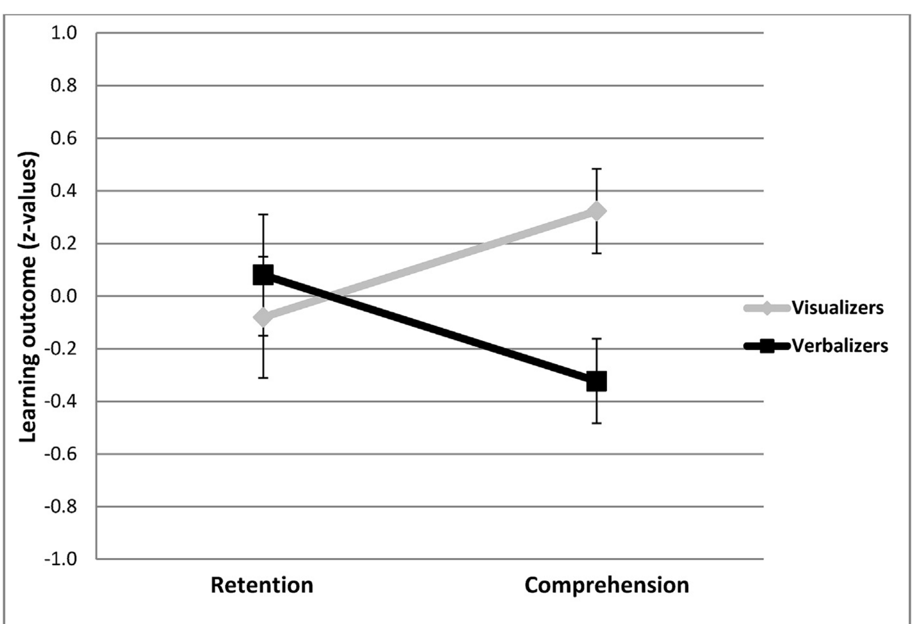

Fig. 7. Learning outcome (z-values) as a function of cognitive style.

comprehension were higher in visualizers than in verbalizers. This result is even more interesting when we consider that both groups displayed an active way of learning, manifested by the number of transitions, but within their preferred mode of information (pictures for visualizers and texts for verbalizers). We might state that both groups invested comparable mental effort in order to understand the topic, but on different areas of interest. We did not observe differences between both groups in the number of transitions between pictures and texts. Further research is needed to figure out whether understanding pictures is crucial to achieve good learning outcomes or whether visualizers processed information from pictures in a way that allowed them to retrieve more useful information.

The lack of differences between both groups of participants in terms of the number of text-to-picture and picture-to-text transitions also confirms that the learning material was self-contained, so participants could freely choose the source of information they prefer (cf. Holsanova et al., 2008; Sweller et al., 1998).

4.1. Limitations

Our study has some obvious limitations. While interpreting the results, we need to remember that the visualizer-verbalizer cognitive style is a dimension. Most people display both styles (visual and verbal) to some extent. Strong visual or verbal cognitive style (especially verbal cognitive style; cf. Cao & Nishihara, 2012) is rather rare; hence the participants of our study, as people with high levels of either visual or verbal cognitive style, represent a selected group, which restricts our findings to this group only. (One might even interpret their strongly manifested visual or verbal cognitive style as a kind of learning disorder, as they did not seem to be able to connect pictorial and textual material adequately.). One should also always bear in mind that the results reported in this study were derived from a rather small sample, which restricts a possible generalization even more. On the other hand, our results do not differ for the two learning sets, in spite of their diversity in terms of content and difficulty, which speaks for a high external validity of our results.

When designing the study, we had to make decisions, which might have made our study simpler and easier to interpret but at the same time widened the range of limitations. Namely, we had to choose certain types of illustrations, levels of difficulty of the vocabulary, and specific topics, risking that our choice would be not optimal for each participant. Hence, the results of our study should always be considered in the context of the choices we made. Further research conducted with many different types of pictures and texts might widen the generalizability of these results.

Furthermore, unlike with pictures, we could not establish relevant and irrelevant parts of stimuli in the text sections of our learning materials, as the eye tracker was not precise enough for such a fine-grained differentiation between textual parts. Further research is needed to enable more precise conclusions about visualizers' and verbalizers' learning behavior concerning text-based information.

The last important limitation of our study worth mentioning is typical for research addressing individual traits. Namely, it is usually not feasible to design a study which can give results directly linkable to cognitive style only. We have to accept the existence of many additional variables influencing the results. One possible way to minimize these difficulties is to add even more control variables to the design, as well as to replicate the study using diverse layouts and topics.

4.2. Conclusions

Our results can be considered as an additional indication that individual differences in visualizer-verbalizer cognitive style do exist (Pashler, McDaniel, Rohrer, & Bjork, 2009). They can be observed in eye movements (Mehigan et al., 2011; Tsianos et al., 2009) and, thus, have a considerable influence on learning behavior, and on learning outcome (cf. Massa & Mayer, 2006; Rogowsky, Calhoun, & Tallal, 2015).

The present study can thus be regarded as a first step towards examining the actual learning behavior of visualizers and verbalizers. However, further research on the influence of the visualizer-verbalizer cognitive style on learning with multimedia using eye-tracking data is highly recommended; especially with randomly chosen participants rather than participants with high levels of either visual or verbal cognitive style.

Further studies might also want to investigate different layouts in terms of types of illustrations, varying difficulties of texts and topics, and their influence on learning outcomes of visualizers and verbalizers. It would also be interesting to investigate, in detail, how verbalizers learn from pictures only and how visualizers learn from text only. A study with more systematic variations of layout and sources of information, providing information, for example, exclusively with text or pictures, might overcome some of the limitations of our research. Does the lack of the preferred type of representation impair the learning effect (e.g., Plass et al., 1998)? This question, among others, still has no clear answer.

In the end, casting more light on the way in which visualizers and verbalizers use pictorial and verbal information will hopefully provide valuable input to teaching, learning, and the design of learning materials including e-learning as well as text books in schools.

Acknowledgements

The project was kindly funded by German Research Foundation (grant no. HO 4303/6-1).

Appendix. Means and standard deviations of all transformed variables.

Variables	Toilet cistern		Learned helplessness	
	M	*SD*	*M*	*SD*
Dwell time – pictures (in ms)	77453.56	27858.03	37782.86	16463.04
Dwell time – texts (in ms)	64917.35	27797.06	45115.53	16423.12
Entry time – irrelevant (in ms)	21438.32	12588.15	22315.25	10794.60
Entry time – relevant (in ms)	27028.86	11182.94	5451.95	3061.61
Transitions across pictures (number)	32.81	16.87	25.78	13.98
Transitions across texts (number)	6.19	5.83	7.28	5.96
Transitions between pictures and texts (number)	34.38	13.94	22.22	9.04
Retention[a]	0.83	0.20	0.68	0.19
Comprehension[a]	0.60	0.23	0.35	0.26

[a] The scale ranged from 0 to 1.

References

Baddeley, A. (1998). *Human memory*. Boston: Allyn & Bacon.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.

Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology, 23*, 638–663.

Cao, J., & Nishihara, A. (2012). Understanding learning style by eye tracking in slide video learning. *Journal of Educational Multimedia and Hypermedia, 21*, 335–358.

Cassidy, S. (2004). Learning styles: An overview of theories, models, and measures. *Educational Psychology, 24*, 419–444.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149–210.

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low – and high – ability children. *Contemporary Educational Psychology, 24*, 95–123.

Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 447–465). Cambridge, England: Cambridge University Press.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and instruction, 21*, 209–249.

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations – a meta-analytic review. *Educational Psychology Review, 22*, 245–269.

Höffler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations – evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior, 27*, 209–216.

Höffler, T. N., Prechtl, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures. *Learning and Individual Differences, 20*, 479–483.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye Tracking. A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2008). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology, 23*, 1215–1226.

Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 44–55.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*, 146–154.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509–539.

Kirby, J. R., Moore, P., & Schofield, N. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology, 13*, 169–184.

Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist, 48*, 169–183.

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition, 33*, 710–726.

Marks, D. F. (1973). Visual imagery in the recall of pictures. *British Journal of Psychology, 64*, 17–24.

Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*, 356–384.

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*, 95–109.

Massa, L., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences, 16*, 321–335.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge, England: Cambridge University Press.

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*, 833–846.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist, 38*, 43–52.

Mehigan, T. J., Barry, M., Kehoe, A., & Pitt, I. (2011, July). Using eye tracking technology to identify visual and verbal learners. In *Proceedings of the 2011 IEEE international conference on multimedia and expo (ICME), 11–15 july 2011, Barcelona, Spain*.

Messick, S. (1984). The nature of cognitive styles: Problems and promises in educational research. *Educational Psychologist, 19*, 59–74.

Mousavi, S., Radmehr, F., & Alamolhodaei, H. (2012). The role of mathematical homework and prior knowledge on the relationship between students' mathematical performance, cognitive style and working memory capacity. *Electronic Journal of Research in Educational Psychology, 10*, 1223–1248.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.

Paivio, A., & Harshman, R. (1983). Factor-analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 37*, 461–483.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles concepts and evidence. *Psychological Science in the Public Interest, 9*, 105–119.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*, 25–36.

Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge, England: Cambridge University Press.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*, 219–226.

Richardson, A. (1977). Verbalizer-visualizer: A cognitive style dimension. *Journal of Mental Imagery, 1*, 109–125.

Riding, R. J. (1997). On the nature of cognitive style. *Educational Psychology, 17*, 29–50.

Riding, R. J., & Cheema, I. (1991). Cognitive Styles – an overview and integration. *Educational Psychology, 11*, 193–215.

Riding, R. J., & Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology, 63*, 297–307.

Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2015). Matching learning style to instructional method: Effects on comprehension. *Journal of Educational Psychology, 107*, 64–78.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*, 100–110.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141–156.

Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology, 74*(1), 1–9.

Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist, 52*, 700–712.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Tsianos, N., Germanakos, P., Lekkas, Z., Mourlas, C., & Samaras, G. (2009, July). Eye – tracking users' behavior in relation to cognitive style within an e-learning environment. In *Proceedings of the ninth IEEE international conference on advanced learning technologies (icalt 2009), 15–17 july 2009, Riga, Latvia*.

Witkin, H. A. (1973). The role of cognitive style in academic performance and in teacher-student relations. *ETS Research Bulletin Series, 1*, i–58.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*, 345–376.