ZGD Zeitschrift für Geographiedidaktik
Journal of Geography Education

Vol. 50(3), 159-179
ISSN 2698-6752
DOI: 10.18452/25715
CC BY NC ND

Which Factors Influence Learners' Visual Attention to Images in Geographic Learning Media?

Welche Faktoren beeinflussen die visuellen Aufmerksamkeit Lernender für Abbildungen in geographischen Lernmedien?

¿Qué factores influyen la atención visual de los alumnos a las imágenes en los medios de aprendizaje geográfico?

Yvonne Behnke ✉

Zusammenfassung Eine Eye-Tracking-Studie mit Fragebogen untersuchte potenzielle Einflussfaktoren auf visuelle Aufmerksamkeitsmuster beim Lernen mit Fotos, Grafiken und Karten in Geographieschulbüchern (u.a. Usability, Interesse, Medienpräferenz). Dafür wurden Daten von achtundfünfzig Lernenden (14-17 Jahre, Gymnasium) analysiert. Stimuli waren drei Geographieschulbuchdoppelseiten zu drei interessensbasiert ausgewählten Themen. Eine Faktorenanalyse extrahierte relevante Faktoren für visuelle Aufmerksamkeitsmuster auf Fotos, Grafiken und Karten der Schulbuchseiten. Usability, Textpräferenz, Bildpräferenz und individuelles Interesse (Thema Erdbeben) und erklären 48,5 % Varianz (relevantester Faktor: Usability). Individuelles Interesse war weniger relevant als situatives Interesse. Daraus werden Vorschläge zur Unterstützung des Lehrens und Lernens mit multimodalen geographischen Medien abgeleitet.

Schlüsselwörter Schulbuchgestaltung, Eye-Tracking, Georaphiedidaktik, visuelle Aufmerksamkeit, Multimedia Learning

Abstract An eye-tracking and questionnaire-based study investigated possible factors influencing visual attention (e.g., usability, interest, and media preferences) when learning with photos, graphics, and maps. Data from 58 learners (14–17 years, secondary school) were analyzed. Three Geography textbook spreads on three interest-led topics (earthquakes, rainforests, agriculture) served as stimuli. Factor analysis identified relevant factors for learners' visual attention to photos, graphics, and maps in the investigated textbook spreads. Usability, text preference, image preference, and individual interest (topic earthquakes) explained 48.5 per cent of the variance (most relevant factor: usability). Individual interest was less relevant than situational interest. Based on these findings, the paper presents suggestions to support teaching and learning with multimodal geographic learning media.

Keywords textbook design, eye-tracking, Geography Education, visual attention, multimedia learning

Resumen Un estudio de seguimiento ocular y basado en cuestionarios investigó los posibles factores que influyen la atención visual (por ejemplo, la facilidad de uso, el interés y las preferencias de medios de educación) al aprender con fotos, gráficos y mapas. Se analizaron datos de 58 estudiantes (14–17 años, escuela secundaria). Tres tiradas de libros de texto de Geografía sobre tres temas de interés (terremotos, selvas tropicales, agricultura) sirvieron como estímulo. El análisis factorial identificó factores relevantes para la atención visual de los estudiantes a las fotografías, gráficos y mapas en los libros de texto investigados. La usabilidad, la preferencia de texto, la preferencia de imagen y el interés individual (terremotos temáticos) explicaron el 48,5% de la variabilidad (factor más relevante: usabilidad). El interés individual era menos relevante que el interés situacional. Con base en estos hallazgos, este artículo presenta sugerencias para apoyar la enseñanza y el aprendizaje con medios de aprendizaje geográfico multimodal.

Palabras clave diseño de libros de texto, eye-tracking, Didáctica de la Geografía, atención visual, educación multimedial

ZGD 3•22 © The author(s) 2022 Reprints & permission: www.zgd-journal.de

1. Introduction

Geography education maintains a particular relationship with images because the creation and interpretation of images determine an essential part of Geography instruction (Thornes 2004). Widdowson and Lambert (2006) argue in the same vein when they refer to images as powerful tools for transmitting geographical concepts, just as Lambert (2008) emphasizes the importance of images in geographical knowledge construction and Uhlenwinkel (2007) highlights the didactic potential inherent in working with pictures in Geography lessons.

As textbooks—despite the variety of digital media available—remain essential tools for students' learning and the most important educational medium in schools (Knight et al. 2017; Fuchs & Bock 2018; Bagoly-Simó et al. 2019), this contribution focuses on images in Geography textbooks.

Studies revealed learners' limited attention to visuals in educational media (Schnotz et al. 2014; Behnke 2016a; Eitel 2016). Many learners face challenges in learning with images (Schnotz et al. 2014; Scheiter et al. 2018; Seufert 2019). Other studies found that visuals in learning may positively affect learning motivation (Schnotz et al. 2009) and trigger situational interest (Magner et al. 2016; Endres et al. 2020).

Given this backdrop, learners' visual attention to images in learning media is considered interrelated with learners' interest and is i.a. driven by motivational factors (Pettersson 2000; Magner et al. 2016). Studies named interest as a factor influencing learning motivation and highlighted relationships between attention and interest (Prenzel et al. 1986; Krapp 2002; Schiefele 2012; Renninger & Hidi 2016).

However, although there is evidence for interrelations between interest and learning indicators, they have so far mainly been researched with text (Renninger & Hidi 2006; Schiefele 2012; Magner et al. 2016). Visual processes that may trigger learners' interests are, in large part, still unclear (Renninger et al. 2019), and how learners visually interact with images in learning media is still a marginal topic in educational research (Ballstaedt 2017).

Another shortcoming of previous studies on visuals in learning media is their lack of a multidisciplinary perspective. The present study includes findings from Visual Communication, Media Studies, Geography Education, and Educational Psychology to examine which parameters affect learners' attention to visuals in Geography textbooks through two questions:

> Q1. How do learners' interests in geographical topics and media influence visual attention to photos, graphics, and maps when learning with Geography textbooks, and what other possible factors are involved?
>
> Q2. How can extracted factors influencing visual attention patterns on text elements and images in geography textbooks be structured, summarized, and characterized?

In summary, this study's aim is to investigate factors affecting students' visual attention when learning with photos, graphics, and maps in Geography textbooks, explicitly addressing the interest factor.

2. Theoretical Background

This section begins with the presentation of the theoretical model *Usability Parameters of Well-designed Geography Textbook Visuals* (Section 2.1, Fig. 1), as the two research questions, the eye-tracking research design, as well as the corresponding questionnaire of the present study were developed based on this model. Sections 2.2 and 2.3 elaborate on selected theoretical approaches from the model (Fig. 1) along with empirical evidence relevant to examine Q1 and Q2. Section 2.4 explores eye-tracking as analysis method for visual attention processes to images in learning media.

2.1 Theoretical Modeling

Based on evidence from a structured literature review supplemented with findings from her exploratory study (Behnke 2016a, 2016b, 2017a), the author developed the theoretical model of *Usability Parameters of Well-designed Geography Textbook Visuals*. The model is based on a synthesis of established theoretical approaches for effective learning with visuals from Design, Educational Psychology, Media Studies, and Geography adapted to Geography Education. Six usability parameters of well-designed textbook visuals were derived: usefulness, interest, helpfulness, comprehensibility, aesthetics, and orientation (Fig. 1). Behnke (2021) elaborates on the model in detail. Fig. 1 only highlights those parameters that contain theoretical approaches relevant to the present paper.

2.1 Interest Research in Education and Learners' Interest in Textbook Visuals

Both Geography Education (Hemmer & Hemmer 2010; Teschner 2011; Hemmer et al. 2020) and Educational Psychology (Schiefele 2012; Magner et al. 2016; Renninger & Hidi 2016) consider learners' interest as an important factor influencing students' learning processes. In the context of education, interest describes learners' positive attitude towards a school subject, a topic, or a learning medium (Krapp 2002). When researching learners' interests, both Educational Psychology and Geography Education refer to the *Person-Object Approach to Interest* (POI; Krapp 2002).

Krapp's (2002) POI relates the pedagogical theory of interest which defines *interest* through the relationship between the person (learner) and the object (learning topic) (Schiefele et al. 1983; Prenzel et al. 1986) to Ryan and Deci's (2000) *Self-determination Theory* (SDT), naming competence, autonomy, and relatedness as key motivational factors for learning. Consequently, Krapp's POI includes affective components, like curiosity, and cognitive factors, such as perceived usefulness (Krapp 2002; Renninger & Hidi 2006; Schiefele 2012; Magner et al. 2016). Thus, Krapp's POI exhibits interrelations between educational theories of interest (pedagogical theory of interest) and motivational theories, such as SDT.

The interest construct differentiates between situational and individual interest. Individual interest is a personal attitude, such as the preference for a learning topic (e.g., earthquake). Situational interest is a reaction to an environmental input, such as a salient image (Krapp 1992; Renninger & Hidi 2006, 2016; Schiefele 2012; Magner et al. 2016). Therefore, situational interest can be interpreted as a usability parameter, triggered by perceived usefulness in achieving a goal, and individual interest as a personal preference, such as towards a learning topic (Krapp & Prenzel 1992).

Consequently, when examining learners' interests, there should be a distinction between situational and individual interest. However, while several studies on students' interest in Geography Education have been conducted to date (Hemmer & Hemmer 2010; Teschner 2011; Hemmer et al. 2020), they have mainly focused on individual interests (topics, methods, media) and less on situa-

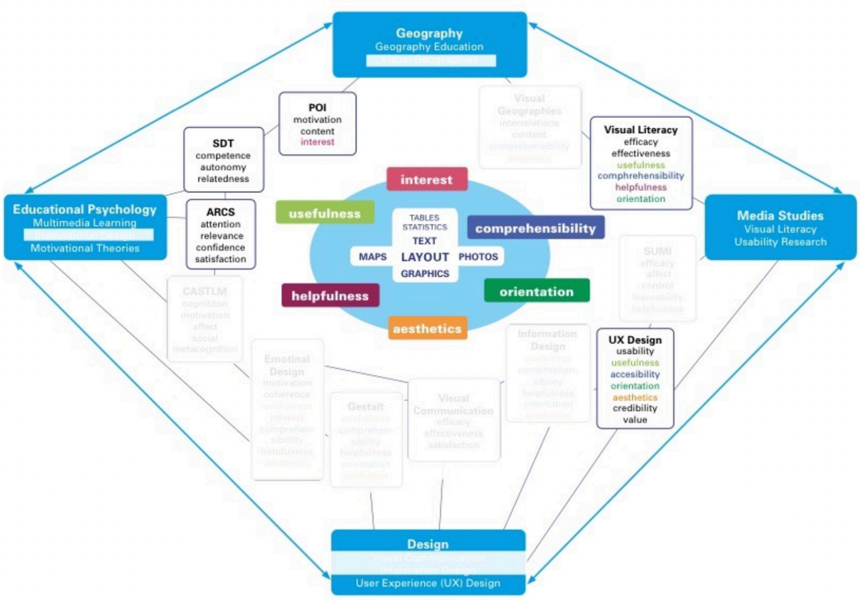


Fig. 1. Usability parameters of well-designed Geography textbook visuals (Source: Behnke 2021, amended)

tional interests, such as media in specific learning situations. Moreover, most studies have measured students' level of interest and not why learners are interested (Renninger & Hidi 2006, 2016). Therefore, Schiefele (2012) emphasized the need for more research to examine the relationships between interest and further factors influencing students' learning.

2.2 Interest, Motivation, Usability, and Visual Attention to Textbook Images

Visual attention, as a prerequisite for successful knowledge construction with images, is the process that turns looking into seeing (Carrasco 2011) by filtering out irrelevant and selecting relevant information in the eye of the beholder (Bischof et al. 2019).

Empirical evidence supports the notion that learners' visual attention to images in learning media can be influenced by various affective, cognitive, and motivational factors (Carrasco 2011). Learners' visual attention to a textbook image can be triggered by situational interest (salient detail) (Pettersson 1995, 2000; Magner et al. 2016); intrinsic factors, such as learners' interest in the topic; motivational factors, such as perceived usefulness to achieve a goal (e.g., content comprehension) (Magner et al., 2016; Wong & Adesope 2021); or extrinsic factors, such as a learning task.

Moreover, on the one hand, many learners face challenges in decoding and interpreting images according to the respective learning context (Weidenmann 1989, 1994; Horz & Schnotz 2009; Baadte & Schnotz 2012). For example, the illusion of full understanding (Peeck 1993) states that learners' visual attention, especially to familiar images, is often limited because they assume they have already captured all information and so miss relevant details to complete a task.

On the other hand, Schnotz et al. (2009) emphasized the motivational role of visuals in learning media in attracting interest, because they allow learners to focus their attention, thus leading to deeper cognitive processing. Therefore, this contribution focuses on intrinsic and motivational factors.

Moreover, motivational factors, such as perceived usefulness also characterize the usability of a textbook image because usability describes the extent to which it serves learners to achieve a goal, such as content comprehension, quickly, successfully, and satisfactorily (Bojko 2014).

Because this study considers motivational factors, motivational models are included in the theoretical considerations, namely Ryan and Deci's (2000) SDT with the key factors competence, autonomy, and relatedness, and Keller's (2010) ARCS model, where attention (A), relevance (R), confidence (C), and satisfaction (S) are key factors of learning motivation (Fig. 1).

In summary, an interrelationship between interest, motivation, usability, and visual attention to textbook images can be deduced from theory and empirical evidence.

2.3 Relationships between Visual Attention and Eye Movements

Eye-tracking measures eye movements—which point of a stimulus is observed, for how long, and in what order (Holmqvist et al. 2010; Bojko 2014). Eye movements can be generally divided into saccades and fixations. A saccade is a rapid eye movement from one point of interest to another, while a fixation is a pause between two saccades, where visual information is gleaned and cognitively processed (Duchowski 2007; Bischof et al. 2019). Thus, new information is acquired during fixations (Rayner 2009). Here, widespread agreement exists on interrelations between visual perception, visual attention, and the cognitive processing of visual information (Holmqvist et al. 2010; Klein & Ettinger 2019; Emhardt et al. 2020).

Although past studies showed that just a few fixations are sufficient to perceive a rough *scene gist* (Loschky et al. 2018), at least 150 ms are needed for cognitive processing, to decode image properties (Rayner 2009; Rayner et al. 2009; Loschky et al. 2018). Moreover, observers can focus only on a small area at a time (de Koning et al. 2010). Therefore, in textbooks, learners continuously need to decide on which detail to focus their attention.

Learners' visual attention may be elicited by exogenous factors (e.g., salient details) or endogenous factors, such as a specific task (Hyönä 2010). Different tasks (visual search, scene perception, text reading) induce different eye movements (Rayner 2009; Emhardt et al. 2020).

However, it is difficult to explain, solely based on eye-tracking data, why a participant observed a stimulus (Bischof et al. 2019). Hence, it is beneficial to interlink eye-tracking data with additional data in the form of a questionnaire. Conversely, it is useful to interlink questionnaire data with information on interactions between learners and learning environments (Sass et al. 2017; Emhardt et al. 2020), because their integration provides more comprehensive insights (Guo et al. 2019).

Therefore, data from participants' eye movements, a questionnaire, and three different Geography textbook spreads were triangulated to answer the research questions as will be elaborated on in the next section.

3. Methods

3.1 Previous Work

An explorative eye-tracking study (Behnke 2016a, 2016b, 2017) analyzed participants' visual attention to text, graphics, and photos in five different Geography textbook spreads, presenting an identical topic with five different designs.

The main finding was that participants mainly focused on text and devoted limited attention to photos and graphics, in line with Scheiter et al. (2014) and Schnotz et al. (2014).

Questionnaire

The questionnaire consisted of 39 questions, 32 of which were included in the quantitative data analysis of the present study. Participants rated the items on a five-point Likert-scale ranging from 5 (very good) to 0 (not at all). Two open questions were analyzed separately.

In the first section, participants rated the visual design of three different Geography textbook spreads (three different designs, three different topics) along the categories of layout, photos, graphics, maps, and text in line with six usability parameters (Fig. 1) for textbook spreads *A*, *B*, and *C* separately. Through two open questions, participants reasoned which of the three textbook spreads they assessed as best and as worst for their learning.

In section two, participants rated their individual preferences for text, photos, graphics, and maps in learning media in general.

Section tree evaluated participants' individual interest in the geographical topics presented on the stimuli (earthquakes, tropical rainforest, agriculture) and, in section four, their individual interest in the school subject Geography. The reliability of the scale as measured by Cronbach's alpha was <= 0.775.

Experiment Design

The study examined how participants observed three different Geography textbook spreads during an eye-tracking experiment. Afterwards, participants assessed the stimuli with a questionnaire. The stimuli addressed three different topics. The study was not intended to assess learning performance. Tasks were utilized to direct visual attention.

Participants

A total of 69 students with normal or corrected-to-normal vision (14–17 years old) from four secondary schools in three German federal states (Thuringia, Brandenburg, and Berlin) participated in the experiment from June to October 2019. Data from eleven students were excluded, ten because of data loss (gaze sample rate below 80 per cent threshold), and one due to missing data in the questionnaire. Thus, data from 58 students were analyzed (29 female, 29 male; M_{age} 15.26 years; SD_{age} 0.80). The study was conducted on site at the respective secondary schools. All students participated voluntarily without any remuneration. The study was carried out in compliance with the legal, ethical, and administrative requirements of the respective federal states, school administrations, and Humboldt-Universität zu Berlin.

Stimuli

The stimuli consisted of three textbook spreads (three different topics, three different values for students' interest). The spreads were presented as they appeared in the printed textbook (two pages, original size, high-resolution PDF, full color). The stimuli were selected using the following five criteria:

First, topics with different interest values, according to Hemmer & Hemmer's (2010) study (five-point Likert-scale: 5 = very high interest to 1 = no interest) were selected.

Spread *A*: earthquake/natural disasters (high interest = 4.21 mean)
Spread *B*: tropical rainforest/ecosystems in

tropical areas (medium interest = 3.21 mean) Spread *C*: agriculture (low interest = 2.81 mean).

Second, each spread utilized text, graphics, photos, and maps to develop the topic. Third, the curricula of the respective federal states covered the topics. Fourth, the topic had been taught in the participants' Geography lessons. Fifth, the spreads were from regular textbooks (2015–18) by three textbook publishers.

Apparatus

A Tobii Pro X3 120 remote eye-tracker with a 120 Hz sampling rate, running Tobii Studio software, collected participants' eye movement data. The stimuli were presented on a Dell Latitude 5580 laptop (15,6"; 1920 x 1080 pixel with a 60 Hz refresh rate) at a viewing distance of approximately 60 cm. Data from both eyes were recorded. An IVT filter algorithm with a minimum fixation duration of 150 ms was applied for fixation detection (Rayner 2009; Rayner et al. 2009). Eye movement data were processed with Tobii Studio software. Statistics were calculated using SPSS.

Procedure and Set Task

After an introduction to the procedure and questions examining participants' vision, a five-point calibration was conducted. Subsequently, the task appeared on a grey screen. Participants were asked to observe the textbook spread which appeared on the screen, and then to determine the issue depicted on the spread. Meanwhile, participants' eye movements were recorded. After the participants observed the stimulus, a multiple-choice question with three possible answers examined participants' content comprehension. Participants' solutions were recorded by the software and noted in the protocol. Following task completion, the next textbook spread appeared on the screen.

Each participant observed three textbook spreads (earthquake, tropical rainforest, agriculture) and completed three multiple-choice tasks. There was no time limit during the experiment. Each participant decided how much time to devote to each task. The test design was randomized. Each participant observed the three pages in a different order.

After the eye-tracking experiment, the participants completed the questionnaire. As a reminder, the three stimuli were available as a color copy in the original size.

3.2 Data Analysis

The data originated from three sources: three Geography textbook spreads, the questionnaire, and data of participants' eye movements (fixations).

Each element of the stimuli was marked as an area of interest (AOI), grouped by media type and color-coded (text = red, map = green, graphic = blue, photo = yellow). AOI allow separate data analysis of participants' eye movement on each element and of the AOI groups.

The data analysis underwent a three-step analysis. First, a descriptive data analysis structured and described the data (Figs. 2–6). Thereafter, an exploratory factor analysis identified groups of highly correlated variables and separated them from less-correlated variables (Backhaus et al. 2016) (Figs. 7–8). Subsequently, a confirmatory factor analysis examined the results from step three, reduced and verified extracted factors (Fig. 9).

Multivariate analysis is a frequently used statistical analysis method in eye-tracking research (Lowe & Boucheix 2016; Sauro & Lewis 2016; Zander et al. 2017; Seufert 2019). Therefore, the present study applied multivariate analysis (factor analysis) to examine and structure the collected data (Backhaus et al. 2016). Among others, de Lucio et al. (1996) and Koonsanit et al. (2021) applied facror analyis on eye-tracking studies.

Step 1 (Descriptive Analysis)

The descriptive analysis structured and described the data to identify relevant parameters for participants' visual attention, such as formal visual characteristics of the stimuli (Fig. 2), participants' interests (Fig. 3), and participants' fixations (Figs. 4–6).

Step 2 (Exploratory Factor Analysis) (Figs. 7–8)

The exploratory factor analysis as a structure-discovering method (Backhaus et al. 2016) examined data from the questionnaire and eye-tracking data to identify factors that explain most of the variance for students' visual attention to the stimuli. Relationships between variables were analyzed and summarized in factor groups (Thomson 2004; Backhaus et al. 2015, 2016). In total, 20 variables were analyzed, namely data on participants'

(a) individual attitude to photos, graphics, maps, and text in learning media;

(b) situational attitude to photos, graphics, maps, and text of the stimuli;
(c) fixations on photos, graphics, maps, and text of the stimuli;
(d) individual interest in the school subject Geography;
(e) individual interest in the topics (earthquake, rainforest, agriculture) of the stimuli.

Step 3 (Fig. 9)
Confirmatory factor analysis is a structure-verifying method (Backhaus et al. 2016). It served to verify the identified factors from the exploratory factor analysis and to reduce influencing variables. This was done by extracting key factors by means of the highest regression-based factor scores (< 0.5) to identify relevant factors influencing visual attention to photos, graphics, and maps on stimuli *A*–*C*.

The interpretation of the results integrated the factor analysis and descriptive analysis of the three data sources (eye-tracking data, questionnaire, textbook spreads) against the theoretical background and the findings of Behnke's (2016a) explorative study.

4. Results

In line with the research questions, this section provides the main findings of the data analysis and determines which factors influenced participants' visual attention to discontinuous text during the eye-tracking test.

4.1 Factors Influencing Visual Attention Processes during Learning with Images in Geography Textbooks with a Particular Focus on Learners' Interests (Q1)

Step 1a: Descriptive Analysis of Formal Parameters
Since visual design characteristics may affect learners' visual attention, image placement and image sizes are specified in addition to image caption and image content and their relevance for topic comprehension. This provides initial explanations for participants' visual attention patterns to the stimuli.

The content of the depicted images varies in its relevance for topic comprehension (Fig. 2). While *A* depicts two relevant photos, *C* displays decorative photos, and *B* contains one decorative photo (M2) and one informative photo, concurrently doubling the information of figure M3.

In terms of their visual design and content, the graphics are heterogeneous. *A* presents a reduced infographic explaining how earthquakes occur. Conversely, the style and appearance of figure M3 in *B* may not be appropriate for the students due to the childish illustration style. The information design style of M7 in *C* appears outdated.

The maps are most heterogeneous both in size and relevance. *B* depicts a small map in the margin column that only allows a general localization. In contrast, *C* depicts a thematic map across the left-hand spread's entire width, providing relevant information regarding agriculture in the US.

Step 1b: Questionnaire: Participants' Interest in Textbook Elements
Fig. 3 presents participants' individual interest in the text, photos, graphics, and maps in textbooks, and their situational interest in the stimuli's textbook elements. The participants rated their interest after the eye-tracking test.

Fig. 3 also shows that situational interest in the graphics and maps of the stimuli was rated up to one grade lower than individual interest. Situational interest in the photos of *A* (3.88) was rated higher than individual interest in photos, which may be due to the relevant information they provide for topic comprehension and their aesthetic appeal (Fig. 2).

Overall, participants rated their situational interest in *A* (earthquake) highest for all four elements. The evaluation of individual interest in text, photos, graphics, and maps was consistent with the ranking of situational interest in the elements of *A*, *B*, and *C*, but differed in values.

Both individual and situational interest in maps were rated lowest. Nevertheless, maps still received a medium rating for individual interest (3.34), while situational interest varied between 3.26 (*A*, *C*) and 2.38 (*B*). *B*'s rating is thus the lowest among all elements of the stimuli (Fig. 2).

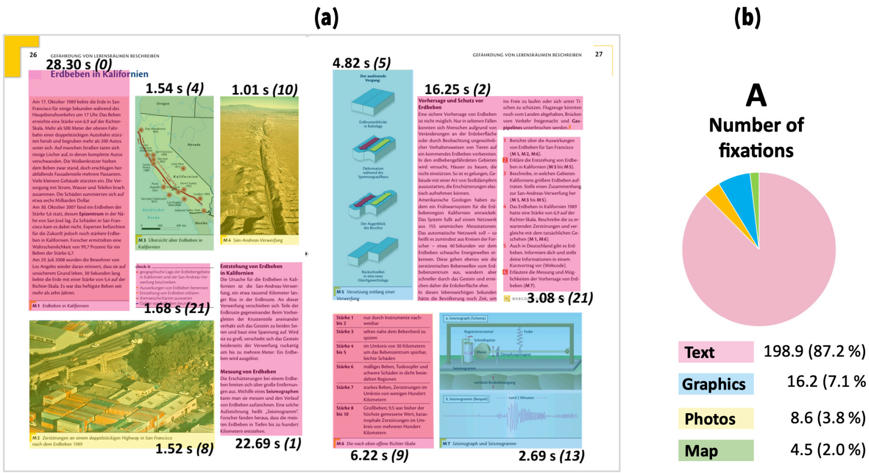
In summary, participants' situational interest was lower for the stimuli than their individual interest in text, photos, graphics, and maps

	Spread *A* (Earthquake in California)	**Spread *B* (Deforestation/Slash & burn)**	**Spread *A* (Agriculture in the USA)**
Stimulus	 © Cornelsen	 © Klett	 © Westermann Gruppe
Photos	*2 photos* **M2** *Placement:* left page, left and middle column, bottom *Size:* two-columned *Content:* Earthquake damage on an urban highway *Caption:* destruction of a highway in San Francisco after the 1989 earthquake *Relevance for topic comprehension:* yes **M4** *Placement:* right page, middle, top *Size:* single-columned *Content:* oblique aerial view of the San Andreas Fault *Caption:* San Andreas Fault *Relevance for topic comprehension:* yes	*2 photos* **M2** *Placement:* left page, middle column, top *Size:* single-columned *Content:* foreground: Yanomami harvesting crops; background: rainforest *Caption:* Yanomami cultivating their crops *Relevance for topic comprehension:* decorative/illustrative **M4** *Placement:* right page, top *Size:* two-columned and margin column *Content:* oblique aerial view, landscape, forest clearing areas in different stages, marks: A–E *Caption:* slash-and-burn landscape with different stages of slash-and-burn agriculture *Relevance for topic comprehension:* doubling information of M5; poor contrast, letters difficult to identify, lacks legend, slash-and-burn islands are difficult to distinguish *Challenge*: relevant information is not recognizable	*3 photos* **M1** *Placement:* left page, top, left column *Size:* Single-columned *Content:* harvested field, combine harvesters; foreground: two farmers *Caption:* US/modern agriculture **M2** *Placement:* left page, top, right column *Size:* single-columned *Content:* red wooden house on a hill, meadow and fence *Caption:* traditional farming **M3** *Placement:* right page, top, *Size:* two-columned *Caption:* oblique aerial view of a large field with carousel irrigation *Caption:* carousel irrigation in the US *Relevance for topic comprehension:* decorative
Graphics	*2 graphics* **M5** *Placement:* right page, top, middle column *Size:* single-columned *Content:* schematic infographics: four steps of the dislocation of Earth crust blocks during earthquake *Caption:* dislocation along a fault line *Relevance for topic comprehension:* yes **M7** *Placement:* right page, bottom, middle and right columns *Size:* two-columned *Content:* infographic. Top: complex infographic; functioning of seismograph; bottom: seismogram *Caption:* seismograph and seismogram *Relevance for topic comprehension:* additional information	*2 graphics* **M3** *Placement:* left page, top, middle column *Size:* single column *Content:* three-part illustrative graphic, three stages of slash-and-burn agriculture, no explanatory text *Caption:* stages of slash-and-burn agriculture *Relevance for topic comprehension:* doubling information of M4 *Challenge:* relevant information might be overlooked due to childish illustration style and missing text labelling **M5** *Placement:* right page, middle column *Size:* single-columned *Content:* bar chart: yield and lifecycle of slash-and-burn agriculture (four years/four crops) *Caption:* yield and lifecycle of slash-and-burn agriculture *Relevance for topic comprehension:* yes *Challenge:* unclear division of axes and columns might cause difficulties in allocation and understanding	*2 graphics* **M5** *Placement:* right page, middle, left column *Size:* single-columned *Content:* six pie charts, US share of world production (wheat, soya, milk, corn, cotton, meat) *Caption:* US share of agriculture of global production *Relevance for topic comprehension:* yes *Challenge:* data need to be decoded and related to other countries; comparative data from other countries are missing **M7** *Placement:* right page, middle, right column *Size:* single-columned *Content:* bar chart with pictograms depicting the increase in agricultural production per farmer since 1940 *Caption:* people supplied by a US farmer *Relevance for topic comprehension:* additional information *Challenge:* year is missing; design is old-fashioned
Maps	*one map* **M3** *Placement:* left page, top, middle column *Size:* single-columned *Content:* reduced thematic map of California earthquakes and San Andreas Fault *Caption:* Earthquake in California *Relevance for topic comprehension:* yes	*one map* **M1** *Placement:* left page, top, bled-off *Size:* margin column (smallest map) *Content:* two small overview maps: one world map, one map of Brazil with Yanomami settlement area *Caption:* Yanomami settlement area *Relevance for topic comprehension:* additional information	*one map* **M3** *Placement:* left page, bottom *Size:* two-columned (largest map, approx. 1/3 textbook spread) *Content:* thematic map of land use in the US; map of Germany for comparison *Caption:* land use in the US heartland *Relevance for topic comprehension:* yes

Fig. 2. Image content and placement (Source: author; for the images, see References)

	Text		Photos		Graphics		Maps	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Individual interest in textbook elements (general)	4.26	0.91	3.84	1.27	4.16	0.97	3.34	1.41
Situational interest in textbook elements of spread *A*	4.05	0.96	3.88	0.82	3.43	1.16	3.26	1.15
Situational interest in textbook elements of spread *B*	3.58	1.21	3.45	1.14	3.05	1.21	2.38	1.30
Situational interest in textbook elements of Spread *C*	3.43	0.96	3.27	1.12	3.36	1.10	3.26	1.24
Mean (*A+B+C*)	**3.69**	**1**	**3.53**	**1.03**	**3.28**	**1.16**	**2.97**	**1.23**

Fig. 3. Questionnaire evaluation: situational interest vs. individual interest (Source: author)



(a) Mean fixation duration (in seconds)

Not fixated (N) = number of participants (out of 58) who did not fixate on the AOI or below the threshold of 150 ms

(b) Pie chart *A*: Number of fixations, percentage distribution of fixations on text, figures, photos, and the map

Fig. 4. Spread *A* (source: author; background image *A*: © Cornelsen, Fischer et al. 2018, pp. 26–27)

in general. Text received the highest ratings and maps the lowest. Participants rated their interest in graphics and photos as medium-high. However, the level of interest in the images differed slightly between the topics.

Step 1c: Fixations on the Textbook Elements of the Stimuli

Participants' fixation duration in seconds, the number of participants who did not fixate on the textbook element, and the proportion of fixations per textbook element served to visualize the mean distribution of participants' visual attention on the stimuli.

The pie charts (Figs. 4–6) show the cumulative number of fixations and their percentage per category (text, figure, photo, maps). Figs. 2–6 visualize participants' observation patterns.

Images and texts are distributed evenly on spread *A* and cover about the same space (each approximately 1.5 columns), but 87 per cent of all fixations were on text.

The largest proportion of fixations was on the text (see Fig. 4b). However, there is a large difference in fixation durations of individual text elements. Participants observed the *check it* box for only 1.68 s and the exercise section for 3.08 s. In all, 21 participants did not observe the *check it* box and the exercise section, whereas all participants observed M1 with a mean fixation duration of 28.3 s.

Participants rated their situational interest in *A*'s photos medium-high (Fig. 3). *A*'s photos depict relevant information (M2: San Francisco earthquake destruction, M5: San Andreas Fault). Nevertheless, the photos received little visual attention. Compared with *C*'s photos, participants observed *A*'s photos longer (M2 1.52 s, M4 1.01 s).

Participants paid little attention to the map despite its prominent placement and relevant content. Fig. 3 shows that the map in *A* received the lowest rating for situational interest among *A*'s elements, but still a medium rating.

Differences in fixation duration and the number of participants who skipped the graph are apparent between figures M5 (4.82 s, not fixed 10) and figure M7 (2.69 s, not fixed 13). The content of the two figures was differently relevant for content comprehension. While M5 explains the origin of earthquakes, M7 offers additional information on how a seismograph works.

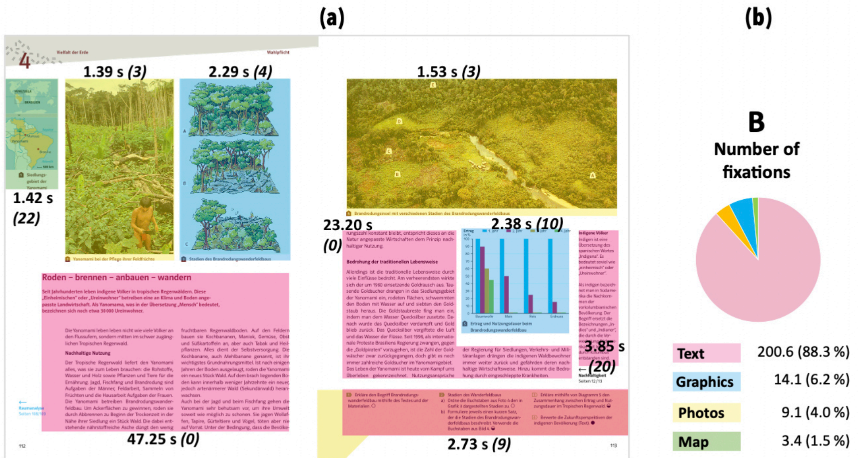
Spread *B* displays images at the top, while text is placed at the bottom of the spread. Like *A*, images and text each occupy about 50 per cent of the print space. Text (88%) shows the highest proportion of fixations among the three stimuli. Here, too, participants' fixations to text differed according to text type. Twenty participants skipped the margin and nine the exercises.

Among the stimuli, *B*'s map was observed most briefly (1.42 s) and skipped most frequently (22). The proportion of fixations on photos was low (4%), the mean fixation duration was short (M2 1.39 s, M4 1.53 s), and three participants skipped the photos, although situational interest in the photos was in the medium range (Fig. 3).

The graphics received little visual attention. Nevertheless, they received the most extended mean fixation duration among *B*'s visuals. However, situational interest in *B*'s graphics was rated lowest among the stimuli's graphics.

Both image content and visual parameters provide possible explanations for the low visual attention to photos, graphics, and maps in spread *B* (Fig. 2).

The image-text distribution of spread *C* differs from that of *A* and *B*. In spread *C*, images occupy approximately two-thirds of the print space and text only one-third. All photos are placed at the top, while graphics and the map are at the bottom. Continuous text is on the left page, while materials and exercises are arranged on

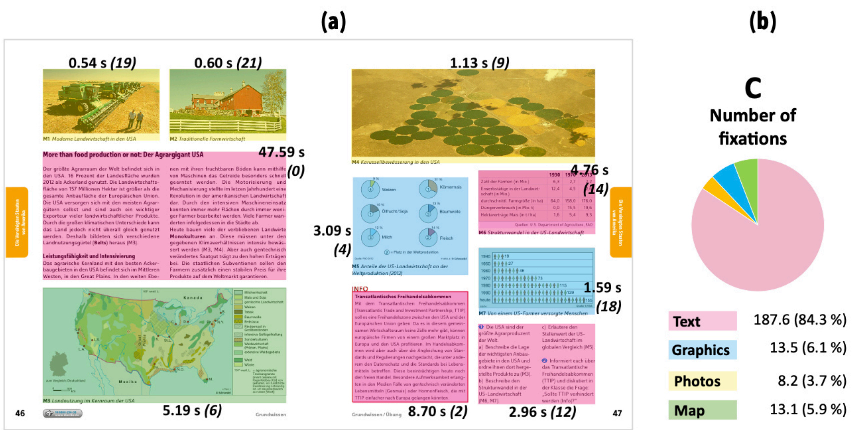


(a) Mean fixation duration (in seconds)

Not fixated (N) = number of participants (out of 58) who did not fixate on the AOI or below the threshold of 150 ms

(b) Pie chart *B*: Number of fixations, percentage distribution of fixations on text, figures, photos, and the map

Fig. 5. Spread *B* (source: author; background image: *B*: © Klett ,Barricelli et al. 2017, pp. 112–113)



(a) Mean fixation duration (in seconds)

Not fixated (N) = number of participants (out of 58) who did not fixate on the AOI or below the threshold of 150 ms

(b) Pie chart *C*: Number of fixations, percentage distribution of fixations on text, figures, photos, and the map

Fig. 6. Spread *C* (source: author; background image: *C*: © Westermann Gruppe Fleischhauer et al. 2015, pp. 46–47)

the right page (Fig. 10). Also, in *C*, text received the most fixations (84.3%) but fewer compared to *A* and *B*. *C* contains most photos, however, mean fixation duration on photos was the shortest and the number of participants who skipped photos was the highest among the stimuli. Learners rated their situational interest at 3.27 (Fig. 3), which was the lowest for situational interest in photos. Here, again, the picture content lacks relevance (Fig. 2), offering a possible explanation for the scant visual attention.

C's map received more visual attention than the maps in *B* and *A*. Nevertheless, six participants skipped the map and it obtained only a medium rating for situational interest (Fig. 3), although it is the largest of the presented maps and depicts relevant information (Fig. 2). A possible explanation may be a lack of competencies in decoding pictorial information.

The graphics received different visual attention: M5 was observed for 3.9 s and skipped by four participants, while M7 was observed for 1.59 s and ignored by 18 participants. Possible reasons may be the relevance of the content and design aspects (Fig. 2).

In summary, participants focused on continuous text. Images received less attention. Differences in participants' visual attention to photos, graphics, and maps can be partly explained by learners' interests (Fig. 3), visual competencies, image content, and visual design parameters (Fig. 2). More precise statements require statistical analysis.

Step 2a: Exploratory Factor Analysis (Figs. 7 & 8)
The exploratory factor analysis investigated and structured interdependencies between 20 identified variables.

A principal component analysis (PCA) with Varimax rotation and Kaiser normalization served as an extraction method to investigate relationships amongst the observed variables. The variables were clustered into seven factor groups. Factor loadings > ± 0.5 were considered relevant and assigned to the respective factor (cf. Backhaus et al. 2016).

Rotated Factor Matrix[a]	Factors						
	1	2	3	4	5	6	7
1 Usability: usefulness	*.867*	.085	.010	-.008	-.016	.083	.009
2 Usability: helpfulness	*.831*	.008	-.022	.011	.113	.264	.228
3 Usability: situational interest	*.785*	.023	-.049	.028	.419	-.031	-.029
4 Usability: comprehensibility	*.779*	-.050	.085	-.153	.179	.137	.193
5 Individual Preference: maps	*.567*	-.131	.009	.300	-.024	.301	.120
6 Individual Interest: *C*	.452	.097	.087	.031	.037	-.024	-.039
7 Text Fixations: *B*	.037	*.920*	.148	-.012	.031	-.261	-.003
8 Text Fixations: *A*	.042	*.919*	-.016	-.021	.165	-.027	.103
9 Text Fixations: *C*	.021	*.836*	.262	.072	.010	.010	-.098
10 Image Fixations: *A*	.137	.140	*.753*	.013	-.034	.118	-.118
11 Image Fixations: *B*	.100	.322	*.615*	-.034	-.307	.151	-.023
12 Individual Interest: *B*	.338	.114	-.570	.242	-.113	-.025	-.011
13 Image Fixations: *C*	.142	.369	.441	.131	.007	-.029	-.064
14 Individual Interest: Geography	.023	.110	.014	*.894*	.034	.077	-.051
15 Individual Interest: *A*	.084	.295	-.192	.377	.302	.110	.248
16 Usability: aesthetics	*.548*	.118	-.167	.082	*.642*	.029	.108
17 Usability: orientation	*.525*	.202	.030	.012	*.629*	-.031	-.005
18 Individual Preference: figures	.287	-.121	.130	.080	.088	*.730*	-.103
20 Individual Preference: text	-.036	.109	-.111	.071	.112	*-.598*	-.392
21 Individual Preference: photos	.220	.058	-.150	.078	.077	.100	*.721*

Note: a = The rotation is converged in nine iterations. Relevant factor loadings are highlighted in italics.

Fig. 7. Results of the exploratory factor analysis (Source: author)

	Initial Eigenvalues			Sum of Load Squares after Extraction			Sum of Load Squares after Rotation		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.270	25.097	25.097	4.982	23.723	23.723	4.079	19.426	19.426
2	3.450	16.427	41.524	3.209	15.279	39.001	2.951	14.053	33.479
3	2.360	11.240	52.764	1.993	9.492	48.493	1.695	8.070	41.549
4	1.778	8.468	61.232	1.445	6.882	55.375	1.464	6.972	48.521

Note: Only factors with eigenvalues < 1 are included in the table.

Fig. 8. Total explained variance and eigenvalues (Source: author)

The quality of the item representation by the factor groups tested a factor matrix and a scree plot. Only factors with eigenvalues >1 were considered significant (Kaiser criterion). Thus, four factor groups (4.079, 2.951, 1.695, 1.464) are relevant (Fig. 8), altogether explaining 48.52 per cent of the variance.

From the eigenvalues, a first ranking of factors' relevance can be deduced (Cleff 2015). The largest explanatory percentage (19.43%) is represented by factor 1, which covers the usability parameters usefulness, helpfulness, situational interest, comprehensibility, aesthetics, and orientation (Fig. 7).

However, exploratory factor analysis does not provide exact statistical correlation tests between factors. Therefore, confirmatory factor analysis was performed.

Step 2b: Confirmatory Factor Analysis (Fig. 9)

A confirmatory factor analysis (principal component analysis with Varimax rotation and Kaiser normalization) tested and confirmed the four extracted factor groups with eigenvalues >1 (Fig. 8).

Rotated Factor Matrix[a]	Factors			
	1	2	3	4
1 Usability: situational interest	*.893*	.066	-.114	.014
2 Usability: helpfulness	*.811*	-.124	.132	.340
3 Usability: comprehensibility	*.806*	-.118	.140	.086
4 Usability: usefulness	*.783*	.005	.149	.114
5 Usability: aesthetics	*.721*	.159	-.273	.202
6 Usability: orientation	*.697*	.274	-.116	.047
7 Text Fixations: *B*	.043	*.949*	.114	.012
8 Text Fixations: *A*	.109	*.842*	.034	.199
9 Text Fixations: *C*	.021	*.812*	.328	.108
10 Image Fixations: *B*	-.002	.248	*.746*	-.034
11 Image Fixations: *A*	.117	.160	*.689*	-.113
12 Image Fixations: *C*	.123	.402	.409	.016
13 Image Interest: *B*	.237	.030	-.357	.272
14 Image Interest: *A*	.125	.279	-.225	*.631*
15 Image Interest: *C*	.425	.090	.106	.021
16 Individual Interest: Geography	-.005	.130	.002	.402
17 Individual Preference: text	-.040	.311	-.326	-.351
18 Individual Preference: maps	.485	-.240	.170	.438
19 Individual Preference: figures	.303	-.290	.356	.306
20 Individual Preference: photos	.240	-.034	-.100	.415

Note: a = The rotation is converged in ten iterations. Relevant factor loadings are highlighted in italics.

Fig. 9. Results of the confirmatory factor analysis (Source: author)

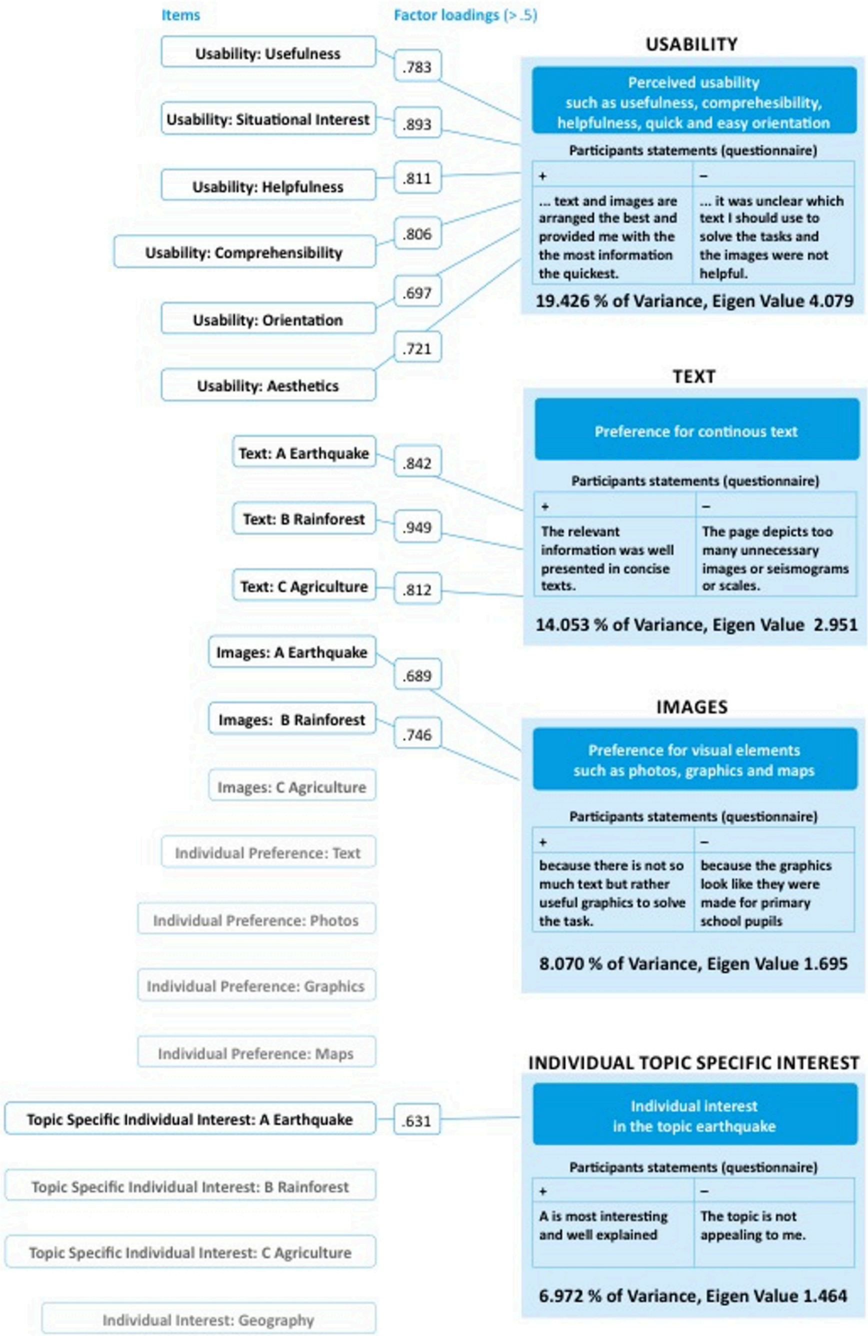


Fig. 10. Factors influencing participants' visual attention to photos, graphics, and maps when observing spreads *A–C* (Source: author)

4.2 How can Influencing Factors for Visual Attention on Continuous Text and Discontinuous Text be Structured, Summarized and Characterized? (Q2)

Step 2c: Factors that Influence Visual Attention when Observing the Stimuli

Based on the results of the confirmatory factor analysis (Fig. 9), Fig. 10 visualizes the classification of the four identified factor groups and the factors included. The group classification is supplemented by participants' comments on the stimuli from the questionnaire that are characteristic of the respective factor group. The classification includes data from 57 participants. Data from one participant could not be included.

Four factor groups with varying relevance were extracted based on a factor loading >0.5 and eigenvalue >1. Among the observed variables, 12 loaded with a factor loading >0.5 on one of the four extracted factor groups.

Based on examples of scan paths, exemplary observational patterns are presented in line with the four extracted factor groups. However, eye movements are highly individual and may be affected by various determinants. Therefore, conclusions about participants' observation patterns should not be

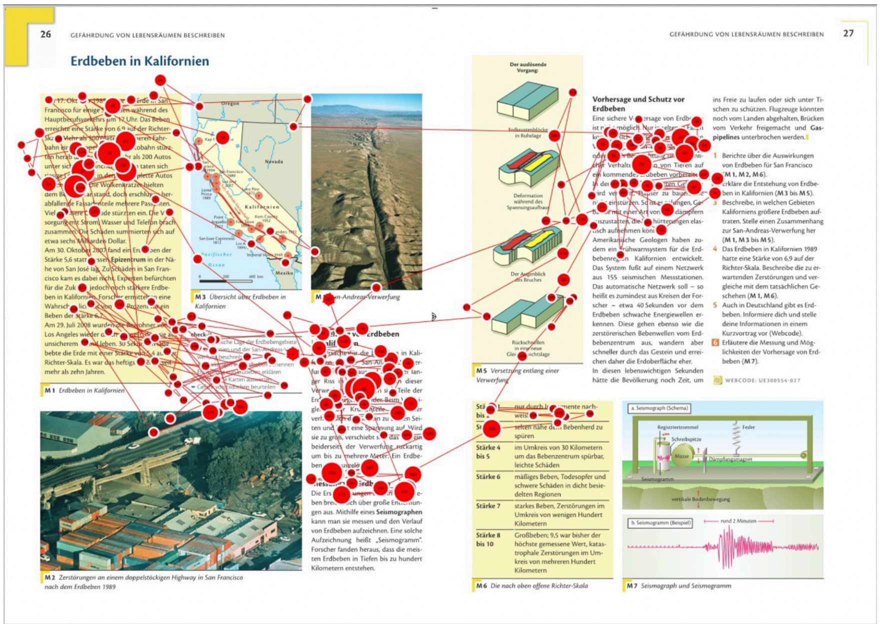


Fig. 11. Factor group usability (exemplary scan path) (Source: author; background image: © Cornelsen, Fischer et. al. 2018, pp. 26-27)

drawn solely based on scan paths but should be considered in the context of further collected data. Hence, it is difficult to define one universal eye movement pattern per factor. Despite the differences, however, commonalities can be identified within a factor group. Statements by the participants (open questions from the questionnaire) complement the examples. Overall, the larger the red dot, the longer the fixation duration. The red lines represent saccades.

Usability

All six usability parameters (Fig. 1) load >0.5. Therefore, usefulness (0.783), helpfulness (0.811), comprehensibility (0.806), aesthetics (0.721), orientation (0.697), and situational interest (0.893) are integrated into the factor usability, explaining the largest share of the variance (19.4%) and obtaining the highest eigenvalue (4.08) among the four factors. Consequently, usability is the most relevant among the examined factors (Fig. 8, Fig. 11).

Responses to two open questions in the questionnaire supported this. Here, participants reasoned which of the three textbook spreads they assessed as best and as worst for their learning. The justifications mainly related to usability, such as an element having been (not) useful or (not) helpful in solving the task, or easy to understand/difficult to understand.

Although the scan paths were heterogeneous, they did have commonalities. Though the visual focus was on the text, task-relevant passages were selected. Likewise, as reflected in the participants' comments, photos, graphics, and the map were fixated according to perceived usability.

> The text provided information, but also the graphics helped to understand the topic and to solve the task (P11); I immediately grasped the topic, the figure explained what the text described (P01) (originals in German; author's translations).

Text

Although the factor loads were high (*B* 0.949, *A* 0.842, *C* 0.812), their eigenvalue (2.951) and explained variance (14.053%) indicated lower relevance of text than of usability (Fig. 12).

Participants frequently fixated the entire continuous text, although they were not explicitly asked to read the text carefully.

Despite differences, the scan paths were more homogeneous within the text factor than in usability. Two main patterns were identified. Either the entire text (including irrelevant sections such as exercises) were fixated and visuals were considered (albeit superficially), or the focus was on continuous text and little or no attention was paid to visuals and materials.

> Information relevant to the task was well presented in concise texts (P09).
>
> The individual text paragraphs are separated, and keywords are highlighted (P010).

Images

Only *A* (0.689) and *B* (0.746) loaded >0.5; *C*, with 0.409, was not included. Nevertheless, 8.1 per cent of the variance can be explained by participants' preference for images (Fig. 13).

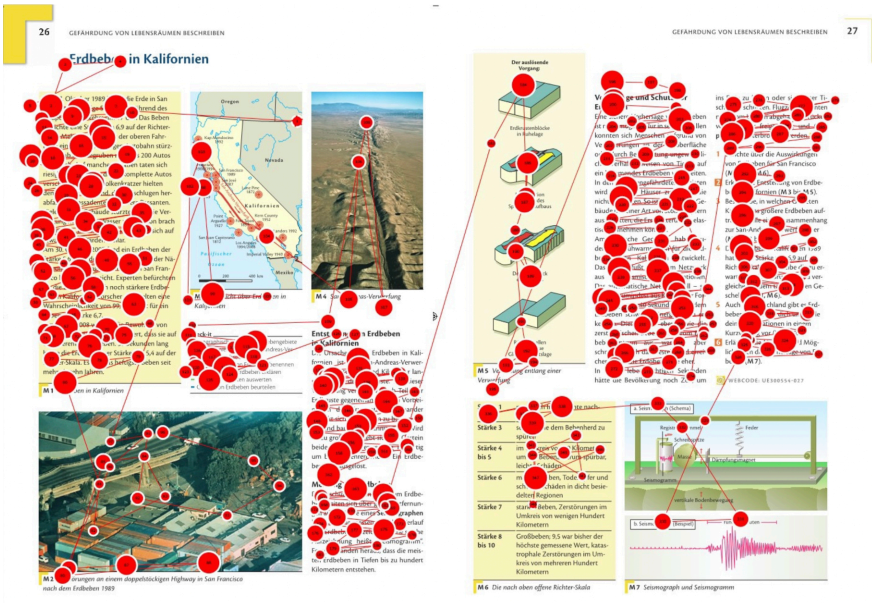


Fig. 12. Factor group text (exemplary scan path) (Source: author; background image © Cornelsen, Fischer et al. 2018, pp. 26-27)

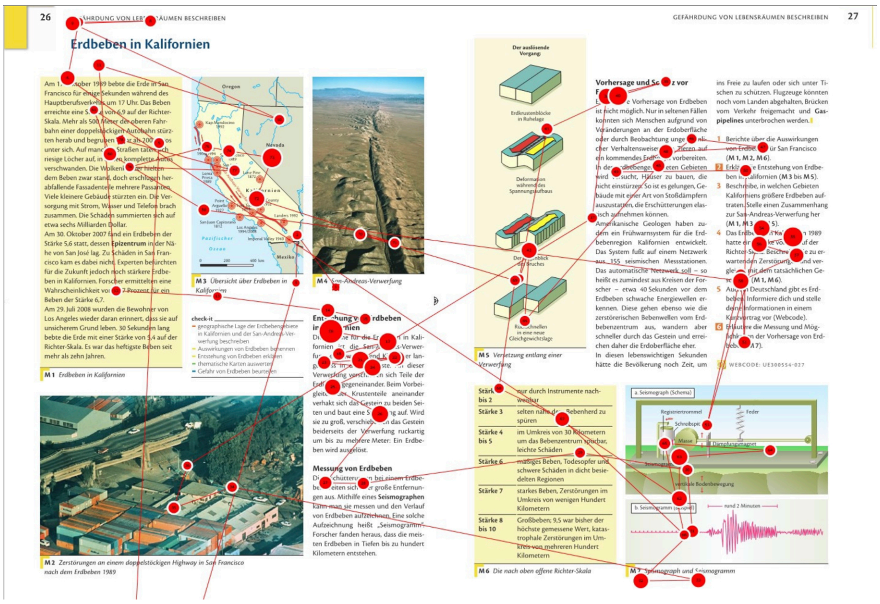


Fig. 13. Factor group images (exemplary scan path) (Source: author; background image © Cornelsen, Fischer et al. 2018, pp. 26-27)

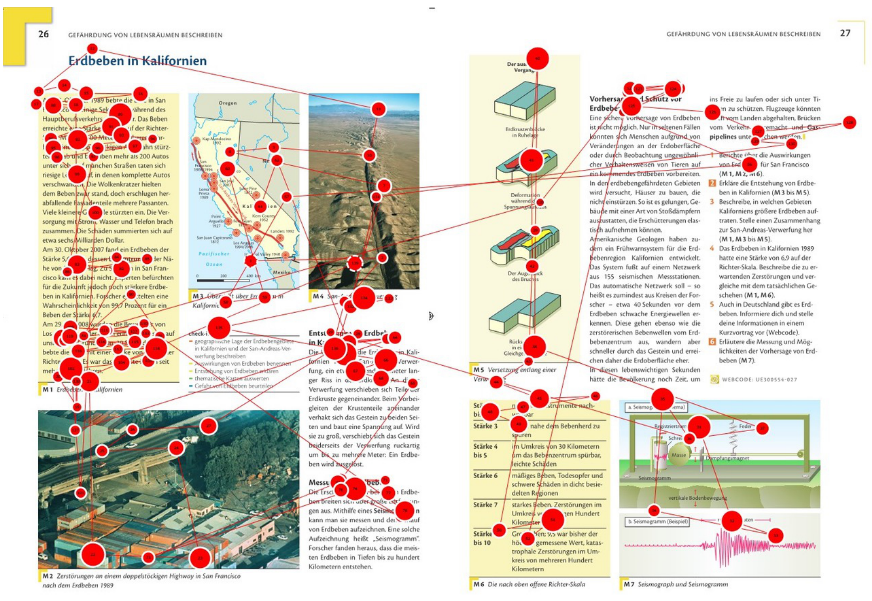


Fig. 14. Factor group individual interest (exemplary scan path) (Source: author; background image © Cornelsen, Fischer et al. 2018, pp. 26-27)

The scan path is characterized by attention to the text but not more attention to visuals than the other factor groups.

The limited attention given to text may be due to challenges in text comprehension and/or alleged easy picture understanding (illusion of full understanding, Peeck 1993).

> On *A*, many different aspects of earthquakes are illustrated with examples in pictures (P34).

Individual interest (in the topic of earthquakes)
Only *A* (0.631) was included in the factor because *B* (0.272) and *C* (0.016) obtained factor loads <0.5. The lowest explained variance of 7.0 per cent and the small eigenvalue (1.5) indicated limited relevance compared to usability, text, and images. Within the interest factor, the distribution of visual attention was heterogeneous. However, commonalities can be found within the factor.

In addition to information relevant to content comprehension, the scan path depicts a focus on topic-specific details. Here, too, fixations are mainly on text but also along the fault line of the San Andreas Fault (M4), on earthquake damage on the highway (M2), or on details explaining how the seismograph works.

> I generally liked *A* the most, because of the picture of the street, because that really happened (P44).
> Spread *A* is the most interesting and well explained with the help of the pictures and figures (P73).

In summary, participants' scan paths combined with their comments provide first insights into how and why participants observed or ignored photos, graphics, and maps depicted in the stimuli.

5. Discussion

As in the exploratory study (Behnke 2016b), participants mainly focused on text. However, they selected among text types and focused on continuous text, while text materials were largely skipped (Figs. 4–6). This suggests that, besides a general preference for text, learners prioritize their visual attention according to relevance and usefulness, which supports motivational theories, such as ARCS (Keller 2010) and SDT (Ryan & Deci 2000).

Overall, the data revealed a discrepancy between medium to medium-high situational interest in photos, graphics, and maps (Fig. 3) of the stimuli and participants' low fixation rate on the stimuli's images. There were differences in the number of participants who skipped photos, graphics, and maps–both within each stimulus and between the stimuli, and in fixation durations (Fig. 5, Fig. 6, Fig. 10).

Image parameters of the stimuli (Fig. 2) displayed critical aspects, such as decorative photos, inadequate information design, or adverse size and placement. Other causes may be redundancies, such as an irrelevant image, as information is also contained in the continuous text (*A*, *B*, *C*), or duplication of the image information (*B*). This reduces the usability of depicted images.

Empirical evidence on multimedia learning reveals that in the case of redundant visual and textual information, learners focus on text and may miss relevant image details (Schnotz et al. 2014), especially if image inclusion is not explicitly demanded and learners are free to decide which representation to choose.

The factor analysis identified four relevant factors determining participants' visual attention to photos, graphics, and maps depicted on the stimuli *A*, *B*, and *C*. Usability, text preference, individual interest in the topic of earthquakes, and image preference explain most of the variance with different degrees of influence. Among the factors, usability had the largest share and is thus considered the most relevant factor (amongst the examined factors) for participants' attention to textbook visuals in the stimuli.

This study illustrates the relevance of usability, but also the inherent potential in optimizing usability parameters, because individual perceived usefulness may be one aspect that controls whether or not learners observe the image. Equally, as argued by Knight et al. (2017), usability is an essential criterion for students' attitude towards a learning environment.

Individual interest in the topic of earthquakes was relevant but explained the smallest proportion of variance. A general preference for text in learning media also affected learners' visual attention, but with a smaller explanatory share than usability. Although extracted as a relevant factor, image preference explained the second smallest share of the variance (Fig. 10).

In summary, reasons for differences in participants' visual attention to depicted visuals can be derived from the individuals' perceived usability of the design and content in relation to learners' goals rather than from individual interest in the topic or media.

6. Conclusions and Future Research

Usability, with the attributes usefulness, comprehensibility, aesthetics, orientation, helpfulness, and situational interest, explains the largest proportion of the variance (Fig. 10) and is therefore considered, among the examined factors, the most relevant influencing factor for visual attention to images in the stimuli. The factor analysis thus confirms the underlying theoretical model (Fig. 1).

Moreover, the study demonstrates that participants considered visuals to be relevant for their learning (Fig. 3), even though their fixation percentage was around 15 per cent (Figs. 4–6) on the stimuli. The data show that learners assess the personal benefit of visuals and orientate their attention and information selection on whether they are helpful, useful, and/or easily accessible.

Consequently, attention to visuals may be promoted by optimizing their usability. Optimal usability of textbook images supports the recognition, processing, and understanding of visual information (Holsanova 2014a, 2014b). The prerequisite is that visuals are meaningfully applied didactically in relation to the learning topic and learning objective and designed appropriately for the target group rather than being purely decorative. In addition, successful image-text integration can be facilitated by support systems, such as cueing (Schneider et al. 2018) and scaffolding (Eitel et al. 2013).

However, successful knowledge construction with visuals requires competencies to access, identify, decode, and utilize pictorial information (Rice & Dallacqua 2019), such as (geographical) picture reading competencies (Jahnke 2012), visual literacy (Avgerinou 2009), and graphicacy (de Vries & Lowe 2010). If these skills are not sufficiently developed, images may also be skipped. This demonstrates the potential of developing visual competencies.

Regarding learners' interests in the images of the three stimuli, a distinction must be made between the usability parameter of situational interest and the factor group individual interest. Situational interest loads highest among the usability factors, which indicates its relevance. Individual interest in earthquakes (spread *A*) explains the smallest share of the variance. However, it is still one of four relevant factors. Studies on learners' interests (Hemmer & Hemmer, 2010; Hemmer et al. 2020) reported that earthquakes are among the most popular geographical topics. This indicates that individual interest, although a small factor, affects visual attention, particularly when a topic is popular, confirming Krapp's (2002) POI, which stresses the influence of motivational aspects on learners' interests.

Personal relevance, concreteness, and ease of comprehension may trigger situational interest in textbook images (Magner et al. 2016), thus indicating that situational interest is moderated by usability factors, such as usefulness, aesthetics, and comprehensibility.

Consequently, situational interest in a particular image triggers potentially more visual attention than individual interest in a geographic topic or medium. This implies that interest and attention are interrelated, and learners' situational interest in photos, graphics, and maps can be promoted through the use of didactically and visually well-designed learning media (Krapp 2002; Magner et al. 2016).

In summary, the study highlights the critical role of usability for textbook visuals for learning, and the need for further research in this area. For this, Rose (2012) suggests methodological approaches for research with visual materials. However, visual attention is complex and may be influenced by various factors, such as task, prior knowledge, and learning environment.

Finally, limitations need to be mentioned. Participants' age was homogeneous (M_{age} 15.26 years, SD_{age} 0.80). Visual information selection strategies may differ with age (Boucheix et al. 2015; Ouwehand et al. 2016; Schnotz et al. 2017). Indeed, a larger sample size would allow more conclusive factor analysis.

Further factors should be investigated, such as tasks, prior knowledge, and different media. Interrelations between usability parameters should be examined to determine which (combinations of) usability parameters are most effective in triggering learners' attention, motivation, and knowledge construction. Moreover, using prototyping in real classroom situations to investigate how optimized visualizations affect visual attention, may also provide valuable insights.

In summary, textbook images are still underestimated, both in their learning potential and in their challenges for learners. Knowledge about learners' visual preferences and their information selection strategies is helpful to promote knowledge construction with images. To promote knowledge construction with images in educational media, learners' visual attention should be guided and supported. Image content in learning media needs to be useful, relevant, and accessible to learners to achieve this. Strengthening learners' visual skills is equally essential.

References

Avgerinou, M. D. (2009). Re-Viewing Visual Literacy in the "Bain d'Images" Era. *TechTrends*, *53*, 28–34.

Baadte, C., & Schnotz, W. (2012). Das Verstehen von Texten mit Bildern. *Weiterbildung*, *4*, 35–37.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2016). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. SpringerGabler.

Backhaus, K., Erichson, B., & Weiber, R. (2015). *Fortgeschrittene Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (Issue 3). SpringerGabler.

Bagoly-Simó, P., Hartmann, J., & Reinke, V. (2019). School Geography under COVID-19: Geographical Knowledge in the German Formal Education. *Tijdschrift voor economische en Sociale geografie. Journal of Economic and Social Geography*, *111*(3), 224–238.

Ballstaedt, S.-P. (2017). Text und Bild: ein didaktisches Traumpaar. *Bildendes Sehen*, *January 2017*, 45–55.

Behnke, Y. (2016a). How Textbook Design May Influence Learning with Geography Textbooks. *Nordidactica – Journal of Humanities and Social Science Education*, *1*, 38–62.

Behnke, Y. (2016b). Wie betrachten und bewerten Lernende Geographieschulbuchseiten? How Do Learners Observe and Assess Geography Textbook Spreads? *Zeitschrift für Geographiedidaktik | Journal of Geography Education (ZGD)*, *44*(3), 5–34.

Behnke, Y. (2017). Learning-Related Challenges Presented by Graphic Visualisations in Geography Textbooks. In F. Dervin, L. A. Kulbrandstad & H. Ragnarsdóttir (Eds.), *Changing Subjects, Changing Pedagogies: Diversities in School and Education* (pp. 68–89). Finnish Research Association for Subject Didactics.

Behnke, Y. (2021). Usability Qualities of 'Well-designed' Geography Textbook Visuals. *Journal of Visual Literacy*, *40*(1), 15–33.

Bischof, W. F., Anderson, N. C., & Kingstone, A. (2019). Temporal Methods for Eye Movement Analysis. In C. Klein & U. Ettinger (Eds.), *Eye Movement Research. An Introduction to its Scientific Foundations and Applications* (pp. 407–448). Springer Nature Switzerland.

Bojko, A. (2014). *Eye-tracking the User Experience*. Rosenfeld.

Boucheix, J., Lowe, R. K., & Bugaiska, A. (2015). Age Differences in Learning from Instructional Animations. *Applied Cognitive Psychology*, *29*(4), 524–535.

Carrasco, M. (2011). Visual Attention: The Past 25 Years. *Vision Research*, *51*(13), 1484–1525.

Cleff, T. (2015). *Deskriptive Statistik Explorative Datenanalyse. Eine computergestützte Einführung mit Excel, Spss und Stata*. SpringerGabler.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention Guidance in Learning from a Complex Animation: Seeing is Understanding? *Learning and Instruction*, *20*(2), 111–122.

De Lucio, J. V., Mohamadian, M., Ruiz, J. P., Banayas, J., & Bernaldez, F. G. (1996).

Visual Landscape Exploration as Revealed by Eye Movement Tracking. *Landscape and Urban Planning*, *34*(2), 135–142.

de Vries, E., & Lowe, R. K. (2010). Graphicacy: What Does the Learner Bring to a Graphic? In K. Scheiter (Ed.), *SIG2–2010: Comprehension of Text and Graphics. Tracing the Mind: How Do We Learn from Text and Graphics? Proceedings of the EARLI SIG Meeting 2010, Tübingen, August 26–28* (pp. 26–28).

Duchowski, A. (2007). *Eye Tracking Methodology. Theory and Practice*. Springer.

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a Picture Facilitates the Process of Learning from Text: Evidence for Scaffolding. *Learning and Instruction*, *28*, 48–63.

Emhardt, S., Wermeskerken, M., Scheiter, K., & Gog, T. (2020). Inferring Task Performance and Confidence from Displays of Eye Movements. *Applied Cognitive Psychology*, *34*(6), 1430–1443.

Endres, T., Weyreter, S., Renkl, A., & Eitel, A. (2020). When and why Does Emotional Design Foster Learning? Evidence for Situational Interest as a Mediator of Increased Persistence. *Journal of Computer Assisted Learning, 36*(4), 514–525.

Fuchs, E., & Bock, A. (2018). Introduction. In E. Fuchs & A. Bock (Eds.), *The Palgrave Handbook of Textbook Studies* (pp. 1–9). Palgrave Macmillian.

Guo, D., Zimmer, W., Matthews, S. D., & McTigue, E. M. (2019). Critical Analysis of Research on the Impact of Visual Literacy for Learning: Strengths, Weaknesses and Recommendations for Improvement. *Journal of Visual Literacy*, *38*(3), 181–198.

Hemmer, I., & Hemmer, M. (2010). Interesse von Schülerinnen und Schülern an einzelnenThemen, Regionen und Arbeitsweisen des Geographieunterrichts – ein Vergleich zweier empirischerStudien aus den Jahren 1995 und 2005. In I. Hemmer & M. Hemmer (Eds.), *Schülerinteresse an Themen, Regionen und Arbeitsweisen des Geographieunterrichts. Ergebnisse der empirischen Forschung und deren Konsequenzen für die Unterrichtspraxis. Geographiedidaktische Forschungen (Band 46)* (S. 65–145). HGD.

Hemmer, M., Hemmer, I., Warnke, M., & Zietzsch, P. (2020). Welche Themen interessieren Schülerinnen und Schüler des Gymnasiums im Geographieunterricht? Ausgewählte Ergebnisse einer empirischen Untersuchung in Nordrhein-Westfalen. *Der Schulgeograph*, *1*(53), 30–34.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Weijer, J. van De (2010). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press.

Holsanova, J. (2014a). In the Eye of the Beholder: Visual Communication from a Recipient Perspective. In D. Machin (Ed.), *Visual Communication. Handbook of Communication Science* (pp. 331–355). De Gruyter.

Holsanova, J. (2014b). Reception of Multimodality: Applying Eye-tracking Methodology in Multimodal Research. In C. Jewitt (Ed.), *Routledge Handbook of Multimodal Analysis* (pp. 285–296). Routledge.

Horz, H., & Schnotz, W. (2009). Multimedialer Wissenerwerb. In E. Wild & J. Möller (Eds.), *Pädagogische Psychologie* (pp. 188–192). Spinger.

Hyönä, J. (2010). The Use of Eye Movements in the Study of Multimedia Learning. *Learning and Instruction*, *20*(2), 172–176.

Jahnke, H. (2012). Geographische Bildkompetenz. *Geographie und Schule*, *36*(195), 27–34.

Keller, J. M. (2010). *Motivational Design for Learning and Performance. The ARCS Model Approach*. Springer.

Klein, C., & Ettinger, U. (2019). *Eye Movement Research. An Introduction to its Scientific Foundations and Applications*. Springer.

Knight, B. A., Casey, M., & Dekkers, J. (2017). Using Electronic Textbooks to Teach Mathematics in the Secondary Classroom: What Do the Students Say? *International Journal of Education and Development Using Information and Communication Technology (IJEDICT), 13*(1), 87–102.

Koonsanit, K., Tsunajima, T., & Nishiuchi, N. (2021). Evaluation of Strong and Weak Signifiers in a Web Interface Using Eye-Tracking Heatmaps and Machine Learning. In K. Saeed & J. Dvorský (Eds.), *Computer Information Systems and Industrial*

Management. CISIM 2021. Lecture Notes in Computer Science, *12883*.

Krapp, A. (1992). Interesse, Lernen und Leistung. Neue Forschungsansätze in der Pädagogischen Psychologie. *Zeitschrift Für Pädagogik*, *38*(5), 747–770.

Krapp, A. (2002). An Educational-psychological Theory of Interest and its Relation to SDT. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 405–430). University of Rochester Press.

Krapp, A., & Prenzel, M. (1992). *Interesse, Lernen, Leistung. Neuere Ansätze der pädagogisch-psychologischen Interessenforschung*. Aschendorff.

Lambert, D. (2008). Introduction – Part 1: What Is Living Geography? In D. Mitchell (Ed.), *Living Geography* (pp. 1–12). Chris Kington.

Loschky, L. C., Hutson, J. P., Smith, M. E., Smith, T. J., & Magliano, J. P. (2018). Viewing Static Visual Narratives through the Lens of the Scene Perception and Event Comprehension Theory (SPECT). In A. Dunst, J. Laubrock & J. Wildfeuer (Eds.), *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods* (pp. 217–238). Routledge.

Lowe, R. K., & Boucheix, J.-M. (2016). Principled Animation Design Improves Comprehension of Complex Dynamics. *Learning and Instruction*, *45*(August), 72–84.

Magner, U., Glogger, I., & Renkl, A. (2016). Which Features Make Illustrations in Multimedia Learning Interesting? *Educational Psychology*, *36*(9), 1596–1613.

Ouwehand, K., van Gog, T., & Paas, F. (2016). Compensatory Effects of Pointing and Predictive Cueing on Age-related Declines in Visuospatial Working Memory. *Memory and Cognition*, *44*(6), 950–965.

Peeck, J. (1993). Increasing Picture Effects in Learning from Illustrated Text. *Learning and Instruction*, *3*(3), 227–238.

Pettersson, R. (1995). The Learning Helix. *Eye on the Future: Converging Images, Ideas, and Instruction. Selected Reading from the Annual Conference of the International Visual Literacy Association (October 18-22, Chicago, 1995)*, 395–407.

Petterson, R. (2000). Attention – An Information Design Perspective. *Document Design*, *2*(2), 114 –130.

Prenzel, M., Krapp, A., & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, *32*(2), 163–173.

Rayner, K. (2009). Eye Movements and Attention in Reading, Scene Perception, and Visual Search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye Movements and Visual Encoding During Scene Perception. *Psychological Science*, *20*(1), 6–10.

Renninger, K. A., Bachrach, J. E., & Hidi, S. E. (2019). Triggering and Maintaining Interest in Early Phases of Interest Development. *Learning, Culture and Social Interaction*, *23*(December 2019), 100260.

Renninger, K. A., & Hidi, S. E. (2016). *The Power of Interest for Motivation and Engagement*. Routledge.

Renninger, S., & Hidi, K. A. (2006). The Four-phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111–127.

Rice, M. F., & Dallacqua, A. K. (2019). Teaching Visual Literacies: The Case of The Great American Dust Bowl. *SANE Journal: Sequential Art Narrative in Education*, *2*(4), Article 1.

Rose, G. (2012). *Visual Methodologies: An Introduction to Researching with Visual Materials*. Sage.

Ryan, R. M., & Deci, E. L. (2000). Self-determination Theory and the Facilitation of Intrinsic Motivation. *American Psychologist*, *55*(1), 68–78.

Sass, S., Schütte, K., & Lindner, M. A. (2017). Test-takers' Eye Movements: Effects of Integration Aids and Types of Graphical Representations. *Computers and Education*, *109*, 85–97.

Sauro, J., & Lewis, J. R. (2016). *Quantifiying the user experience. Practical statistics for user research.* Elsevier Inc.

Scheiter, K., Schubert, C., & Schüler, A. (2018). Self-regulated Learning from Illustrated Text: Eye Movement Modelling to Support Use and Regulation of Cognitive Processes during Learning from Multimedia. *British Journal of Educational Psychology*, *88*(1), 80–94.

Scheiter, K., Schüler, A., Gerjets, P., Stalbovs, K., Schubert, C., Huk, T., & Hesse, F. W.

(2014). Welche Rolle spielt neben Merkmalen des Instruktionsdesigns die fachspezifische und aufgabenspezifische Motivation beim Lernen mit Multimedia im naturwissenschaftlichen Unterricht? *Zeitschrift für Erziehungswissenschaften*, *17*, 279–296.

Schiefele, H., Prenzel, M., Krapp, A., Heiland, A., & Kasten, H. (1983). *Zur Konzeption einer pädagogischen Theorie des Interesses*. Institut für Empirische Pädagogik und Pädagogische Psychologie, Hochschule der Bundeswehr München.

Schiefele, U. (2012). Interests and Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1623–1626). Springer.

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A Meta-analysis of how Signalling Affects Learning with Media. *Educational Research Review*, *23*, 1–24.

Schnotz, W., Fries, S., & Horz, H. (2009). Motivational Aspects of Cognitive Load Theory. In M. Wosnitza, S. A. Karabenick, A. Efklides & P. Nenninger (Eds.), *Contemporary Motivation Research. From Global to Local Perspectives* (pp. 69–96). hogrefe.

Schnotz, W., Mengelkamp, C., Baadte, C., & Hauck, G. (2014). Focus of Attention and Choice of Text Modality in Multimedia Learning. *European Journal of Psychology of Education*, *29*, 483–501.

Schnotz, W., Wagner, I., Ullrich, M., Horz, H., & McElvany, N. (2017). Development of Students' Text-Picture Integration and Reading Competence across Grades 5 – 7 in a Three-tier Secondary School System: A Longitudinal Study. *Contemporary Educational Psychology*, *51*, 152–169.

Seufert, T. (2019). Training for Coherence Formation when Learning from Text and Picture and the Interplay with Learners' Prior Knowledge. *Frontiers in Psychology*, *10*(193), 1–11.

Teschner, M. (2011). *Schulbuch und Schülerinteresse – Zur Bedeutung, Wahrnehmung und Bewertung von Erdkundeschulbüchern aus der Perspektive von Schülerinnen und Schülern*. Leibniz Universität Hannover.

Thomson, B. (2004). *Exploratory and Confirmatory Factor Analysis. Understanding Concepts and Applications*. APA.

Thornes, J. E. (2004). The Visual Turn and Geography (Response to Rose 2003 Intervention). *Antipode*, *36*(5), 787–794.

Uhlenwinkel, A. (2007). Schüler ins Bild setzen. *Praxis Geographie*, *37*(1), 28–31.

Weidenmann, B. (1989). When Good Pictures Fail. An Information-processing Approach to the Effect of Illustrations. In H. Mandl & J. R. Levin (Eds.), *Knowledge Acquisition from Text and Pictures* (pp. 157–170). Elsevier Science Publishers B. V.

Weidenmann, B. (1994). *Wissenserwerb mit Bildern*. Huber.

Widdowson, J., & Lambert, D. (2006). Using Geography Textbooks. In D. Balderstone (Ed.), *Secondary Geography Handbook* (pp. 146–159). Geographical Association.

Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of Emotional Designs in Multimedia Learning: A Replication and Extension Study. *Educational Psychology Review, 33*(2), 357–385.

Zander, S., Wetzel, S., Kühl, T., & Bertel, S. (2017). Underlying Processes of an Inverted Personalization Effect in Multimedia Learning – An Eye-Tracking Study. *Frontiers in Psychology* (December 15), 1–9.

Stimuli

A: Fischer, P. , Flath, M., Jung, L., Maroske, R., Mathesius-Wendt, U., Rudyk, E., & Theissen, U. (2018) *Unsere Erde 2 NRW Gymnasium*. Cornelsen (pp. 26–27).

B: Barricelli, M., Kucharzyk, K., Kusserow, K., Meissner, B. Pflüger, S., & Werner, S. (2017) *Terra Geographie 7/8 Berlin und Brandenburg*. Klett (pp. 112–113).

C: Fleischhauer, T., Franz, S., Gerlach, A., Gutberlet, M., Köhler, P., & Steller, R. (2015). *Seydlitz/Diercke Geographie. Thüringen 9/10*. Westermann (pp. 46–47).

Autorin

✉ **Yvonne Behnke**
Humboldt-Universität zu Berlin
Geographisches Institut/Didaktik der Geographie
Unter den Linden 6
10099 Berlin
kontakt@behnke-design.de