Kunz *et al. Smart Learning Environments (2024) 11:44*
https://doi.org/10.1186/s40561-024-00333-6

Smart Learning Environments

RESEARCH

Open Access

Investigation of students' use of online information in higher education using eye tracking

Ann-Kathrin Kunz[1], Olga Zlatkin-Troitschanskaia[1], Susanne Schmidt[1], Marie-Theres Nagel[1*] and Sebastian Brückner[1]

*Correspondence:
marie.nagel@uni-mainz.de

[1] Department of Law and Economics, Johannes Gutenberg-University Mainz, Jakob-Welder-Weg 9, 55099 Mainz, Germany

Abstract

To successfully learn using freely available (and non-curated) Internet resources, university students need to search for, critically evaluate and select online information, and verify sources (defined as Critical Online Reasoning, COR). Recent research indicates substantial deficits in COR skills among higher education students. To support students in learning how to critically use online information for their learning, it is necessary to better understand the strategies and practices that might elicit less critically-reflective judgments about online information and thus account for such deficits. To this end, using eye tracking data, we investigate how the COR behaviors of students who critically-reflectively evaluate the credibility of online information ('high performers') differ from those of students who do not critically-reflectively evaluate it ('low performers'): 19 students were divided into high and low performers according to their performance in the newly developed Critical Online Reasoning Assessment (CORA). The fixation and dwell times of both groups during CORA task processing were compared regarding time spent on the different processing steps and eye movements on the visited web pages. The results show noticeable differences between the two groups, indicating that low performers indeed approached the task rather heuristically than systematically, and that COR skills require targeted and effective training in higher education.

Keywords: Online information processing, Critical Online Reasoning Assessment, Eye tracking, Response process patterns, Higher education, Two-process theory

Introduction

Research background and study objective

In recent years, eye tracking has been increasingly used in educational research and practice, e.g. to analyze domain-specific understanding and expertise in computer-based assessments (e.g., Han et al., 2017; Klein et al., 2020), to investigate the effectiveness of learning methods (e.g., Lee & Wu, 2017; Luo et al., 2017), and to examine the usability of digital learning environments (e.g., Erdogan et al., 2023) and multimedia learning content (for detailed reviews, see Alemdag & Cagiltay, 2018; Coskun & Cagiltay, 2021). Eye-tracking studies generally use closed formats (ready-made websites or interaction interfaces) to research how students deal with (preselected) learning content (e.g., Sharma et al., 2020; Ben Kheder et al., 2018; Navarro et al., 2015). So far however, little is known about how students interact with freely accessible information on the Internet (e.g., Schmidt et al., 2020), despite online information from websites being increasingly used in both formal and informal learning contexts to acquire knowledge and achieve learning goals (Gadiraju et al., 2018; Maurer et al., 2020). To address this research gap, we aim to investigate how students interact in an *open* assessment environment with unrestricted Internet search access. This corresponds to real learning environment, in which students navigate real websites on the Internet.

The almost unlimited access to information poses challenges, since there is no guarantee regarding the quality of information found on the Internet (Gerjets et al., 2011). The Internet is characterized by a high heterogeneity of information and a plethora of sources that differ considerably in terms of the expertise and (hidden) interests of their originators (Metzger, 2007). For this reason, there is a risk of integrating incorrect or erroneous information into the learning process (Kahne et al., 2016), which can lead to faulty mental models in domain learning (in a university context; Zlatkin-Troitschanskaia et al., 2019). A critically-reflective approach when using online information is therefore essential. Students should be able to judge online information based on evidence-based arguments rather than superficial characteristics (of websites) or personal beliefs (McGrew et al., 2018). This requires effectively searching for information, evaluating that information for credibility, and verifying it by consulting other sources (Brand-Gruwel et al., 2009; Britt & Rouet, 2012; McGrew et al., 2018). Recent research indicates substantial deficits among higher education students in these skills (e.g., McGrew et al., 2019; Walraven et al., 2009; Wineburg et al., 2018). Accordingly, young adults should consistently be supported in learning how to use online information critically (Kahne et al., 2016). To meet this demand in higher education, it is necessary to better understand the practices that might elicit less critically-reflective judgments about online information and thus account for deficits among higher education students.

Research focus and questions

Previous studies show that students make use of cognitive heuristics to assess online information (Flanagin & Metzger, 2007; Walraven et al., 2009; Zhang, Cole & Belkin, 2011). According to the two-process theory (Evans, 2006), heuristics are automated processes that are mostly experience-based and require little cognitive effort (Gronchi & Giovannelli, 2018; Horstmann et al., 2009). In contrast, systematic processes require a higher cognitive effort. It is assumed that heuristics are less likely to lead to critically reflective judgments (Evans, 2006; Flanagin & Metzger, 2007). So far, students' critically reflective judgments regarding the credibility of online information are often assessed using only "self-report-based approaches" (List & Alexander, 2018a, p.199), which do not necessarily provide information about students' actual approaches (Fogg et al., 2002). Therefore, in this paper, we focus on the following research question:

RQ 1: How do the actual behaviors of students who critically-reflectively evaluate the credibility of online information ('high performers') differ from those of students who do not critically-reflectively evaluate it ('low performers')?

A direct, real-time access to students' information processing is supposed to give more comprehensive insights. Increasingly, studies on Internet behavior use eye tracking (ET) for such access to gain insights into information processing based on eye movements during confrontation with websites (Orquin & Mueller Loose, 2013; Rayner, 1998). Fundamental assumptions that qualify eye movements as indicators of processing are the eye-mind and immediacy theories (Just & Carpenter, 1976, 1980), which assume that there is a close relationship between the fixation of objects (incl. objects on a screen) with the eyes and the cognitive processing of these objects. However, Holmqvist et al. (2011) emphasize that eye movement data must be embedded in a research context so that they can be interpreted meaningfully. Thus, the following RQ arises:

RQ 2: How can differences in the approaches to online information of high-performing students compared to low-performing students be operationalized through eye tracking data?

Since the focus of cognitive information processing is on attention-related processes (Orquin & Mueller Loose, 2013), one of the best-known ET measures is fixation. Fixations are stable, minimal eye movements within an area that occur when an individual maintains their gaze on an object of interest (Duchowski, 2007). According to Holmqvist et al. (2011), they are well studied as indicators of processing depth. In particular, the study of the length of fixations is widely used in research practice. Attentional ET analysis also examines fixations within specific areas of materials used in assessment (Bera, Soffer, & Parsons, 2019) referred to as Areas of Interest (AOIs). By defining them, additional ET measures can be determined (Holmqvist et al., 2011). The dwell time and the length of fixation on an AOI are commonly used to operationalize the overall processing of these areas (Gerjets et al., 2011; Raney et al., 2014). Based on the selection of these ET measures for operationalization, the following RQ is examined:

RQ 3: Which differences can be identified in terms of the length of identified process steps and the length of fixations on AOIs between low performers and high performers?

Theoretical and conceptual framework

Students information processing strategies

Numerous studies show that students use simplified heuristics rather than systematic procedures to assess the credibility of online information (e.g., Barzilai & Zohar, 2012; Iding et al., 2009; Metzger & Flanagin, 2013; Sundar, 2008; Walraven et al., 2009; Winter & Krämer, 2014; for an overview, see Zlatkin-Troitschanskaia et al., 2021a, 2021b). In doing so, they often unconsciously allow their judgments to be misguided not by objective criteria but by superficial characteristics that are of little or no relevance to the credibility of information and information sources (Barzilai & Zohar, 2012; Metzger et al., 2010). Simplistic inferences are already made when searching for information by interpreting the order of search results as a signal of credibility (Gerjets et al., 2011; Walraven et al., 2009; Zhang et al., 2011). On websites, judgments about credibility are often made based on superficial external features. Fogg (2003) and Wathen and Burkell (2002) found that an (initial) judgment about the credibility of an online source is primarily made based on the site presentation, i.e., visual design elements such as images or the color scheme. Metzger et al. (2010) summarize the following findings from research: "(…) information seekers are likely to cope with the perceived costs of information search and overload by seeking out strategies that minimize their cognitive effort and time through the use of cognitive heuristics."

Walraven et al. (2009) examined students' evaluation criteria using think-aloud protocols and found, first, that few of them consciously used any criteria at all. Second, criteria for evaluating online information that students considered useful in previous interviews were not used in practice. Flanagin and Metzger (2007) also emphasize that although students are skeptical about online information and report that it should be verified, such verification does not occur. According to Brem et al. (2001), even students who attempt to evaluate online information in a critically reflective manner often have difficulty applying objective evaluation criteria. Apparently, there is a "dubious association," (List & Alexander, 2018, p. 209), i.e., a clear difference between what students think is the correct way to deal with online information and what they *actually do* (Fogg et al., 2002). This could be due to the fact that the assessment of online information is also based on unconscious processes, which can be taken into account by means of eye tracking.

Eye tracking to operationalize cognitive processes

Eye tracking (ET) enables the identification of the position of the eyes as they move over a stimulus. A stimulus is any material that the eyes are confronted with, e.g., text, web pages, images or videos (Scheiter & Van Gog, 2009). The most common video-based corneal reflex method today uses an infrared reflection to visualize and record the cornea and its position relative to the pupil (Djamasbi, 2014; Goldberg & Wichansky, 2003).

In recent years, the use of ET has received increasing attention in the context of investigating processes of dealing with online information (Granka et al., 2008). Existing ET studies investigate the effects of different multimedial content on recipients (Beymer, Orton & Russell, 2007; Chuang & Liu, 2012), online information seeking behavior (Granka, Joachims & Gay, 2004; Zhou & Ren, 2016), evaluation and selection of online information (Gerjets et al., 2011; Sülflow & Schäfer, 2019), and problem resolution processes when using the Internet (Horstmann et al., 2009).

Fundamentally, ET is based on the dogmatic assumption that eye movements reveal information about individuals' cognitive processes (Just & Carpenter, 1980). ET is a periactional method, which means that data are collected during the subject's actions (simultaneously), thus allowing direct and immediate access to their cognitive processes (Roldan, 2017). Compared to verbal methods, ET is hardly reactive, i.e., there is little or no influence on the behavior of test persons during the assessment. In addition, ET reduces the general problem of self-reports, i.e., that respondents could make untruthful statements regarding their procedures due to social desirability (Neuert & Lenzner, 2019; Sülflow & Schäfer, 2019). For this paper, the most decisive advantage of using ET methodology is that subjects are not even consciously aware of many of their cognitive processes and therefore cannot report them (Neuert & Lenzner, 2019; Scheiter & Van Gog, 2009). ET opens up the possibility of not having to rely on (subjective) information from test persons, which can be erroneous and incomplete. This, in turn, makes it possible to obtain evidence about cognitive processes during the assessment of online information on the basis of objective data (Granka et al., 2008; Wang et al., 2014).

Conceptual framework of COR

To validly assess a critical-reflective approach to using online information among university students, we based our study on the framework of "Critical Online Reasoning" (Molerov et al., 2020). This construct includes three interconnecting facets: (i) Online Information Acquisition, (ii) Critical Information Evaluation, and (iii) Reasoning based on Evidence, Argumentation, and Synthesis (for details, Molerov et al., 2020). To measure and promote this skill, we developed a new tool 'Critical Online Reasoning Assessment' (CORA). The assessment focuses specifically on students' ability to search for and evaluate online information and make a reasoned decision using selected information to solve a problem/answer a question presented in a CORA task. This framework was comprehensively validated according to the Standards for Educational and Psychological Testing by AERA et al. (2014) (Molerov et al., 2020; Nagel et al., 2022; Schmidt et al., 2020). CORA included (at the time the eye tracking study was conducted) six tasks, each with a completion time of 10 min and providing students with a description of the context and a website to evaluate. They are asked to conduct an open-ended web search, evaluate online information, and write an open-ended response (a short text) for each task. Two of the six tasks were used in the eye tracking study presented here. (Fig. 8 in the Appendix shows one of the tasks used in this paper).

Modeling the processes involved in COR

Descriptive perspective

For the analysis of students' use of online information, processes can be structured by the descriptive model of Information Problem Solving on the Internet (IPS-I) by Brand-Gruwel et al. (2009). An information problem could be the question of whether and why a website and its information are (not) credible (for examples, see CORA task in the Appendix) (Fig. 1).

Only a few studies have investigated the processes underlying the solving of information problems on the Internet in terms of the individual activities required (Brand-Gruwel et al., 2017; Collins-Thompson et al., 2016). Brand-Gruwel et al. (2009) compared PhD students to first-year students in terms of the underlying processes while solving a task for which they used online information. The main differences in approach were that

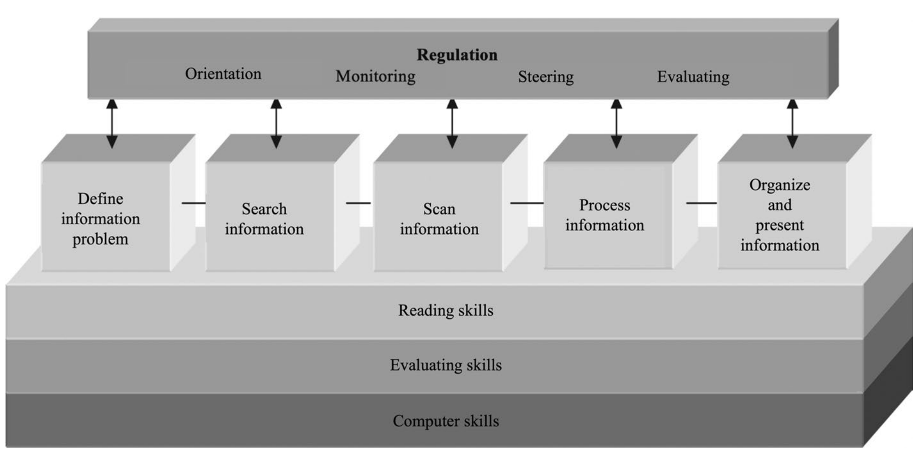

Fig. 1 IPS-I model according to Brand-Gruwel et al. (2009)

PhD students spent more time defining the information problem. In addition, they made a decision regarding the credibility of information at a later stage. In a similar study, Wineburg and McGrew (2017) found that professional fact-checkers read "laterally", leaving a website after a quick scan to first verify the credibility of the website based on an online search and thus through content from other sources, whereas undergraduate students read "vertically" and only stayed on a single website. List and Alexander (2017) describe "sampling" as a concept similar to lateral reading, where the focus is on selecting the best source of information according to certain criteria by quickly scanning sources to select the optimal information. Empirically, however, they show that so-called "satisficing" is more common, where the focus is on content engagement, which is characterized by few sources accessed without revisits and linear reading. Zhou and Ren (2016) had similar findings. They showed that high-achieving students switched more frequently between search results and web pages in the process of seeking information before “landing” on a web page, which was interpreted as stronger metacognitive engagement.

In summary, the results from these studies suggest that high-performing students spend more time reading tasks and activating prior knowledge (defining information problem) and searching for information (sampling, lateral reading) than low-performing students. In this respect, scanning processes could occur comparatively more often in higher-performing students than in lower-performing students, since the latter use fewer sources and engage more with one website.

Processual perspective

According to Flanagin and Metzger (2007), the perception of information credibility can occur not only through heuristic processing of easily accessible cues, but also through systematic processes. Cho et al. (2018) describe processes as components of larger thinking operations, regardless of their degree of complexity, organization and intentionality. They are therefore not necessarily goal-oriented. Systematic processes are consciously employed processes and take place when a subject selects, coordinates and applies various goal-oriented thoughts and actions (Afflerbach et al., 2008). The differentiation of processing into heuristic and systematic processes is known as the heuristic-analytic theory of reasoning (Evans, 2006). According to this theory, heuristic processes are fast, unconscious, automatic, experience-based and occur with little cognitive effort (Gronchi & Giovannelli, 2018; Horstmann et al., 2009; Kahneman, 2011). Systematic processes require more cognitive effort: They are analytical and based on a weighing of positive and negative aspects of different options (Chen & Chaiken, 1999; Evans & Stanovich, 2013). Researchers assume that cognitive effort is a prerequisite for being able to adequately assess the credibility of online information (Afflerbach & Cho, 2009; Bråten & Strømsø, 2011; Metzger et al., 2010). The construction of coherent mental representations of content from sources is associated with a considerable systematic effort (Bråten & Strømsø, 2011; Britt et al., 1999; Perfetti et al., 1999; Stadtler & Bromme, 2014). The evaluation of sources also requires "deep-level processing" (List & Alexander, 2018). Researchers also argue that heuristics can be just as effective and efficient as more cognitively demanding strategies of inference and decision-making (Gigerenzer & Todd, 1999;

Wirth et al., 2007). Following the definition of COR, however, it is assumed that systematic processes are necessary for critically dealing with online information.

Systematic processing does not mean that no heuristic processes take place. According to default-intervention models, these processes take place one after the other (Evans, 2006; Evans & Stanovich, 2013). These models state that heuristic processes are always activated first as a default mode and that systematic, conscious processes can intervene in these intuitive processes. According to Evans (2006), judgements are thus either determined by heuristic processes or the systematic approach is actively used to suppress the default reproductions led by the heuristic system and to engage in conscious, strategic deliberation instead.

Figure 2 illustrates this principle in relation to the evaluation of online information: A website automatically evokes heuristic processes that take place on the basis of superficial features of the website. The subject then decides (consciously or unconsciously) whether the analytical system intervenes. This decision can depend on the task structure, the time available and the subject's intelligence. According to Cho et al. (2018), and in line with the definition of COR, a systematic activation includes the analysis of the source with regard to the expertise, competence and trustworthiness of the author and the organization or person operating the website, and the content analysis of the "main ideas" of the texts. In this way, a fusion with the initial (heuristic) source judgement takes place. Ultimately, a critical, systemic processing through the recursive use of source references and content analysis should result in the final judgement (Molerov et al., 2020).

Operationalizing COR processes for eye movement diagnostics

Selecting an eye movement metric

There are numerous movement-, position-, quantity- and distance-based ET metrics for analyzing eye movement data. For research on gaze behavior, two metrics are most frequently used: fixations and saccades (Beymer et al., 2007; Poole & Ball, 2006). During a fixation, information is extracted and encoded by the observer, with the eyes remaining relatively immobile for approximately 100–800 ms (Duchowski, 2007; Raney et al., 2014). Saccades are rapid eye movements of 10–20 ms between fixations that occur

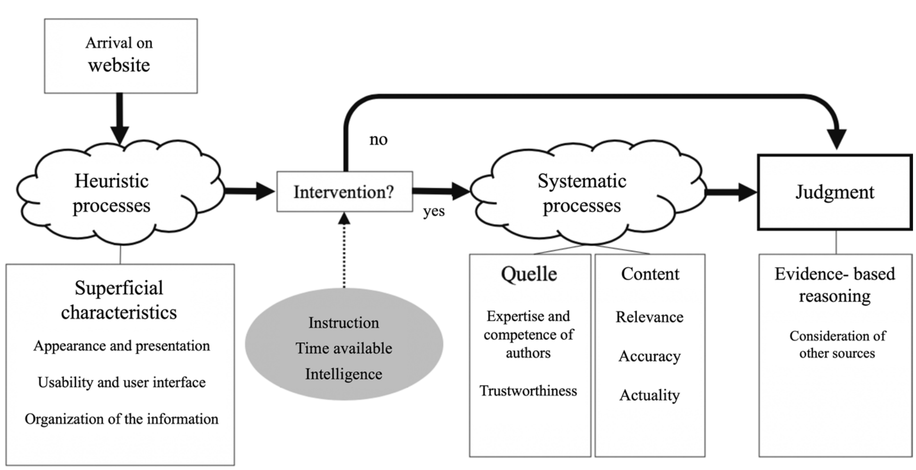

Fig. 2 Default intervention model for dealing with online information (own illustration)

when attention is directed from one object to another (Duchowski, 2007). It is assumed that little to no information is acquired and processed during saccades. Fixations are suitable for studying how information is processed as they make it possible to distinguish superficial scanning from deeper processing of information (Glöckner & Herbold, 2011). They are therefore particularly interesting for cognitive psychological studies and are also used in most ET studies that investigate the handling of online information (Horstmann et al., 2009; Raney et al., 2014; Sülflow & Schäfer, 2019; Wang et al., 2014; Zhou & Ren, 2016). Moreover, fixations have been well studied and validated as indicators of information processing depth compared to other metrics (Holmqvist et al., 2011).

The relationship between eye fixations and cognitive processing has been explored for over two centuries (Wade, 2015). The key finding is that increased processing demands are associated with eye fixation on specific objects or changes in fixation patterns (Raney et al., 2014). According to the "Eye-Mind Assumption" (Just & Carpenter, 1976, 1980), what the eyes fixate on is also what is actually being processed. The “Immediacy Assumption” (Just & Carpenter, 1976) suggests that the duration of a fixation is the same as the duration for which the corresponding object is processed. Accordingly, the speed of fixation shifts also corresponds to the speed of processing. Eye fixations are therefore a major focus of ET research since the attentional allocations implied by them are considered a reliable proxy for the level of processing (Rayner, 1998; Velichkovsky, 1999; Wang et al., 2014). The duration of fixations is to be preferred from a theoretical perspective, since longer fixations are not only an indicator of greater interest on the part of the viewer and a higher level of complexity (Cyr & Head, 2013; Poole & Ball, 2006), but are also most frequently used as an indicator of deeper and cognitively more complex processing (Holmqvist et al., 201; Rayner, 1998; Velichkovsky, 1999; Wang et al., 2014). Research on dealing with online information shows that superficial levels of processing (e.g., scanning a website) are associated with shorter fixations of up to 250 ms, while deeper processing (e.g., systematic integration of information) is associated with longer fixations of over 500 ms (Glöckner & Herbold, 2011).

Fixations (initially) refer to arbitrary areas of the stimulus, which are defined as “Areas of Interest” (AOI). When examining fixations on AOIs, the metric ‘dwell’ is often used (Gidlöf et al., 2013; Klein et al., 2019). A dwell includes all directly consecutive eye movements that are located within an AOI (Holmqvist et al., 2011). The dwell time, i.e., the time spent attending an AOI, can be seen as the counterpart to the fixation duration for a given AOI. Though, dwell time on an AOI is not indicative of fixation duration within that AOI, however, for the analysis of eye movements within AOIs, both the dwell time as well as the total fixation duration within the AOIs are examined, as the latter can provide additional information about the depth of processing.

Identifying the distinct process steps in *CORA* task-solving

Since students’ processing procedures while solving a CORA task can show a high degree of variability due to the open assessment format, they are divided into individual, empirically distinct processing steps that enable comparability between students. For this purpose, the five constitutive steps of the IPS-I model (see 2.3.1) are transferred specifically to CORA task processing. Accordingly, it is assumed that the definition of the information problem mainly takes place when reading the task, as a goal-directed action can only take place following this step (Vermetten et al., 2005). For this reason, the overall duration of all the phases, i.e., the time that students spend on the CORA task and that they do not use to take notes and write answers is assigned to the step of *defining the information problem*.

Even if information can, in principle, be searched for on any website (e.g., through the search function), a search engine is usually used for this purpose (Wirth et al., 2007). Therefore, the step of *searching for information* is represented by the time students spend on the website of a search engine. This is a simplified indicator; there is also the possibility that students may already be thinking about a search strategy or search terms before they go to the search engine. However, it is assumed that, in most cases, this is mainly done while visiting the search engine (Hoppe et al., 2018; Pifarré et al., 2018).

Both *scanning and processing of information* can take place on the same websites. Thus, eye movement indicators come into play for differentiation. More superficial processing, which is common when scanning information, is associated with shorter fixations, while information integration, which requires deeper, more elaborate processing, is associated with longer fixations (Sect. 3.2.1). According to Glöckner and Herbold (2011), the key figures of up to 250 ms per fixation for short fixations and over 500 ms per fixation for longer fixations are used in our study. Shorter fixations on the web pages are interpreted as an indicator for scanning processes of information, while the duration of longer fixations is used as an indicator for processing information.

Students organize and present information mainly by writing answers and citing evidence. Therefore, all the time students spend on writing texts and inserting copied statements and URLs as part of answering the CORA tasks is assigned to this processing step. Table 1 summarizes the operationalization of process steps during CORA task-solving.

Defining areas of interest (aois)

For the theoretical derivation of AOIs, the MAIN model by Sundar (2008) is used as a framework, which represents an approach to understanding credibility evaluation in the use of online media. Sundar (2008) recommends the model to advance the study of credibility heuristics in research and identifies four "cues" that have significant psychological effects on the assessment of credibility of online information: modality cues (M), agency cues (A), interactivity cues (I) and navigability cues (N).

Modality cues refer less to the content and more to the structure of a webpage, namely the differences in the effects that visual, auditory and textual elements have on the

Table 1 Operationalization of process steps in the CORA task-solving

Constitutive steps	Application to CORA task-solving	Unit of analysis for operationalization
Define information problem	Reading the task	Time spent on the task not writing
Searching for information	Use of search engine	Time spent using search engines
Scanning information	Scanning the web pages used in CORA	Duration of all fixations of up to 250 ms on all visited web pages
Process information	Deeper processing of information on web pages used in CORA	Duration of all fixations of more than 500 ms on all visited web pages
Organise and present information	Write notes and reply/insert links	Time spent on the task while writing

subject. Overall, different modalities (image, text, video and audio) evoke certain heuristics that can have both positive and negative effects on the judgement of the credibility of the information, depending on the relationship between them. In particular, visual elements such as images and graphics feature highly as design elements in various frameworks for evaluating the credibility of online information. Wathen and Burkell (2002) cite factors such as external appearance in the form of graphics and color design. Fogg (2003) refers to surface credibility, which is assessed by superficially processing the page structure based on its individual modules, as "most common". Text modality on a website is classified as any continuous text that contains information and is not assigned to any other cue. This corresponds to the complete introductory part of the text.

Agency cues refer to the source of the information, which is reflected both by the "identity" of a website and by the author of the information. Thus, the organization that operates a website can play a role in assessing the site's credibility. At the same time, references to the author can also evoke heuristics, which, depending on the context, can have both a positive effect and a negative effect. Assessing the credibility of the source is an essential component of COR and is emphasized in various empirical studies (e.g. Metzger & Flanagin, 2013; Fogg et al., 2002; Winter & Krämer, 2014; Elsweiler & Kattenbeck, 2019). Other frameworks also emphasize the importance of source information: In Wathen and Burkell (2002), information about the author (and the resulting assessment of their expertise and trustworthiness) is an essential part of evaluating the credibility of a message. Fogg (2003) places more emphasis on information about the organization running the website, which is grouped under "presumed credibility". According to Metzger and Flanagin's (2015) factors of credibility evaluation, source and message cues are two of four categories of assessing credibility on a website. They include qualification, references, contact information, motives and the author's reputation. Areas on the website that give information about the operating organization as well as those that give information about the author are therefore considered AOIs.

Many web pages contain attributes called *interactivity cues*. Such interactivity elements allow the user's needs to be specified, as they make the medium "responsive". Sundar (2008) assigns dialogue boxes, search functions, menu bars and communication possibilities to interactivity. Such interactive elements are also mentioned by Metzger and Flanagin (2015) as a source cue. Since empirical studies show the effect of interaction elements on credibility (e.g., Jahng & Littau, 2016; Johnson & Kaye, 2016), these too should be defined as AOIs on a website.

Finally, *navigability cues* consist of interface features such as cross-references and access to other content. The navigation design is expressed in the use of hyperlinks, the increased use of which, according to Sundar, can lead to an "elaboration heuristic", which leads recipients to a deeper processing of the content by clicking on the links. References to further information (e.g., citations and sources) can also be grouped among these cues. Metzger and Flanagin (2015) also list citations and links among the so-called message cues as characteristics of a website for assessing its credibility, specifying these in terms of their quality. Thus, citations of research sources and links to external authorities would increase credibility. Further qualitative studies confirm the relevance of cross-references and external links as criteria for assessing credibility (Eysenbach & Köhler, 2002; Freeman & Spyridakis, 2004). Therefore, the section with cross-references to other contents of a website as well as the citation of a study in a text and the references below the text are also considered AOIs. Figure 3 shows the application of the criteria to build AOIs.

In addition, there is a wide range of so-called "checklists" for the correct handling of information on the Internet by learners, both for the school (Klicksafe Initiative, (2020); State Institute for Teacher Education and School Development, 2012; State Agency for Civic Education, 2005) and university sector (Leibniz Technical Information Library (2021); Ulm University, (2008); Bielefeld University, (2008)). These lists include criteria that should be taken into account by the recipient to adequately assess the credibility of online information. The criteria from the abovementioned selection of three checklists each for university and school practice were also analyzed to consider further practical criteria for assessing the credibility of information when forming AOIs.

Study and evaluation design

Research context and sample

The eye movements during the processing of the CORA tasks were recorded with a Tobii Pro X3-120 Eye Tracker fixed to the screen and a sampling rate of 120 Hz. The stimuli generated, which included the CORA tasks with the links to the websites, were tested in a pretest with two test persons to ensure a technically smooth process as well as the comprehensibility of the content and the time frame. No noteworthy anomalies were found.

In the winter semester 2019/2020, ET data were collected from 32 students from two German universities, who were selected based on theoretically predefined criteria

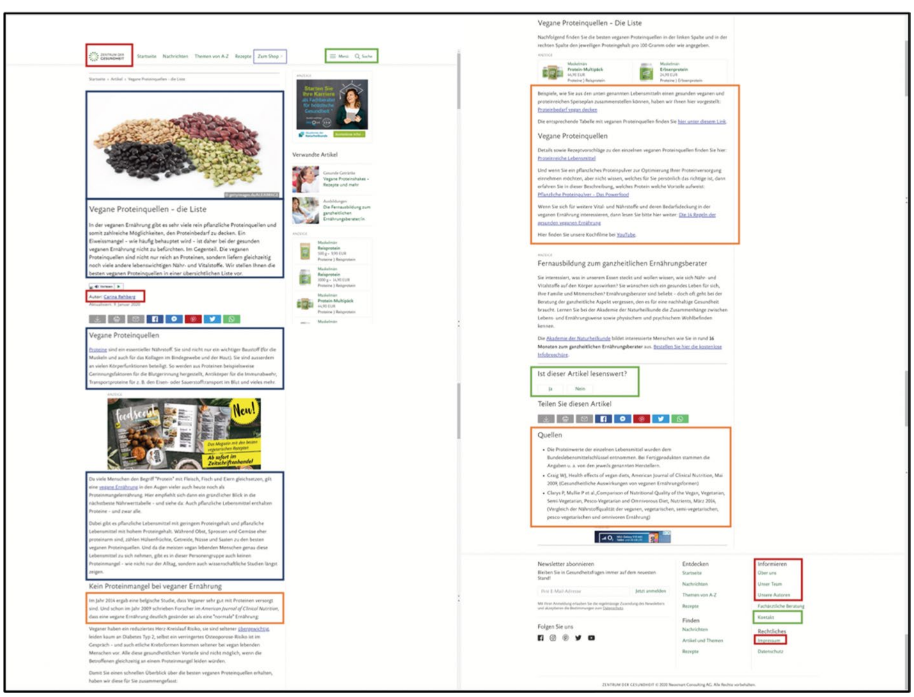

Fig. 3 Application of the cues of the MAIN model to the website from CORA

(gender, age, course of study, study semester) in a cross-sectional design (see Table 2). All students were given the same two CORA tasks to complete.

To study eye movement data, data quality must be ensured (Holmqvist et al., 2012). Holmqvist et al. (2011) recommend a maximum average deviation of the measured gaze points from the actual gaze points of 0.5°, while they describe values above 1.0° as "unacceptable". Data quality is particularly important for eye movement diagnostics within AOIs, as lower precision and accuracy increases the likelihood that fixations will not be assigned to the correct AOI, especially if there is insufficient distance between AOIs (Holmqvist et al., 2012). According to Orquin et al. (2016), values deviating from the recommended accuracy would have to be compensated for by increasing the size of the AOI, i.e., creating a buffer distance to other sections of the stimulus. In a "real-world environment" (Holmqvist et al., 2011), however, it is not possible to influence the properties of the stimulus, which makes it difficult to enlarge the margins. In our study, therefore, all students whose precision and/or accuracy values were below the acceptable value of 1.0° were excluded from the analysis to avoid interpretation errors. The exclusion of test persons with unacceptable values reduced the sample by nine students. Two further students had to be excluded from the analysis due to missing values in the socio-demographic part of the study. Thus, data from 19 subjects were used for this paper.

To answer the RQs, the sample was separated into two groups based on their CORA scores to investigate differences regarding their processing procedures. From an educational practice perspective, we specifically focus on low-performing students, as there is a potential need for support among this group. The distribution of the CORA scores shows that a large proportion of the students in the sample have deficits in the COR facets. Students who did not argue at all or hardly argued critically and reflectively in their evaluation of the online information in CORA (low performers, LP) and scored less than one point on a scale from 3 (max.) to 1 (min.) were grouped together and contrasted with the group of students who argued (at least) partially critically and reflectively (high performers, HP) and scored at least one point or higher. Table 2 shows the socio-demographic distribution as well as the group size of LPs and HPs.

Table 2 Sample description

Attribute	Group	Distribution
Gender	Male female	5 (26.3%) 14 (73.7%)
Course of Study	Business Studies Business Education	7 (36.8%) 12 (63.2%)
Age	Up to 22 years (median) Older than 22 years	10 (52.6%) 9 (47.4%)
Semester	Up to second semester (median) From third semester	11 (57.9%) 8 (42.1%)
Performance in CORA	Low Performer (LP) [a] High Performer (HP) [b]	12 (63.2%) 7 (36.8%)

N = 19 [a]scored less than one point on CORA. [b]scored at least one point and higher on CORA

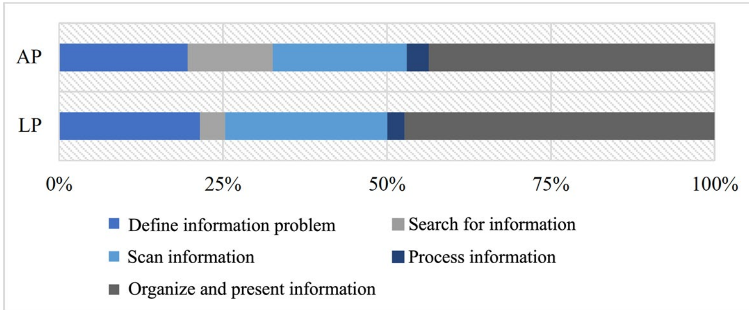

Fig. 4 Average relative duration of the process steps (based on Strobel et al., 2018)

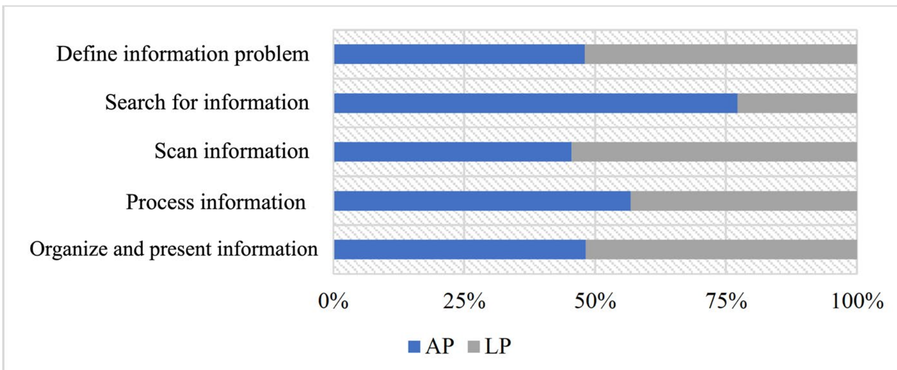

Fig. 5 Distribution of the total duration of each IPS-I step between the LP and HP groups

Results

Process steps

Firstly, a descriptive comparison of the distribution of the processing time on the five processing steps according to the IPS-I model between the two performance groups (LP vs. HP) as well as the distribution of the total duration of the individual processing steps between the two groups shows that the most salient differences relate to the processing steps that took place during the use of search engines and websites (Figs. 4, 5).

Secondly, in relation to their total processing time, LPs spent proportionately less time on searching for further external information compared to HPs. The majority of the duration of the search processes carried out by all students in total (77.2%) is accounted for by the group of HPs. The average absolute difference in the duration for the information search between LPs and HPs is significant (37.10 s), however, at an α-level of 10% ($U=21.00$, $Z=-1.876$, $p=0.061$). The effect of group membership (Cohen's $d=0.892$) is large (Cohen, 1988).

With regard to the duration of website use, LPs' was 71.5 times longer on average in relation to the total duration of website use (MEAN = 303.97; SD = 129.28); 71.4% of the time was spent on the webpage directly linked in the CORA (MEAN = 216.96; SD = 115.16), while HPs on average only spent 58.2% (MEAN = 168.78; SD = 72.77) of their time spent on webpages (MEAN = 289.90; SD = 80.65) on this one webpage. Although the difference between the mean values of the relative time spent on this website in relation to other websites was not significant (t (17) = − 1.136, p = 0.272), there is still at least a medium effect (Cohen's d = -0.54) Table 3.

While the average difference in the time spent scanning information processes between HPs and LPs (16.57 s) was not significant (t (17) = − 0.871, p = 0.396), there was at least a small to medium effect of group membership (Cohen's d = − 0.414). Although HPs spent less time on websites, they processed the information on the websites slightly longer and/or more often deeply. However, at 3.28 s, the difference in mean is small in absolute terms, not significant (t (17) = 0.807, p = 0.431), and the effect size was relatively small (Cohen's d = 0.384).

When considering all processing steps, LPs switched significantly less frequently between the individual processing steps of the IPS-I model. While HPs switched 13.43 times on average between defining the information problem, searching for information, using web pages, and presenting the information, this was only the case 8.75 times for LPs (t(17) = 2.065, p = 0.054).

Eye movement on the web page

In relation to their total time on the website, LPs spent a large part (64.3%) of their time looking at the (introductory) text, while HPs devoted less than half (41.2%) of their time on the website to this AOI. Measured in time, more than three quarters of all gaze movements (78.0%) were allotted to the LPs. In contrast, the difference in dwell time for the graph is initially hardly noticeable in relative terms compared to other AOIs, since the share of dwell time in the total viewing time of the webpage was small (0.7% for LP and 2.2% for HP). There was also only a moderate difference in the distribution of all eye movements on the graphic between the two groups (44.4% for LP and 55.6% for HP), which, however, speaks for a longer viewing time for HPs. The absolute mean of the dwell time of the graph is nevertheless significantly higher for HPs (mean difference 1.39 s) than for LP (U = 15.00, Z = -2.286, p = 0.022) (Fig. 6).

LPs looked at the AOI containing the citation section of the study longer than HPs, both proportionally (12.5% of the webpage viewing time) and in relation to total eye movements of all students (76.0%). In absolute terms, the total dwell time of the HPs

Table 3 Mean comparison of the duration of the processing steps

Constitutive step according to the IPS-I model	Mean of LP (N = 12) (SD)	Mean difference HP to LP (SE)	*p*-value	Effect size Cohen's *d*
Define information problem	86.28 (45.58)	− 6.73 (19.82)	0.738[a]	− 0.162
Searching for information	15.54 (24.45)	+ 37.10 (17.96)	0.061[b]	0.892
Scanning information	99.27 (45.21)	− 16.57 (19.03)	0.396[a]	− 0.414
Process information	10.48 (8.82)	+ 3.28 (4.06)	0.431[a]	0.384
Organize and present information	189.88 (65.46)	− 13.40 (38.13)	0.730[a]	- 0.167

All time values in seconds and representation based on Keller et al. (2015)[a] Calculation by *t* test for independent samples.
[b] Calculation by Mann–Whitney U-test with indication of asymptotic significance, since the conditions N > 12 apply to the total sample and N > 6 to both groups (Gibbons & Chakraborti, 2003)

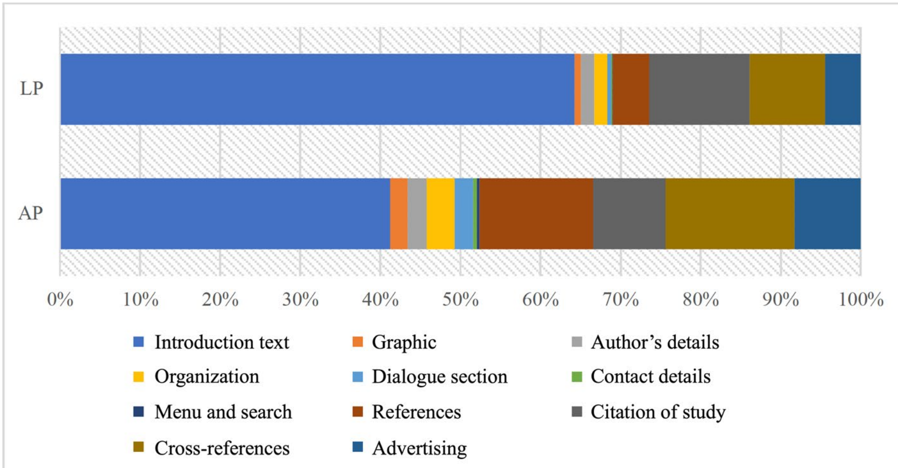

Fig. 6 Proportional distribution of dwell time on the AOIs in relation to the total time on the website

was significantly shorter, on average by 17.89 s, than the LPs' ($t(17) = -2.312$, $p = 0.034$), which corresponds to a large effect (Cohen's d = 1.1). The same direction of effect was also observed for the dwell time on the cross-references, although the mean difference (24.49 s) was only significant at an α-level of 10% (U = 20.50, Z = − 1.818, $p = 0.069$). For the dwell time on the AOI sources, the direction of effect was different. LPs looked at the sources for a shorter amount of time than HPs (share of 14.2% of the viewing time of the website), both proportionally (4.5%) to the total viewing time of the website and in relation to all eye movements of the students within the AOI (41.9%).

For the total dwell time over all advertisements no significant difference between LPs and HPs were found Table 4.

Based on the p-values in combination with the effect size of the mean differences in the context of the comparison of the dwell time, a medium to high significance of the differences between HPs and LPs with regard to six AOIs (text, graphics, organization, sources, citation of the study and cross-references) could be highlighted. In order to draw conclusions about the depth of processing of the content within the AOIs, the next step was to examine whether differences between the performance groups could also be found with regard to the length of fixations.

HPs fixated on the graphic significantly longer, on average by 0.98 s (U = 17.00; Z = − 2.132, $p = 0.033$, Cohen's d = 1.108). Also, the HPs' fixation time in the source area was 11.52 s longer on average than the LPs' (U = 22.50, Z = -1.655, $p = 0.098$). Lastly, the HPs fixated on the information about the organization for a mean of 2.15 s longer than the LPs. However, the areas with the description of the study (U = 19.50, Z = − 1.902, $p = 0.057$) and with the cross-references to other contents of the website (U = 20.50, Z = -1.818, $p = 0.069$) were fixated on significantly longer by the LPs than by the HPs at the 10% level. Apart from the effect of group membership on the fixation duration in the text area, all the effects mentioned are high (Cohen, 1988) Table 5.

These findings are corroborated by the comparative visualization of the mean fixation duration of the two groups on the website shown in Fig. 7. The red squares represent areas within which the LPs had a longer fixation duration on average, while green

Table 4 Mean value comparison of the total dwell time

AOI	Mean of LP (N = 12) (SD)	Mean difference HP to LP (SE)	*p* value	Effect size Cohen's *d*
Introduction text	62.01 (11.89)	− 27.39 (18.92)	0.166[a]	− 0.689
Graphic	0.52 (0.22)	+ 1.39 (0.63)	0.022[b]	1.229
Author's details	0.93 (0.28)	− 0.11 (0.45)	0.815[a]	− 0.113
Organizational information	1.20 (0.51)	+ 1.92 (1.22)	0.132[a]	0.752
Dialogue section	0.47 (0.24)	+ 0.10 (0.35)	0.778[a]	0.136
Contact details	0.04 (0.03)	+ 0.08 (0.10)	0.415[a]	0.398
Menu and search	0.03 (0.02)	+ 0.01 (0.03)	0.619[a]	0.241
References	4.99 (1.59)	+ 13.74 (7.05)	0.188[b]	0.630
Citation of the study	27.53 (5.06)	− 17.89 (7.74)	0.034[a]	− 1.100
Cross-references	34.42 (9.57)	− 24.49 (13.29)	0.069[b]	− 0.917
Advertising	4.12 (1.14)	− 0.17 (2.06)	0.935[a]	− 0.039

All times in seconds and representation based on Keller et al. (2015) [a]Calculation by *t* test for independent samples. [b]Calculation by Mann–Whitney U-test with indication of asymptotic significance, since the conditions N > 12 apply to the total sample and N > 6 to both groups (Gibbons & Chakraborti, 2003)

Table 5 Mean comparison of fixation duration for relevant AOIs

AOI	Mean of LP (N = 12) (SD)	Mean difference HP to LP (SE)	*p* value	Effect size Cohen's d
Introduction text	41.30 (33.15)	− 17.28 (14.83)	0.260[a]	− 0.554
Graphic	1.45 (0.47)	+ 0.98 (0.44)	0.033[b]	1.108
Organizational information	2.76 (0.61)	+ 2.15 (1.27)	0.069[b]	0.917
References	14.23 (2.71)	+ 11.52 (7.23)	0.098[b]	0.817
Citation of the study	6.81 (17.95)	− 11.13 (5.61)	0.057[b]	− 0.970
Cross references	7.34 (21.28)	− 13.95 (6.77)	0.069[b]	− 0.917

All times in seconds and representation based on Keller et al. (2015) [a]Calculation by *t* test for independent samples. [b]Calculation by Mann–Whitney U-test with indication of asymptotic significance, since the conditions N > 12 apply to the total sample and N > 6 to both groups (Gibbons & Chakraborti, 2003)

squares are those areas within which HPs fixated on the content longer overall. Equivalent to the previous mean comparison, a comparatively longer total fixation duration can be seen for LPs in the areas of the study and the cross-references, while for HPs,

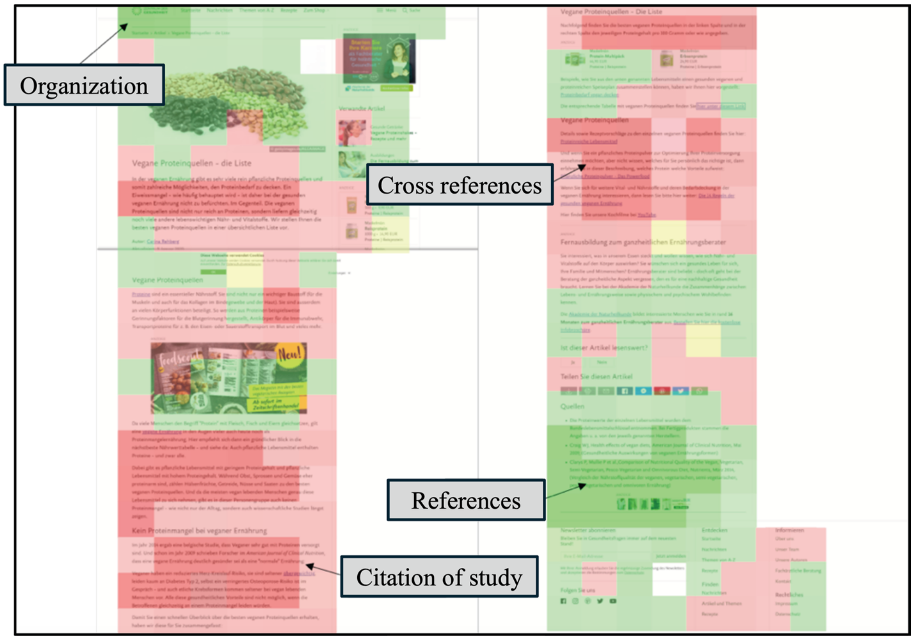

Fig. 7 Visual comparison of the distribution of mean fixation duration between the performance groups on the website

longer fixation durations can be located within the areas with the information on the organization as well as the sources.

Discussion, limitations, and implications for future research

In this study, we investigate the three RQs via a combination of the descriptive and processual approach. Regarding RQ1 and RQ2, we were able to demonstrate that actual student behaviors exhibited when solving the CORA tasks using unrestricted Internet search could be operationalized through eye tracking data. The eye movements of students who critically and reflectively evaluate the credibility of online information ('high performers') differ from those of students who do not critically and reflectively evaluate ('low performers') as theoretically expected. Investigating RQ3, we were able to show, that students who assessed online information less critically-reflectively (LPs) processed the areas that are particularly important for evaluating credibility (source references and information on the organization or author) more briefly or more quickly. Moreover, they processed less relevant areas (e.g., information internal to the website such as cross-references) for longer instead of looking for further external information to verify the content presented. As a result, they tended to spend proportionately less time than the HPs on obtaining information from other websites. This approach corresponds more closely to the characteristics of heuristic processes, since it was mainly the information presented and less information from different sources that was combined to form a mental model — this is often associated with less cognitive effort (Britt & Rouet, 2012). Compared to HPs, a less recursive analysis of sources and content took place (Cho et al., 2018), which is illustrated by the lower number of processing steps.

Examining the findings presented here, we are able to conclude that using the ET measurement method the study provided us with significant insights into the differences in approaches between higher- and lower-performing students when solving the online information problem task using unrestricted Internet search. In this way, our study demonstrates that this method can be used successfully in such open information and learning environments without the predefined AOIs (e.g. compared to several studies with closed tests, e.g., Klein et al., 2020). Within the framework of the assessment design, we therefore created an ecologically highly valid test environment. In addition to quantitative data analysis, case-by-case qualitative investigations of individual processing procedures based on ET records provided more specific insights into the differences in procedures in the future research.

At the same time, against the background of the use of the ET methodology, some limitations of this study must be pointed out. Although the duration of fixations is considered a viable indicator of the depth of processing, there are hardly any guidelines on what length of fixations classify as deep or less deep processing (Horstmann et al., 2009). The chosen threshold values according to Glöckner and Herbold (2011), which were used to distinguish scanning and processing processes, are based on study results on dealing with online information and not on cognitive psychological theories. Specifying which fixation lengths are selected for which levels of processing can have a significant impact on the results of data analysis.

Eye-movement data should generally be interpreted through theory-driven operationalization (Granka et al., 2008; Holmqvist, 2011). In the context of a theoretical embedding, fixation duration could be an indicator of stronger interest, higher content complexity or other phenomena instead of deeper processing and more effective information integration (Cyr & Head, 2013; Poole & Ball, 2006). In addition, eye movements are different individually and can therefore be interpreted differently from person to person (Rakoczi, 2012). The causality between eye movements and cognitive processes should be further focused on in future studies by integrating qualitative, competing methods during processing, e.g., the method of thinking aloud (Bojko, 2013; Gerjets et al., 2011).

Moreover, this study on the use of online information is based on a sample of only 19 students. Although this sample size corresponds to a quite common size in relevant national and international eye-tracking research, a larger sample would be desirable for additional analyses of task solving processes, such as a multilevel modeling. Therefore, future research projects need to draw from a larger and more representative sample, and ensure high data quality during the surveys, since the generalizability of findings can be significantly impaired by this (Holmqvist et al., 2012).

In addition, in the presentation of the default intervention model in dealing with online information, we noted that the decision on whether to use systematic processes in assessing websites depends, for example, on instruction, available time, and intelligence (Evans, 2006; Wathen & Burkell, 2002). In particular, the time available in the CORA could play a major role, as it was limited to ten minutes per task. Furthermore, factors such as (task-related) prior knowledge, personal beliefs, and motivation can also influence the approach to assessing information (Britt & Rouet, 2012;

Kammerer et al., 2013; Metzger, 2007; Scheiter & Van Gog, 2009). These influencing factors should be investigated in future studies.

Despite these limitations, our study offers starting points for further hypothesis-testing investigation of students' critical and reflective handling of online information to explore those deficits identified here and their causes in more detail. In particular, this study complements the existing research base on university students' heuristics in dealing with real Internet information by using the ET measurement method to gain a deeper understanding of how online websites and content is actually used to solve a generic information problem. Upon analyzing the AoIs further, this study substantially contributes to other existing studies that have published similar findings using other assessment methods (e.g., Flanagin & Metzger, 2007; Walraven et al., 2009; Zhang, Cole & Belkin, 2011) and indicate high potential of the ET method to investigate learning processes in real online information environments that increasingly dominate the current university landscape.

Regarding practical implications, the results suggest that students often do not proceed to evaluate online information in a critically reflective manner. In view of the increasing degree of digitalization and the greater importance of the ability to evaluate content from the World Wide Web, there is obviously a need for support in educational practice. Based on the results of our study it may be useful to systematically educate students about indicators that influence the credibility of online information on websites such as source provenance, as well as learning helpful procedures to validate the information at hand such as cross-checking.

Appendix

See Fig. 8.

Task 1: Vegan Protein Sources

Take about 10 minutes to complete this task.

Visit the following website:

https://www.zentrum-der-gesundheit.de/vegane-proteinquellen.htlm

Is this a reliable source of information about vegan protein sources?

Note: You can use any information on this website, as well as search the internet. Justify your answer with evidence from the internet sources used and include the associated URLs.

Please write down your answer to task 1 and the URLs here:

Fig. 8 Task of the CORA

Abbreviations

AOI	Areas of Interest
COR	Critical Online Reasoning
CORA	Critical Online Reasoning Assessment
ET	Eye tracking
HP	High performers
IPS-I	Information Problem Solving on the Internet
LP	Low performers
RQ	Research question

Acknowledgements
We would like to thank all students from the Medical Faculty of Goethe University Frankfurt, and from the Faculty of Law and Economics at Johannes Gutenberg University who participated in this study. We would like to thank the reviewer and the editor who provided constructive feedback and helpful guidance in the revision of this manuscript. We would also like to thank all students who participated in this study.

Author contributions
A-KK carried out the assessments, conducted the analyses, and co-wrote the article. OZ-T co-developed the assessment, supervised the analyses, and co-wrote the article. SS and SB co-developed and carried out the assessment and supported with the data analysis. M-TN carried out the assessment. All authors read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. This study was part of an RMU project, which was funded by the RMU fund. The funding did neither influence the design of the study, nor the collection, analysis, or interpretation of data or writing the manuscript.

Availability of data and material
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The studies involving human participants were reviewed and approved by the State Officer for Data Protection and Freedom of Information Rhineland-Palatinate. The participants provided their written informed consent to participate in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 21 August 2023 Accepted: 11 September 2024
Published online: 09 October 2024

References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teaclier, 61*(5), 364–373.
Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education, 125*, 413–428.
Barzilai, S., & Zohar, A. (2012). Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cognition and Instruction, 30*(1), 39–85. https://doi.org/10.1080/07370008.2011.636495
Ben Khedher, A., Jraidi, I., & Frasson, C. (2018). Static and dynamic eye movement metrics for students' performance assessment. *Smart Learning Environments, 5*, 1–12.
Bera, P., Soffer, P., & Parsons, J. (2019). Using eye tracking to expose cognitive processes in understanding conceptual models. *MIS Quarterly: Management Information Systems., 43*(4), 1105–1126. https://doi.org/10.25300/MISQ/2019/14163
Beymer, D., Orton, P. Z. Russell, D. M. (2007). An eye tracking study of how pictures influence online reading. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 456–460
Bielefeld University (2020). Checkliste zur Bewertung von Internetquellen [Checklist for evaluating internet sources]. Retrieved from https://www.uni-bielefeld.de/gesundhw/studienberatung/guide/assessment_internet_sources.pdf.
Bojko, A. (2013). *Eye Tracking The User Experience - A Practical Guide to Research*. Brooklyn, New York: Rosenfeld Media

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers and Education, 53*(4), 1207–1217. https://doi.org/10.1016/j.compedu.2009.06.004

Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L. & van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning, 33*(3), 234–251

Bråten, I., & Strømsø, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning, 6*(2), 111–130. https://doi.org/10.1007/s11409-011-9075-7

Brem, S. K., Russell, J., & Weems, L. (2001). Science on the web: Student evaluations of scientific arguments. *Discourse Processes, 32*, 191–213.

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. Goldman, A. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209–233). Lawrence Erlbaum Associates Publishers.

Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes* (pp. 276–314). Cambridge: Cambridge University Press.

Cho, B.-Y., Afflerbach, P., & Han, H. (2018). Strategie Processing in Accessing, Comprehending, and Using Multiple Sources Online. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 133–150). Routledge.

Chuang, H. H., & Liu, H. C. (2012). Effects of different multimedia presentations on viewers' information-processing activities measured by eye tracking technology. *Journal of Science Education and Technology, 21*(2), 276–286. https://doi.org/10.1007/s10956-011-9316-1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Collins-Thompson, K., Rieh, S. Y., Haynes, C. C. Syed, R. (2016). Assessing Learning Outcomes in Web Search: A Comparison of Tasks and Query Strategies. CHIIR '.16 163–172

Coskun, A., & Cagiltay, K. (2021). A systematic review of eye-tracking-based research on animated multimedia learning. *Journal of Computer Assisted Learning, 38*, 581–598.

Cyr, D., & Head, M. (2013). The impact of task framing and viewing timing on user website perceptions and viewing behavior. *International Journal of Human-Computer Studies, 71*(12), 1089–1102. https://doi.org/10.1016/j.ijhcs.2013.08.009

Djamasbi, S. (2014). Eye Tracking and Web Experience 2. *Visual System: How Do We See? AIS Transactions on Human-Computer Interactionransactions on Human-Computer Interaction, 6*(2), 16–31. https://doi.org/10.17705/1thci.00060

Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice*. Springer. https://doi.org/10.1007/978-1-84628-609-4

Elsweiler, D., & Kattenbeck, M. (2019). Understanding credibility judgements for web search snippets. *Aslib Journal of Information Management, 71*(3), 368–391. https://doi.org/10.1108/AJIM-07-2018-0181

Erdogan, R., Saglam, Z., Cetintav, G., & Karaoglan Yilmaz, F. G. (2023). Examination of the usability of Tinkercad application in educational robotics teaching by eye tracking technique. *Smart Learning Environments, 10*(1), 27.

Evans, J. S. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*(3), 378–395.

Evans, J. S., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science, 8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ (Clinical Research Ed.), 324*, 573–577. https://doi.org/10.1136/bmj.324.7337.573

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media and Society, 9*(2), 319–342. https://doi.org/10.1177/1461444807075015

Fogg, B. J., Soohoo, C., Danielson, D., Marable, L., Stanford, J. & Tauber, E. R. (2002). How do people evaluate a Web site's credibility: Results from a large study. Retrieved from https://dejanmarketing.com/media/pdf/credibilityonline.pdf

Fogg, B. J. (2003). Credibility and the world wide web. *Persuasive Technology*. https://doi.org/10.1016/b978-155860643-2/50009-3

Freeman, K., & Spyridakis, J. (2004). An Examination of Factors That Affect the Credibility of Online Health Information. *Technical Communication, 51*, 239–263.

Gadiraju, U., Yu, R., Dietze, S. Holtz, P. (2018). Analyzing knowledge gain of users in informational search sessions on the web. CHIIR 2018 Proceedings of the 2018 Conference on Human Information Interaction and Retrieval. https://doi.org/10.1145/3176349.3176381

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye tracking data. *Learning and Instruction, 21*(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Gibbons, J. D., & Chakraborti, S. (2003). *Nonparametric statistical inference* (4th ed.). Marcel Dekker.

Gidlöf, K., Wallin, A., Dewhurst, R., & Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research., 6*(1), 1–14. https://doi.org/10.16910/jemr.6.1.3

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Glöckner, A., & Herbold, A.-K. (2011). An eye tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making, 24*, 71–98. https://doi.org/10.1002/bdm.684

Goldberg, J., & Wichansky, A. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 493–516). Elsevier.

Granka, L., Joachims, T. Gay, G. (2004). Eye tracking analysis of user behavior in WWW search. *Proceedings of Sheffield SIGIR - Twenty-Seventh Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 478–479. https://doi.org/10.1145/1008992.1009079

Granka, L., Pellacini, F., & Pan, B. (2008). Eye Tracking and Online Search: Lessons Learned and Challenges Ahead. *Journal of the American Society for Information Science and Technology, 59*(7), 1041–1052. https://doi.org/10.1002/asi.20794

Gronchi, G., & Giovannelli, F. (2018). Dual process theory of thought and default mode network: A possible neural foundation of fast thinking. *Frontiers in Psychology, 9*, 1–4. https://doi.org/10.3389/fpsyg.2018.01237

Han, J., Chen, L., Fu, Z., Fritchman, J., & Bao, L. (2017). Eye-tracking of visual attention in web-based assessment using the force concept inventory. *European Journal of Physics*. https://doi.org/10.1088/1361-6404/aa6c49

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 University Press.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. *Eye Tracking Research and Applications Symposium (ETRA), 1*(212), 45–52. https://doi.org/10.1145/2168556.2168563

Hoppe, A., Holtz, P., Kammerer, Y., Yu, R., Dietze, S. & Ewerth, R. (2018). Current Challenges for Studying Search as Learning Processes. *Learning & Education with Web Data*, *LILE2018*, 2–5. Abgerufen unter https://lile2018.wordpress.com/

Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation? An eye tracking analysis of instruction-induced decision modes. *Judgment and Decision Making, 4*(5), 335–354. https://doi.org/10.2139/ssrn.1393729

Iding, M. K., Crosby, M. E., Auernheimer, B., & Klemm, E. B. (2009). Web site credibility: Why do people believe what they believe? *Instructional Science, 37*(1), 43–63. https://doi.org/10.1007/s11251-008-9080-7

Klein, P., Küchemann, S., Brückner, S., Zlatkin-Troitschanskaia, O., & Kuhn, J. (2019). Student understanding of graph slope and area under a curve: a replication study comparing first-year physics and economics students. *Physical Review Physics Education Research, 15*(2), 1–17

Klicksafe Initiative (2020). Glaubwürdigkeitscheck [Credibility check]. Retrieved from https://www.klicksafe.de/suchmaschinen/quellenkritik-und-bewertungskompetenz/

Jahng, M. R., & Littau, J. (2016). Interacting is believing: Interactivity, social cue, and perceptions of journalistic credibility on Twitter. *Journalism and Mass Communication Quarterly, 93*(1), 38–58. https://doi.org/10.1177/1077699015606680

Johnson, T. J., & Kaye, B. K. (2016). Some like it lots: The influence of interactivity and reliance on credibility. *Computers in Human Behavior, 61*, 136–145. https://doi.org/10.1016/j.chb.2016.03.012

Just, M. A., & Carpenter, P. A. (1976). *Eye Fixations and Cognitive. Cognitive Psychology, 8*(4), 441–480.

Just, M. A., & Carpenter, P. A. (1980). Psychological Review. *Psychological Review, 87*(4), 329–354. https://doi.org/10.1093/mind/xxv.3.415-b

Kahne, J., Hodgin, E., & Eidman-Aadahl, E. (2016). Redesigning Civic Education for the Digital Age: Participatory Politics and the Pursuit of Democratic Engagement. *Theory and Research in Social Education, 44*(1), 1–35. https://doi.org/10.1080/00933104.2015.1132646

Kahneman, Watson K. D. (2011). Thinking, Fast and Slow New York NY: Farrar Straus and Giroux. *Canadian Journal of Program Evaluation., 26*(2), 111–3.

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior, 29*(3), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

Keller, J., Gorges, M., Horn, H. T., Aho-Özhan, H. E. A., Pinkhardt, E. H., Uttner, I., Kassubek, J., Ludolph, A. C., & Lulé, D. (2015). Eye tracking controlled cognitive function tests in patients with amyotrophic lateral sclerosis: A controlled proof-of-principle study. *Journal of Neurology, 262*(8), 1918–1926. https://doi.org/10.1007/s00415-015-7795-3

Klein, P., Lichtenberger, A., Küchemann, S., Becker, S., Kekule, M., Viiri, J., Baadte, C., Vaterlaus, A., & Kuhn, J. (2020). Visual attention while solving the test of understanding graphs in kinematics: An eye-tracking analysis. *European Journal of Physics*. https://doi.org/10.1088/1361-6404/ab5f51

Lee, W.-K., & Wu, C.-J. (2017). Eye movements in integrating geometric text and figure: Scanpaths and given-new effects. *International Journal of Science and Mathematics Education, 16*, 1–16.

Leibniz Technical Information Library (TIB) (2021). Checkliste bei der Bewertung von Internetquellen [Checklist in evaluating internet resources]. Retrieved from https://www.tib.eu/fileadmin/Data/documents/learning-work/checklist-in-evaluating-internet-sources.pdf.

List, A., & Alexander, P. A. (2017). Text navigation in multiple source use. *Computers in Human Behavior, 75*, 364–375.

List, A., & Alexander, P. A. (2018). Corroborating students' self-reports of source evaluation. *Behaviour & Information Technology, 37*(3), 198–216. https://doi.org/10.1080/0144929X.2018.1430849

Luo, L., Kiewra, K. A., Peteranetz, M. S., & Flanigan, A. E. (2017). Using Eye-Tracking Technology to Understand how Graphic Organizers Aid Student Learning. In C. Was, F. Sansosti, & B. Morris (Eds.), *Eye-tracking technology applications in educational research* (pp. 220–238). GI Global.

Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O., & Jitomirski, J. (2020). *Positive and Negative Media Effects on University Students' Learning: Preliminary Findings and a Research Program*. In O. Zlatkin-Troitschanskaia (Ed.). Frontiers and Advances in Positive Learning in the Age of Information (PLATO) (pp. 109–119). Springer.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can Students Evaluate Online Sources? Learning From Assessments of Civic Online Reasoning. *Theory & Research in Social Education, 46*(2), 165–193.

McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology, 2019*, 1–16.

Metzger, M. J. (2007). Making Sense of Credibility on the Web: Models for Evaluating Online Information and Recommendations for Future Research Miriam. *Journal of the American Society for Information Science and Technology, 58*(13), 2078–2091. https://doi.org/10.1002/asi.20672well

Metzger, M. J. & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics, 59*, 210–220. https://doi.org/10.1016/j.pragma.2013.07.012

Metzger, M. J., & Flanagin, A. J. (2015). Psychological Approaches to Credibility Assessment Online. In S. S. Sundar (Ed.), *The Handbook of the Psychology of Communication Technology.* New Jersey: John Wiley Sons.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication, 60*(3), 413–439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Molerov, D., Zlatkin-Troitschanskaia, O., Nagel, M. T., Brückner, S., Schmidt, S., & Shavelson, R. (2020). Assessing University Students' Critical Online Reasoning Ability: A Conceptual and Assessment Framework with Preliminary Evidence. *Frontiers in Education, 5*(1), 1–29. https://doi.org/10.3389/feduc.2020.577843

Nagel, M.-T., Zlatkin-Troitschanskaia, O., & Fischer, J. (2022). Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates. *Frontiers in Education*. https://doi.org/10.3389/feduc.2022.914857

Navarro, O., Molina Díaz, A. I., Lacruz Alcocer, M., & Ortega Cantero, M. (2015). Evaluation of Multimedia Educational Materials Using Eye Tracking. *Procedia - Social and Behavioral Sciences, 197*, 2236–2243.

Neuert CE, Lenzner T. 2019 Use of eye tracking in cognitive pretests. Leibniz Institute for the Social Sciences, Mannheim. https://doi.org/10.15465/gesis-sg_en_025

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206. https://doi.org/10.1016/j.actpsy.2013.06.003

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Lawrence Erlbaum Associates Publishers.

Pifarré, M., Jarodzka, H. M., Brand Gruwel, S., & Argelagós, E. (2018). Unpacking cognitive skills engaged in web-search: How can log files, eye movements, and cued-retrospective reports help? An in-depth qualitative case study. *International Journal of Innovation and Learning, 24*(2), 152. https://doi.org/10.1504/ijil.2018.10014361

Poole, A., & Ball, L. (2006). Eye tracking in human-computer interaction and usability research: Current status and future prospects. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). IGI Publishing.

Rakoczi, G. (2012). Eye Tracking in Forschung und Lehre. Möglichkeiten und Grenzen eines vielversprechenden Erkenntnismittels. In G. Csanyi, F. Reichl, & A. Steiner (Hrsg.), *Digitale Medien - Werkzeuge für exzellente Forschung und Lehre* (S. 87–98). Münster u.a.: Waxmann.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments, 83*, 1–7. https://doi.org/10.3791/50780

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, 124*(3), 372–422.

Roldan, S. M. (2017). Object Recognition in Mental Representations: Directions for Exploring Diagnostic Features through Visual Mental Imagery. *Frontiers in Psychology, 8*, 833. https://doi.org/10.3389/fpsyg.2017.00833

Scheiter, K., & Van Gog, T. (2009). Using Eye Tracking in Applied Research to Study and Stimulate the Processing of Information from Multi-representational Sources. *Applied Cognitive Psychology, 23*, 1209–1214. https://doi.org/10.1002/acp.1524

Schmidt, S., Zlatkin-Troitschanskaia, O., Roeper, J., Klose, V., Weber, M., Bültmann, A.-K., & Brückner, S. (2020). Undergraduate students' critical online reasoning - process mining analysis. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2020.576273

Sharma, K., Giannakos, M., & Dillenbourg, P. (2020). Eye-tracking and artificial intelligence to enhance motivation and learning. *Smart Learning Environments, 7*, 1–19.

Stadtler, M., & Bromme, R. (2014). The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. Braasch (Eds.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (pp. 379–402). MIT Press.

State Agency for Civic Education (2005). Mit dem Internet unterrichten - Informationen für Lehrerinnen und Lehrer [Teaching with the Internet - Information for teachers]. Retrieved from http://www.politikundunterricht.de/1_05/baustein_d.pdf

State Institute for Teacher Education and School Development (2012). Checkliste zur Bewertung von Internetquellen für Schülerinnen und Schüler [Internet resource assessment checklist for students]. Retrieved from https://li.hamburg.de/contentblob/3461588/aeeb63b90b0c1ca82dbb0737d318392c/data/pdf-internetquellen-bewerten-in-der-profiloberstufe.pdf;jsessionid=D9DD030256B00DF294BB41138BA5330E.liveWorker2

Sülflow, M., & Schäfer, S. (2019). Selective attention in the news feed: An eye tracking study on the perception and selection of political news posts on Facebook. *New Media Society, 21*(1), 168–190. https://doi.org/10.1177/1461444818791520

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *Digital Media, Youth, and Credibility*. https://doi.org/10.1162/dmal.9780262562324.073

Ulm University (2008). Bewertungskriterien für Internetquellen [Evaluation criteria for internet sources]. Retrieved from https://www.uni-ulm.de/fileadmin/website_uni_ulm/kiz/bib/schuelermaterial/criteria_internet_sources.pdf.

Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition: Converging evidence from three domains of research. In B. H. Challis & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (p. 203). Amsterdam: John Benjamins Publishing Company.

Vermetten, Y., Brand-Gruwel, S., & Wopereis, I. G. J. H. (2005). Information Problem Solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior, 21*, 487–508.

Wade, N. (2015). How Were Eye Movements Recorded Before Yarbus? *Perception, 44*(8–9), 851–883. https://doi.org/10.1177/0301006615594947

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers and Education, 52*(1), 234–246. https://doi.org/10.1016/j.compedu.2008.08.003

Wang, Q., Yang, S., Liu, M., Cao, Z., & Ma, Q. (2014). An eye tracking study of website complexity from cognitive load perspective. *Decision Support Systems, 62*, 1–10. https://doi.org/10.1016/j.dss.2014.02.007

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology, 53*(2), 134–144. https://doi.org/10.1002/asi.10016

Wineburg, S. McGrew, S. (2017). Lateral Reading: Reading Less and Learning more when Evaluating Digital Information Wineburg, Sam and mcgrew, sarah, Lateral Reading: Reading Less and Learning More When Evaluating Digital Information. Stanford History Education Group Working paper no 20

Wineburg, S., Breakstone, J., McGrew, S., & Ortega, T. (2018). Why Google Can't Save Us. In O. Zlatkin-Troitschanskaia (Ed.), *Frontiers and Advances in Positive Learning in the Age of InformaTiOn* (pp. 221–228). Wiesbaden: Springer Fachmedien Wiesbaden.

Winter, S., & Krämer, N. C. (2014). A question of credibility - Effects of source cues and recommendations on information selection on news sites and blogs. *Communications, 39*(4), 435–456. https://doi.org/10.1515/commun-2014-0020

Wirth, W., Böcking, T., Karnowski, V., & Von Pape, T. (2007). Heuristic and systematic use of search engines. *Journal of Computer-Mediated Communication, 12*(3), 778–800. https://doi.org/10.1111/j.1083-6101.2007.00350.x

Zhang, X., Cole, M. & Belkin, N. (2011). Predicting users' domain knowledge from search behaviors. In W.-Y. Ma, J.-Y. Nie, R. Baeza-Yates, T.-S. Chua, & W. B. Croft (Hrsg.), *Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval*. New York: ACM Press. 1225–1226

Zhou, M. & Ren, J. (2016). Use of Cognitive and Metacognitive Strategies in Online Search: An Eye tracking Study. *International Conferences ITS, ICEduTech and STE 2016*. Abgerufen von https://files.eric.ed.gov/fulltext/ED571583.pdf

Zlatkin-Troitschanskaia, O., Brückner, S., Fischer, J., Molerov, D., & Schmidt, S. (2021a). Performance Assessment and Digital Training Framework for Young Professionals´ Generic and Domain-Specific Online Reasoning in Law, Medicine and Teacher Practice. *Journal of Supranational Policies of Education, 13*, 9–36. https://doi.org/10.15366/jospoe2021.13.001

Zlatkin-Troitschanskaia, O., Brückner, S., Molerov, D., & Bisang, W. (2019). What Can We Learn from Theoretical Considerations and Empirical Evidence on Positive and Negative Learning in Higher Education? Implications for an Interdisciplinary Research Framework. In O. Zlatkin-Troitschanskaia (Ed.), *Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO)* (pp. 281–303). Springer.

Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., & Krstev, J. (2021b). Students' online information use and learning progress in higher education – A critical literature review. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2021.1953336

Publisher's Note