What does it take to deal with academic literature? Epistemic components of scientific literacy

Hannes Münchow[a], Tobias Richter[a] & Sebastian Schmid[b]

[a] University of Würzburg (Germany)

[b] University of Regensburg (Germany)

To appear in: Zlatkin-Troitschanskaia, O., Pant, H. A., Toepper, M., & Lautenbach, C. (Eds.) (2019). *Student learning in German higher education: Innovative modelling and measurement approaches and research results*. Wiesbaden, Germany: Springer.

Abstract

The skills required for understanding and evaluating academic literature include a broad repertoire of different reading strategies, which are rarely explicitly taught and go beyond classical learning strategies that foster learning from expository texts. This chapter proposes a taxonomy of strategies for reading academic literature, which distinguishes between two different processing goals (receptive vs. epistemic) and processing modes (systematic vs. heuristic). Recent research on epistemic-systematic reading strategies, diagnostic instruments to assess these strategies and training interventions to foster these strategies is described in more detail. Finally, the chapter provides an outlook on further research that includes epistemic-heuristic reading strategies as another key component of scientific literacy.

Keywords: argument comprehension, argument evaluation, epistemic learning strategies, receptive learning strategies, scientific literacy

Author note:

Preparation of this paper was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF; grant 01PK15009A and grant 01PK15009B).

Corresponding authors: Hannes Münchow & Tobias Richter, University of Würzburg, Department of Psychology IV, 97070 Würzburg, Germany

E-mail: hannes.muenchow@uni-wuerzburg.de, tobias.richter@uni-wuerzburg.de

1 What does it take to deal with academic primary literature? More than deep-level learning strategies

University students are often advised to organize the information they are reading, to relate it to relevant prior knowledge, and to reread passages they find difficult to understand (e.g., Pintrich 2004). Other suggestions to improve learning from text include the use of practice tests (Karpicke and Roediger 2010) or the distribution of rereading across time (Rawson and Kintsch 2005). This advice is based on a broad body of research on learning strategies, reading strategies, and self-regulated learning in cognitive and educational psychology, rooted in theories of human cognitive architecture (Weinstein and Mayer 1986) and supported by a wealth of correlational studies and training experiments (for reviews, see Dunlosky et al. 2013; Richardson et al. 2012). Certainly, the learning strategies that have been in the focus of cognitive and educational research for many years can help students to deeply comprehend expository texts, remember the information conveyed by these texts, and retrieve it from long-term memory at a later test. However, there is doubt whether these strategies suffice for appropriately dealing with other types of academic literature and in other types of reading situations. Academic primary literature, in particular, might call for the application of strategies that have rarely been studied in previous research.

In the first part of this chapter, we briefly sketch *a taxonomy of strategies* that goes beyond traditional models of reading and learning strategies (e.g., Pintrich 2004; Weinstein and Mayer 1986) and that covers a broader range of reading situations that university students are typically confronted with. We seek to approximate a more comprehensive conceptualization of reading strategies that together form the core of scientific literacy. Scientific literacy can be defined as "the ability to understand and critically evaluate scientific content to achieve one's goals" (Britt et al. 2014, p. 104). In line with this definition, we distinguish two generic reading goals, which we call *receptive and epistemic goals*, and two modes of understanding and evaluating scientific content, the *heuristic and the systematic mode*. Receptive reading goals involve comprehending and memorizing information (e.g., for recalling this information in a later test); epistemic reading goals involve the use of texts for the acquisition of knowledge the reader regards as plausible or true (Richter 2003). The *epistemic reading strategies* that are functional for accomplishing epistemic reading goals are novel in the sense that they are not covered by classical taxonomies of reading learning strategies (e.g., Weinstein and Mayer 1986) and have only recently come to the fore of educational psychology research (Barzilai and Zohar 2014; Richter and Schmid 2010). In the second part of the chapter, we focus on systematic epistemic reading strategies. More specifically, we give an overview of recent research from our lab on the characteristics of systematic epistemic reading strategies, how these strategies can be assessed, and how they can be taught to university students through systematic, computer-based trainings.

1.1 A Taxonomy of Reading Strategies for Academic Literature

Scientific content is mainly communicated through written texts. These texts are typically written by scientists to be read by scientists (Goldman and Bisanz 2002). Yet, university students of almost all disciplines are required to read academic literature as part of their studies. To illustrate the scope of the learning strategies required to comprehend academic literature, it is instructive to consider the way scientists themselves read scientific texts. The extant studies (Bazerman 1985; Berkenkotter and Huckin 1995) indicate that scientists routinely employ a broad repertoire of reading strategies, which include both heuristic and epistemic strategies. The physicists interviewed by Bazerman (1985), for example, distinguished between core reading and peripheral reading. In their peripheral reading, they scanned texts for particular words, skipped sections, and evaluated publications by their authors. Remarkably, projected onto classical taxonomies of learning strategies, these reading behaviors routinely exhibited by experts in the domain of physics would have to be regarded as superficial processing, although they apparently help the physicists to achieve certain goals: for example, to find specific information in a text or to select those texts for a closer reading that seem to be most informative or trustworthy.

Our taxonomy covers the variety of reading situations and goals that scientists and students alike are confronted with by posing two generic reading goals and two modes of processing scientific information. First, we distinguish between two types of goals: receptive vs. epistemic goals. Readers with a receptive goal strive to learn facts, understand the text contents, or find specific information, whereas readers with an epistemic goal strive to gain an adequate picture of the state-of-affairs described in the text or develop an own standpoint on the issues discussed in the text (Richter 2003, 2011; Richter and Schmid 2010). Receptive reading goals can be fully described within the classical information processing framework of cognitive psychology (Neisser 1967); learning strategies that help achieving receptive learning goals improve cognitive processes such as encoding, manipulation, storage, and retrieval of information. However, how and to what extent the processed information corresponds to state-of-affairs in the world, whether it is true or whether there are good reason to assume that it is true, is irrelevant for the attainment of receptive reading goals. For example, students studying for an exam need to comprehend and memorize information in expository texts, but it is often irrelevant to their study goal whether the theories described in the text are actually valid. By contrast, for epistemic reading goals, these criteria are essential. Thus, epistemic reading goals involve the acquisition of knowledge in a classical (philosophical) sense, that is, the acquisition of true and justified beliefs (e.g., Ichikawa and Steup 2018). For example, a student who aims at identifying the most powerful theory in a given field of study would follow an epistemic reading goal.

Second, we distinguish between systematic (deep-level) or heuristic (surface-level) strategies (following to the distinction made in two-process models of information processing, e.g., Petty and Wegner 1999). Systematic strategies involve controlled processes, are cognitively demanding, and, if successful, lead to deeper understanding of a scientific issue or domain. By contrast, heuristic strategies can be applied fast and demand less cognitive resources. If successful, heuristic strategies result often in a specific decision, for example, that a piece of information matches the answer of a question, that a statement is implausible, that a document is trustworthy, or that it is worthwhile to read a text more thoroughly. The distinction between systematic and heuristic strategies does not imply that systematic strategies lead to better results than heuristic strategies. Rather, it depends on the processing goal and other conditions (such as the available time, reader's expertise, and the information provided by the text) whether a systematic or heuristic strategy is appropriate.

Combining the dimensions epistemic-receptive and systematic-heuristic yields a 2x2 table with four categories of strategies (Figure 1). The four categories are described next with examples.

1.1.1 Receptive-Systematic Strategies

The strategies in this category serve to enrich or structure information to facilitate later recall. Examples are classical learning strategies such as rehearsal, organization and elaboration, as described in the educational and cognitive research on reading and learning strategies (e.g. Dunlosky et al. 2013; Pintrich 2004).

1.1.2 Receptive-Heuristic Strategies

The strategies in this category serve to gain a first impression of the text content. Readers can rely on genre knowledge about the types of information texts in a particular genre typically provide. Once a reader has selected a particular text for further scrutiny, he or she can turn to receptive-heuristic strategies such as *skimming* (cursory reading to extract the gist) and *scanning* (cursory reading to find specific information). These strategies are particularly important in the phase of literature research and require generic knowledge of canonic text structures (Dillon 1991).

1.1.3 Epistemic-Systematic Strategies

The strategies in this category serve to validate the argumentation of a text. In some models of learning strategies, they are covered by the construct *critical evaluation* (e.g. Pintrich 2004). At a more specific level, it is possible to distinguish between *consistency checking* and *knowledge-based validation* (Richter and Schmid 2010). These strategies require the reader to identify the functional components of arguments such as *claim* and *ground* (reasons), to evaluate the acceptability of the reasons, and to evaluate the internal consistency of the argument, i.e. the relevance and sufficiency of the reasons (Larson, Britt and Kurby 2009; Toulmin 1958). University students often fail to identify the argumentative function of text passages (e.g., Norris et al. 2003).

1.1.4 Epistemic-Heuristic Strategies

The strategies in this category serve to gain a quick preliminary evaluation of the credibility of the text. Examples include the use of source information (e.g., publication outlet and funding, Zimmerman et al. 2001), fast judgments of the plausibility of claims (Voss et al. 1993), or the selective processing of belief-consistent information (Maier and Richter 2013; Richter and Maier 2017). Such strategies are particularly important when readers lack domain-specific content knowledge or the cognitive or motivational resources necessary for deep processing. Contrary to scientists, students often neglect source information when making epistemic judgments (Zimmerman et al. 2001), which may hamper their comprehension of scientific information (e.g., Strømsø, Bråten and Britt 2010).

University students need to possess a broad knowledge of strategies in all four categories for competently dealing with academic literature. However, they also need to know when and for what purpose they can use a particular strategy (conditional knowledge, Lorch et al. 1993). Skimming and scanning, for example, are particularly important in the phase of literature research to find out whether a particular publication is relevant for the question at hand. Only if this condition is met, the reader should turn to systematic strategies. To give another example, readers following an epistemic reading goal might not be able to adequately judge the arguments in a text as they lack pertinent prior knowledge. In that case, the epistemic reading goal might be served better by making a heuristic judgement about the credibility of the source, even if the text presents the scientific issue in a way that it seems easy to comprehend (Scharrer et al. 2014).

1.2 Epistemic-Systematic Reading Strategies: Assessment, Training, and Relevance for Studying at the University

In the remainder of this chapter, we present an overview of research from our own lab on epistemic reading strategies, i.e. the strategies required for comprehending and evaluating the arguments presented in scientific texts. Epistemic reading strategies go beyond a receptive elaboration of the texts' contents and enable students to acquire knowledge about scientific issues as opposed to identifying and memorizing information. Such reading strategies are usually not explicitly taught in school, which contributes to the problems many first-year students encounter when reading academic primary literature. Deficits in epistemic strategies were revealed, for example, in student-scientist comparisons that examined how psychology students and scientists (advanced doctoral students and postdocs) comprehended and evaluated arguments (von der Mühlen et al. 2016a, 2016b). In these studies, scientists were superior in decoding the functional structure of informal arguments, in identifying implausible arguments, and in recognising errors in argumentation. These findings underline the need for diagnostic instruments for the assessment of epistemic-systematic strategies and for training interventions that foster these strategies in university students. In a number of studies to be described next we developed and evaluated such assessments and trainings for undergraduates in the social sciences.

1.2.1 Assessment of Epistemic-Systematic Reading Strategies

For assessing students' epistemic-systematic reading strategies, we developed two computer-based diagnostic instruments: the Argument Structure Test (AST, Münchow et al., in press) and the Argument Judgement Test (AJT, Münchow et al. 2019).

1.2.1.1 Argument Structure Test

The Argument Structure Test assesses students' ability to identify the structural components of informal arguments, which may be considered the building blocks of academic literature. The test consists of eight short informal arguments (*M* = 104 words, *SD* = 24 words) taken from typical psychological texts. The arguments are composed of argument components according to Toulmin (1958). According to this model, arguments consist of up to five functionally different components: The claim or the statement that is being argued for, one or several (empirical, theoretical, or practical) reasons that support the claim, a warrant that states why the reason(s) should support the claim, a backing that justifies the warrant empirically or theoretically, and the rebuttal that limits the validity of the claim (e.g., by referring to exceptions). In empirical sciences, the claims found in scientific texts are most often theoretical in nature (e.g., an explanatory assumption) and the reasons are findings from empirical studies. This principle is also followed in the arguments used in the Argument Structure Test. Respondents' task is to assign the five functional arguments components to one of the sentences of each of the eight arguments. The arguments used in the Argument Structure Test differ with regard to several characteristics that are known to affect argument comprehension. First, the position of claim and reasons is varied. In half of the arguments, the claim is located in the first sentence of the argument (claim-first arguments), whereas the other half of the arguments start with a reason (reason-first arguments). Claim-first arguments follow the canonical order and are easier to comprehend as the claim is the key to building up a mental representation of the argument (Britt and Larson 2003). Moreover, argument complexity is varied systematically. Two arguments are simple arguments that contain only three of the five argument components distinguished by Toulmin, whereas the remaining six arguments contain all five argument components (complex). The two simple arguments are claim-first arguments.

Respondents first read an argument as a continuous text (Figure 2). Afterwards they are shown the same argument again and are asked to assign the sentences of the argument to one of the five argument components (claim, reason, warrant, backing, and rebuttal) via a dropdown menu. Students usually need 20 to 30 minutes to complete all items of the Argument Strucure Test. The total number of accurately assigned argument components serves as a score of the reader’s ability to decode the functional structure of informal arguments.

Psychometric properties of the Argument Structure Test were evaluated in a study with a convenience sample of 225 psychology undergraduates and teacher students. In this study, the internal consistency (Cronbach's α) of the Argument Structure Test reached .76, with a wide range of item difficulties (M = .69, SD = .16). The items of the Argument Structure Test showed a good fit to the Rasch model (1-PL model, Andersen Likelihood-Ratio test with a mean-split of the sample: $\chi^2[df = 38, N = 225] = 46.81, p = .130$), with no indication of interdependencies between items within a specific argument.

To examine the construct validity of the Argument Structure Test, we estimated an explanatory item response model (LLTM, Fischer 1974) to predict the test’s item difficulties through theoretically relevant item characteristics (order of argument components and arguments complexity). As expected, claim-first arguments and less complex arguments were easier to decode. Finally, the observed item difficulties based on the Rasch model could be predicted well through item difficulties estimated with the LLTM ($R^2 = .82$, Figure 3), which is strong evidence for the construct validity of the instrument.

Evidence for the criterial validity of the test was obtained through correlations of the test scores with criterial performance measures as well as the students' epistemological beliefs (scales Structure and Variability of the Connotative Aspects of Epistemic Beliefs Questionnaire, CAEB; Stahl and Bromme 2007). The test scores of the Argument Structure Test were moderately and signifantly ($p < .05$) correlated with verbal intelligence ($r = .40$) assessed with the subtests sentence completion, analogies and commonalities of the Intelligence Structure Test (IST 2000 R, Amthauer, Brocke, Liepmann and Beauducel 2001) and students' Grade Point Average in the school-leaving certificate (Abitur; $r = .17$). Moreover, students who scored higher in the Argument Structure Test were more likely to see knowledge in psychology as structured but changeable, as indicated by significant correlations with the scales Structure ($r = .20$) and Variability ($r = -.33$, reverse scored) of the CAEB. In sum, the Argument Structure Test is a reliable and valid instrument for assessing the ability to comprehend arguments in scientific texts.

1.2.1.2 Argument Judgement Test

The Argument Judgement Test assesses students' abilities to accurately judge the plausibility of informal arguments and to identify common argumentation fallacies. The test consists of two parts. In Part 1 of the Argument Judgement Test, readers are presented two short expository texts about smoking behavior (550 words) and objective self-awareness (404 words). Each text consists of 15 short informal arguments containing a claim and one or several reasons. Twenty out of 30 arguments are plausible, that is, these arguments contain strong and internally consistent reasons that support the claim. The remaining ten arguments contain one of five common argumentation fallacies (i.e., contradiction, false dichotomy, wrong example, circular reasoning, overgeneralization; Dauer 1989), resulting in poor arguments. The readers' task in Part 1 of the Argument Judgment Test is to evaluate whether the presented arguments are plausible or implausible by pressing a corresponding key on the keyboard. The number of correctly judged arguments serves as test score. In Part 2, the arguments that the respondents judged as implausible in Part 1 are presented again, and respondents are asked to assign each of these arguments to one of the five common argumentation fallacies via a dropdown menu. Each fallacy is briefly explained on screen. Participants can also select the answer options *I don't know*, *I was wrong, there is no error*, or *None of the above-mentioned errors, but ...*, where they can enter text in a text box for the last option. The number of correctly assigned arguments is used as a test score in Part 2. Furthermore, a combined score of the responses in Part 1 and Part 2 can be formed, which is a measure of the ability to evaluate the plausibility of informal arguments. An example item of the Argument Judgement Test is shown in Figure 4.

The psychometric properties of the Argument Judgment Test were explored in a study (Münchow et al. 2019) based on the same convenience sample of 225 psychology and teacher students used for examining the Argument Structure Test. In this study, Part 1 turned out to be relatively easy (item difficulty: $M = .74$, $SD = .10$), whereas Part 2 was relatively difficult (item difficulty: $M = 36$; $SD = .22$). However, a combined score of Part 1 and Part 2 responses lead to a wide distribution of item difficulties. The items of the combined scale showed a good fit with the Rasch model (1-PL model, Andersen LR-test based on a mean-split of the sample: $\chi^2[df = 25, N = 225] = 27.53$, $p = .330$). The internal consistency for the combined score was acceptable (with a WLE reliability coefficient of .63), and the stability (test-retest reliability) within an interval of 13 months reached .60 in an independent sample of 22 psychology students.

Construct validity of the Argument Judgment Test was examined by estimating an explanatory item response model revealing that implausible arguments were more difficult to detect than plausible arguments. In addition, we estimated linear mixed model (items X participants) with the students' response times in Part 1 of the Argument Structure Test as the dependent variable. The main idea behind these analyses was that argument evaluation is a rational, effortful activity, which should lead to longer response times in implausible arguments, if these arguments were indeed recognized as implausible. Indeed, there was a significant interaction ($p < .05$) effect of response accuracy and argument plausibility that followed the expected pattern (Figure 5). These results provide some evidence that effortful processing is needed for accurately evaluating implausible arguments, whereas plausible arguments can be evaluated more efficiently.

With regard to criterial validity, test scores correlated moderately and significantly ($r = .36$) with verbal intelligence (assessed with the verbal subtests of the IST 2000 R, Amthauer et al. 2001) and the students' current average grade as well as the student's epistemological beliefs (again measured with the CAEB). Moreover, AJT test scores were significantly associated with students' academic success measured via the students' current grade average at university, even if verbal intelligence and the students' Grade Point Average from school leaving certificates were controlled for. The increment of explained variance was 20%.

1.2.2 Training of Epistemic-Systematic Reading Strategies

The studies conducted with the Argument Structure Test and the Argument Judgment Test suggest that large individual differences exist in the ability to comprehend and evaluate informal arguments (Münchow et al. 2019, in press). Moreover, first year students show deficits in epistemic-systematic strategies (e.g. von der Mühlen 2016a), presumably because these strategies are not explicitly taught in school, where scientific knowledge is often presented as a "monolith of facts" (Osborne 2010, p. 464) rather than the product of rational, argument-based discourse. Thus, there is a need for effective trainings to foster epistemic reading strategies in university students. We addressed this need by developing and evaluating two computerized trainings, an *Argument Structure Training* to enhance students' argument decoding skills and an *Argument Judgment Training* to foster students' skills of evaluating the internal consistency and validity of the argumentation of a text. Both trainings convey conceptual knowledge about arguments (their functional components or normative criteria for their evaluation), which is presented via illustrated texts, audio examples, and short video clips. Various exercises during and after the theoretical parts give participants the opportunity to apply and practice the content they have learned. Participants receive direct feedback after each task and can either redo that specific task, return to the corresponding theory block, or continue training. The trainings last about 45-60 minutes, but there is no limit to the training time. The evaluation of the trainings was based on the Argument Structure Test and the Argument Judgement Test, respectively.

1.2.2.1 Argument Structure Training

The argument structure training imparts strategies for evaluating scientific arguments by training the identification and allocation of functional argument components. The theoretical input focuses on the use and purpose of informal arguments, the Toulmin (1958) model of argumentation, and linguistic connectors and key words that help to correctly identify certain argument components.

The effectiveness of the argument structure training was evaluated in two experimental pre-post-test studies with a follow-up test four weeks after the training. Both studies employed an active control group that received a computerized speed reading training that involved reading but did not train epistemic-systematic reading strategies in any way. Participants of the first training experiment were 53 psychology students at the beginning of their studies. Students in the training condition outperformed participants in the control condition in their ability to identify and assign less typical argument components (i.e. warrants) and to correctly identify argument components in arguments with a less typical structure, i.e. reason-first arguments (medium-sized effects significant at $p < .05$). Moreover, the training intervention was especially effective for students that, according to their grades, were more successful in their studies (von der Mühlen et al. 2018). However, there were no differences between the two conditions in their argument structure decoding skills at follow-up measures, which indicates that the training effects did not remain stable over a period of four weeks. We therefore conducted a second training experiment to evaluate effectiveness of the Argument Structure Training plus a 15-minute booster training session in the week before the follow-up tests. Analyses of the data from this study are still in progress.

1.2.2.2 Argument Judgment Training

The argument judgment training teaches strategies for the normatively appropriate evaluation of arguments, especially strategies for evaluating the relevance and completeness of reasons for the justification of an argument's claim, and trains students to recognize typical errors in argumentation, such as circular reasoning or overgeneralization. Figure 6 shows example pages of the argument structure training.

The Argument Judgment Training was also evaluated in an pre-post-test experimental design with an active control group (speed-reading training) and a follow-up after four weeks (as yet unpublished study). Psychology students and teacher students participated in this study. Similarly to the results for the Argument Structure Training, participants performed better in the training condition than in the control condition at post-test, but the effect disappeared in the post-test after four weeks.

In sum, both trainings developed to foster epistemic-systematic reading strategies in university students produced immediate effects on the trained skills, but the effects were not stable over time. Moreover, evidence for transfer effects, for example on study performance, is still lacking. We are planning future studies to address these issues. In these studies, we plan to add additional instructional measures such as practice tests (Greving and Richter 2018) and interleaved presentation of content (Brunmair and Richter 2019) to the training that promise to foster long-term and transfer effects.

1.3 Conclusion and Outlook

In this chapter, we outlined a taxonomy of reading strategies that differentiates between epistemic-systematic strategies, epistemic-heuristic strategies, receptive-systematic strategies, and receptive-systematic strategies, and described tests and trainings that target epistemic-systematic strategies. Epistemic-systematic strategies are particularly relevant for adequately dealing with scientific literature, but have been neglected in previous research. Other tests and trainings from our lab also cover epistemic-heuristic strategies. The tests for the assessment of these strategies (e.g., the Credibility Judgment Test; von der Mühlen et al. 2016a) involve different sets of texts, which are presented for a limited time to prevent students from engaging in systematic text processing. Thus, unlike other researchers (e.g., in the field of sourcing research; Wineburg 1991; Brante and Strømsø 2018; see also Schoor et al. in this volume), we followed an approach of employing different types text materials and procedures for the assessment of heuristic and systematic components of scientific literacy. An advantage of this approach is that it allows analyzing relationships between these components. The correlation between epistemic-systematic competencies and epistemic-heuristic competencies was significant, but moderate, $r = .58$, $p < .05$ (von der Mühlen et al. 2016). It thus seems that scientific literacy is composed of distinguishable facets (Britt et al. 2014).

The Argument Structure Test and the Argument Judgement Test, and the corresponding trainings, were constructed as research tools. Nevertheless, the findings based on these tests and trainings indicate that they could also be used for practical purposes. The reliabilities of the tests are too low to warrant responsible individual-level selection decisions, but high enough to detect deficits on the group-level or to evaluate university courses. With regard to the trainings, our studies demonstrate that even short-term interventions can improve students' epistemic-systematic competencies. From a practical perspective, more extensive interventions are needed to ensure sustainable effects. For example, our trainings could be used as a starting point of regular reading courses, to familiarize students with the structural components of informal arguments and common argumentation fallacies. Subsequent sessions could then be devoted to the analysis and discussion of the way arguments are typically laid out in concrete exemplars of academic literature from the respective domain, to forge links between content knowledge, genre knowledge, and knowledge about argumentation. Such courses would not only provide students with disciplinary content knowledge, but also enable them to understand, and evaluate, the arguments on which this knowledge rests.

References

Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *I-S-T 2000 R – Intelligenz-Struktur-Test 2000 R* [Intelligence Structure Test 2000 R]. Göttingen, Germany: Hogrefe.

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist, 49*, 13–35.

Bazerman, C. (1985). Physicists reading physics: Schema-laden purposes and purpose-laden schema. *Written Communication, 2*, 3–23.

Berkenkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/ culture/ power*. Hillsdale, NJ: Erlbaum.

Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review, 30*, 773–799.

Britt, M. A., & Larson, A. A. (2003). Constructing representations of arguments. *Journal of Memory and Language, 48*, 794–810.

Britt, M. A., Richter, T., & Rouet, J. F. (2014). Scientific literacy: The role of goal-directed reading and evaluation in understanding scientific information. *Educational Psychologist, 49*, 104–122.

Brunmair, M. & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin*. doi: 10.1037/bul0000209

Dauer, F. W. (1989). *Critical thinking: An introduction to reasoning*. New York, NY: Oxford University Press.

Dillon, A. (1991). Reader's models of text structures: The case of academic articles. *International Journal of Man-Machine Studies, 35*, 913–925.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58.

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests* [Introduction to the theory of psychological testing]. Bern: Huber.

Goldman, S. R., & Bisanz, G. L. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J.A. León, & A.C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 417–436). Mahwah, NJ: Erlbaum.

Greving, S. & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology, 9*:2412. https://doi.org/10.3389/fpsyg.2018.02412

Ichikawa, J. J., & Steup, M. (2018). The analysis of knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* [Online Document]. Retrieved from https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/

Karpicke, J. D., & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38*, 116–124.

Larson, A. A., Britt, M. A., & Kurby, C. A. (2009). Improving students' evaluation of informal arguments. *Journal of Experimental Education, 77*, 339–366.

Lorch, R. F., Lorch, E. P., & Klusewitz, M. A. (1993). College students' conditional knowledge about reading. *Journal of Educational Psychology, 85*, 239-252. http://dx.doi.org/10.1037/0022-0663.85.2.239

Maier, J., & Richter, T. (2013). Text-belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction, 31,* 151–175.

Münchow, H., Richter, T., von der Mühlen, S., & Schmid, S. (2019). The ability to evaluate arguments in scientific texts: Measurement, cognitive processes, nomological network and relevance for academic success at the university. *British Journal of Educational Psychology, 89*, 501–523. https://doi.org/10.1111/bjep.12298

Münchow, H., Richter, T., von der Mühlen, S., Schmid, S., Bruns, K. & Berthold, K. (in press). Verstehen von Argumenten in wissenschaftlichen Texten: Reliabilität und Validität des Argumentstrukturtests (AST) [Comprehension of arguments in scientific tests: Reliability and validity of the argument structure test]. *Diagnostica*.

Neisser, U. (1967). *Cognitive Psychology*. New York, NY: Psychology Press.

Norris, S. P., Phillips, L.M., & Korpan, C.A. (2003). University students' interpretation of media reports of science and its relationship to background knowledge, interest, and reading difficulty. *Public Understanding of Science, 1*2, 123–145.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science, 328*(5977), 463–466. https://dx.doi.org/10.1126/science.1183944

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), Dual process theories in social psychology (pp. 41–72). New York: Guilford Press.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407.

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology, 97*, 70-80. https://doi.org/ 10.1037/0022-0663.97.1.70

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 352–387.

Richter, T. (2003). *Epistemologische Einschätzungen beim Textverstehen* [Epistemic validation in text comprehension]. Lengerich: Pabst.

Richter, T. (2011). Cognitive flexibility and epistemic validation in learning from multiple texts. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexiblity* (pp. 125–140). Berlin: Springer.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A Two-step Model of Validation. *Educational Psychologist, 52*, 148-166. http://dx.doi.org/10.1080/00461520.2017.1322968

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning, 5*, 47–65.

Scharrer, L., Stadtler, M., & Bromme, R. (2014). You'd better ask an expert: Mitigating the comprehensibility effect on laypeople's decisions about science-based knowledge claims. *Applied Cognitive Psychology*, *28*, 465–471.

Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction, 17*, 773–785.

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction, 20,* 192–204.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

von der Mühlen, S., Richter, T., Schmid, S., & Berthold, K. (2018). How to improve argumentation comprehension in university students: Experimental test of a training approach. *Instructional Science, 47*, 215–237.

von der Mühlen, S., Richter, T., Schmid, S., Schmidt, L. M., & Berthold, K. (2016a). The use of source-related strategies in evaluating multiple psychology texts: A student-scientist comparison. *Reading and Writing, 8*, 1677–1698.

von der Mühlen, S., Richter, T., Schmid, S., Schmidt, L. M., & Berthold, K. (2016b). Judging the plausibility of arguments in scientific texts: A student-scientist comparison. *Thinking & Reasoning, 22*, 221–246.

Voss, J. F., Fincher-Kiefer, R., Wiley, J., & Silfies, L. N. (1993). On the processing of arguments. *Argumentation, 7*, 165–181.

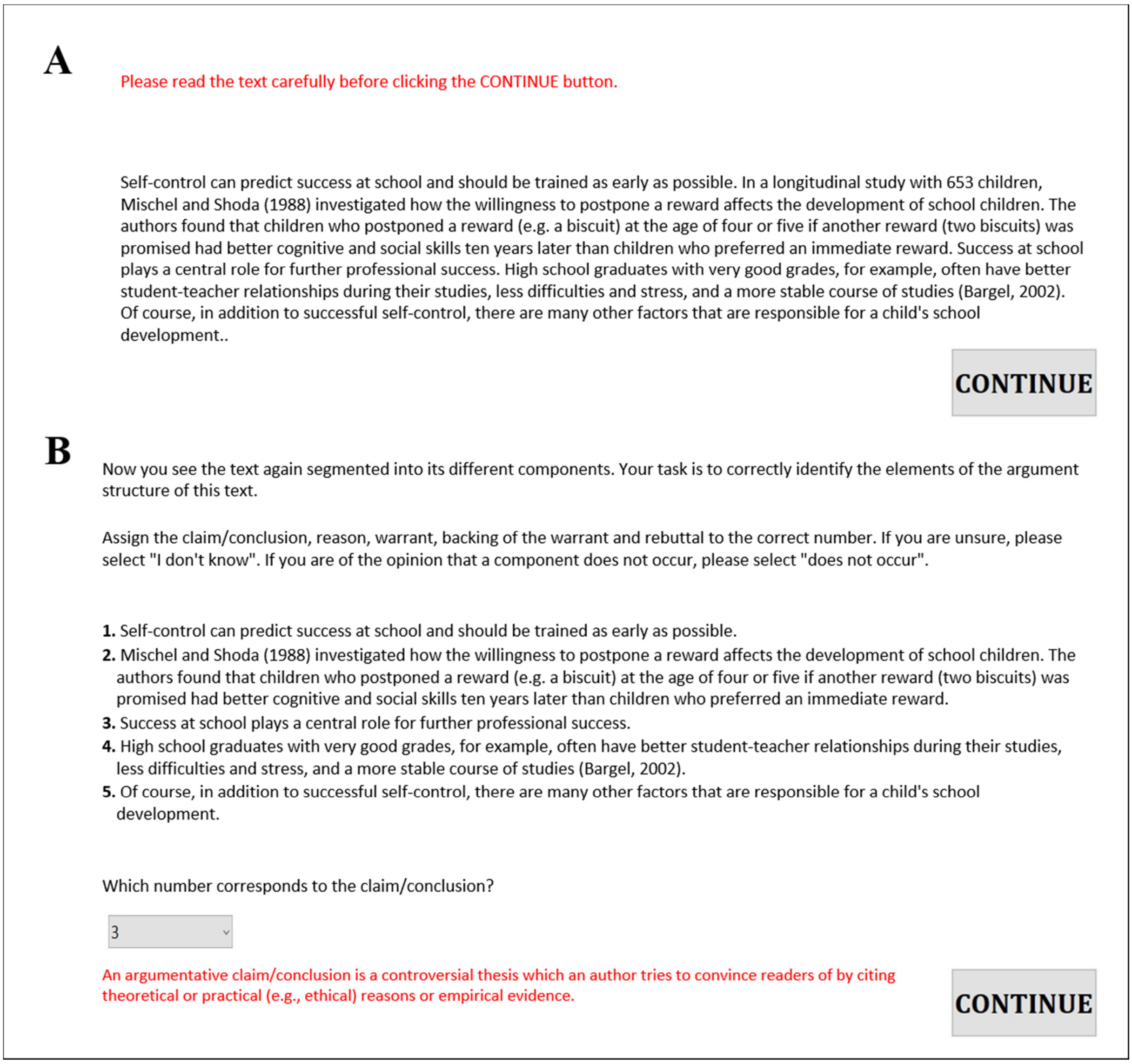
Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.), *Handbook of research in teaching* (pp. 315–327). New York, NY: Macmillan.

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*, 73–87.

Zimmerman, C., Bisanz, G. L., Bisanz, J., Klein, J. S., & Klein, P. (2001). Science at the supermarket: A comparison of what appears in the popular press, expert's advice to readers, and what students want to know. *Public Understanding of Science, 10*, 37–58.

Processing Mode	**Processing Goal**	
	Receptive	Epistemic
Systematic	e.g., organisation	e.g., evaluation of the consistency of arguments
Heuristic	e.g., scanning for certain information	e.g., using source information

Figure 1.1 Overview of taxonomy of reading and learning strategies required to deal with academic literature.



A

Please read the text carefully before clicking the CONTINUE button.

Self-control can predict success at school and should be trained as early as possible. In a longitudinal study with 653 children, Mischel and Shoda (1988) investigated how the willingness to postpone a reward affects the development of school children. The authors found that children who postponed a reward (e.g. a biscuit) at the age of four or five if another reward (two biscuits) was promised had better cognitive and social skills ten years later than children who preferred an immediate reward. Success at school plays a central role for further professional success. High school graduates with very good grades, for example, often have better student-teacher relationships during their studies, less difficulties and stress, and a more stable course of studies (Bargel, 2002). Of course, in addition to successful self-control, there are many other factors that are responsible for a child's school development..

CONTINUE

B

Now you see the text again segmented into its different components. Your task is to correctly identify the elements of the argument structure of this text.

Assign the claim/conclusion, reason, warrant, backing of the warrant and rebuttal to the correct number. If you are unsure, please select "I don't know". If you are of the opinion that a component does not occur, please select "does not occur".

1. Self-control can predict success at school and should be trained as early as possible.
2. Mischel and Shoda (1988) investigated how the willingness to postpone a reward affects the development of school children. The authors found that children who postponed a reward (e.g. a biscuit) at the age of four or five if another reward (two biscuits) was promised had better cognitive and social skills ten years later than children who preferred an immediate reward.
3. Success at school plays a central role for further professional success.
4. High school graduates with very good grades, for example, often have better student-teacher relationships during their studies, less difficulties and stress, and a more stable course of studies (Bargel, 2002).
5. Of course, in addition to successful self-control, there are many other factors that are responsible for a child's school development.

Which number corresponds to the claim/conclusion?

3

An argumentative claim/conclusion is a controversial thesis which an author tries to convince readers of by citing theoretical or practical (e.g., ethical) reasons or empirical evidence.

CONTINUE

Figure 1.2 Example item for the Argument Structure Test. (A) Argument presented as a continuous text. (B) Argument separated by sentences (translated from German). (Figure adapted from Münchow et al., in press)

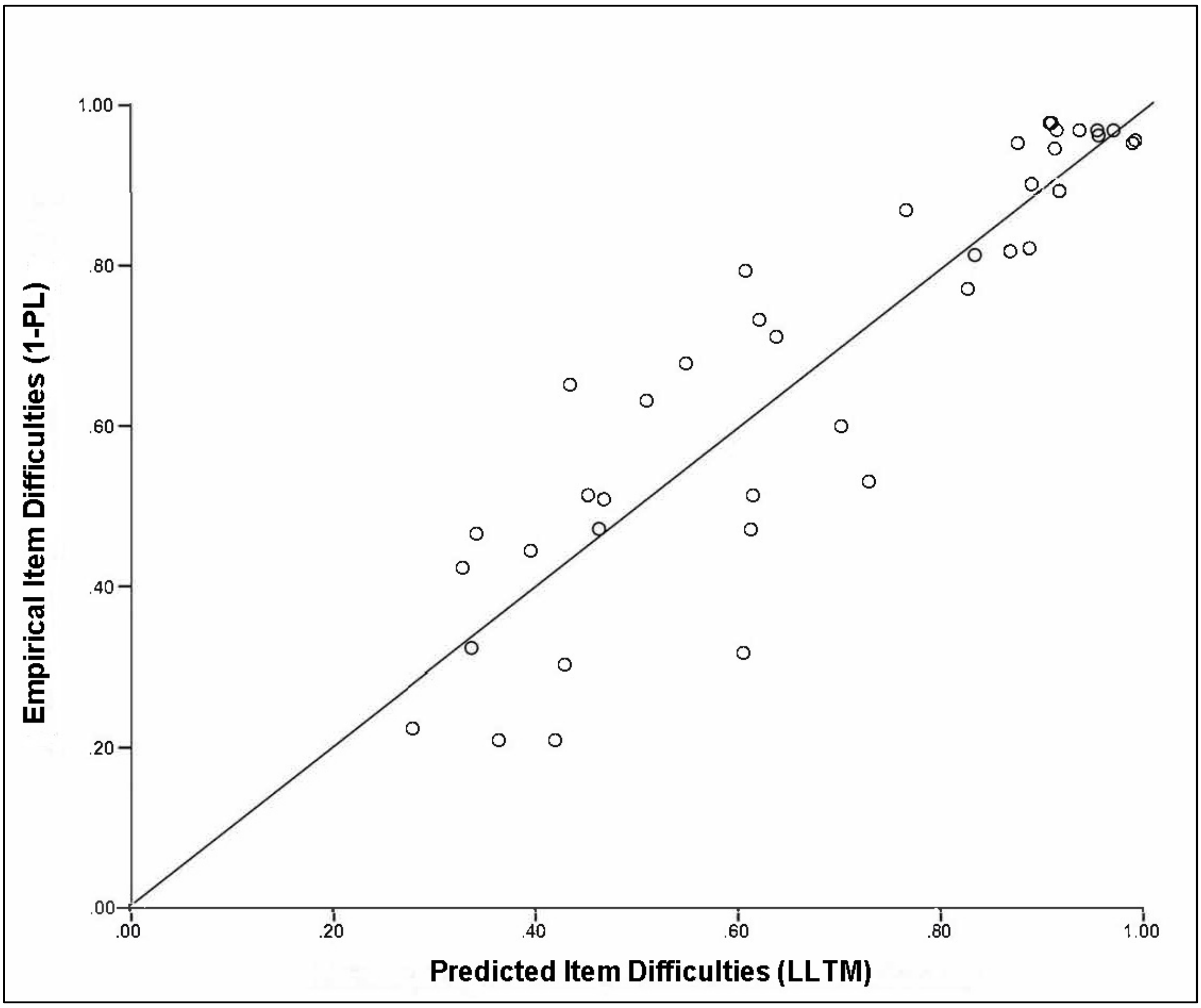


Figure 1.3 Scatterplot of predicted and empirical item difficulties for the Argument Structure Test. (Figure adapted from Münchow et al. in press, p. 3)

A

The construct of inherited nicotine sensitivity seems to play a central role here. This construct refers to the fact that some people react more strongly to nicotine because they are more sensitive to nicotine.

Press the P key if the sentence seems plausible, or the Q key if the reasoning does not seem plausible.

“Q”
Implausible

“P”
Plausible

B

The construct of inherited nicotine sensitivity seems to play a central role here. This construct refers to the fact that some people react more strongly to nicotine because they are more sensitive to nicotine.

You have declared the above sentence implausible. Please select the argumentation error you think is involved.

- Circular Reasoning
  [The attempt to prove the correctness of a premise with the help of a (logical) conclusion drawn from this premise. The premises shall prove the conclusion and at the same time the conclusion shall prove the premise.]
- Classical False Conclusion - Overgeneralization
  [A premise is followed by a false, hasty conclusion by generalizing or overrating results.]
- Classical False Conclusion – Contradiction
  [A premise is followed by an incompatible conclusion.]
- Wrong Example
  [A false or inappropriate example is cited as evidence for an allegation.]
- Wrong Dichotomy
  [A contradiction is suggested, but it's not really a contradiction.]
- I don’t know.
- I was wrong, there is no error.
- None of the above mentioned errors, but

NEXT

Figure 1.4 Example item for (A) Part 1 and (B) Part 2 of the Argument Judgement Test (translated from German). (Figure adapted from Münchow et al. 2019, p. 6)

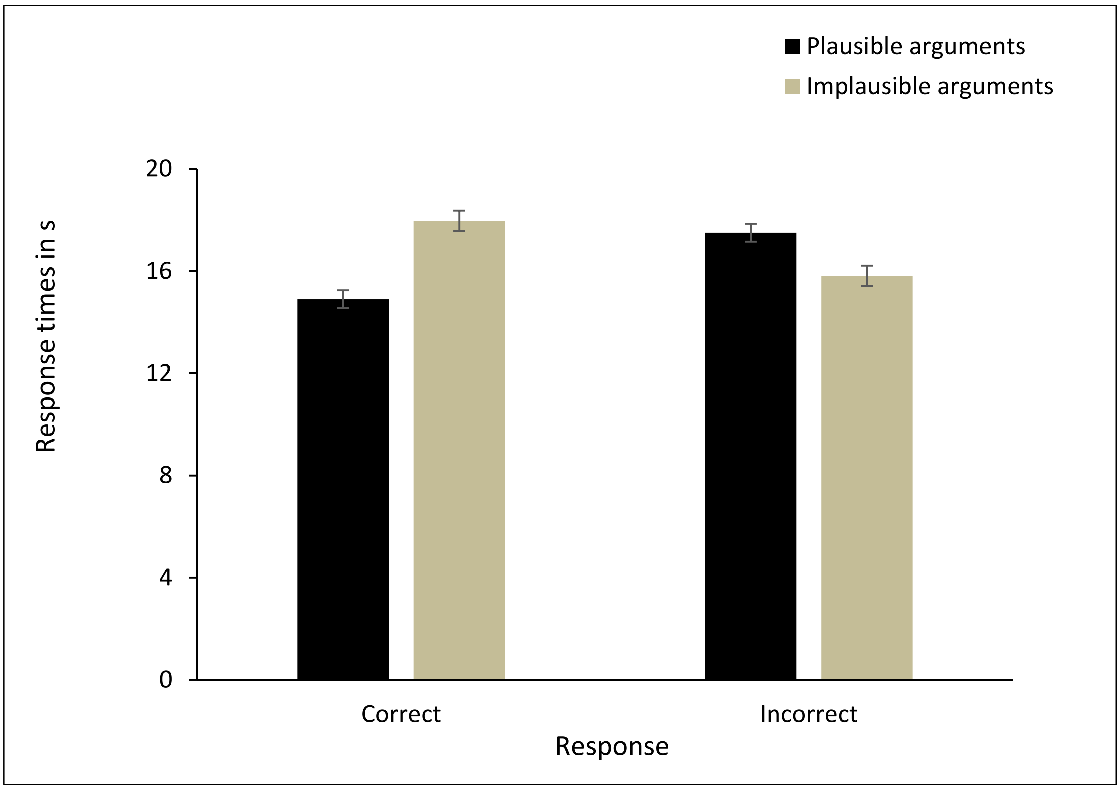


Figure 1.5 Response times (in seconds) in Part 1 of the Argument Judgement Test for plausible versus implausible arguments and response accuracy. (figure adapted from Münchow et al. 2019)

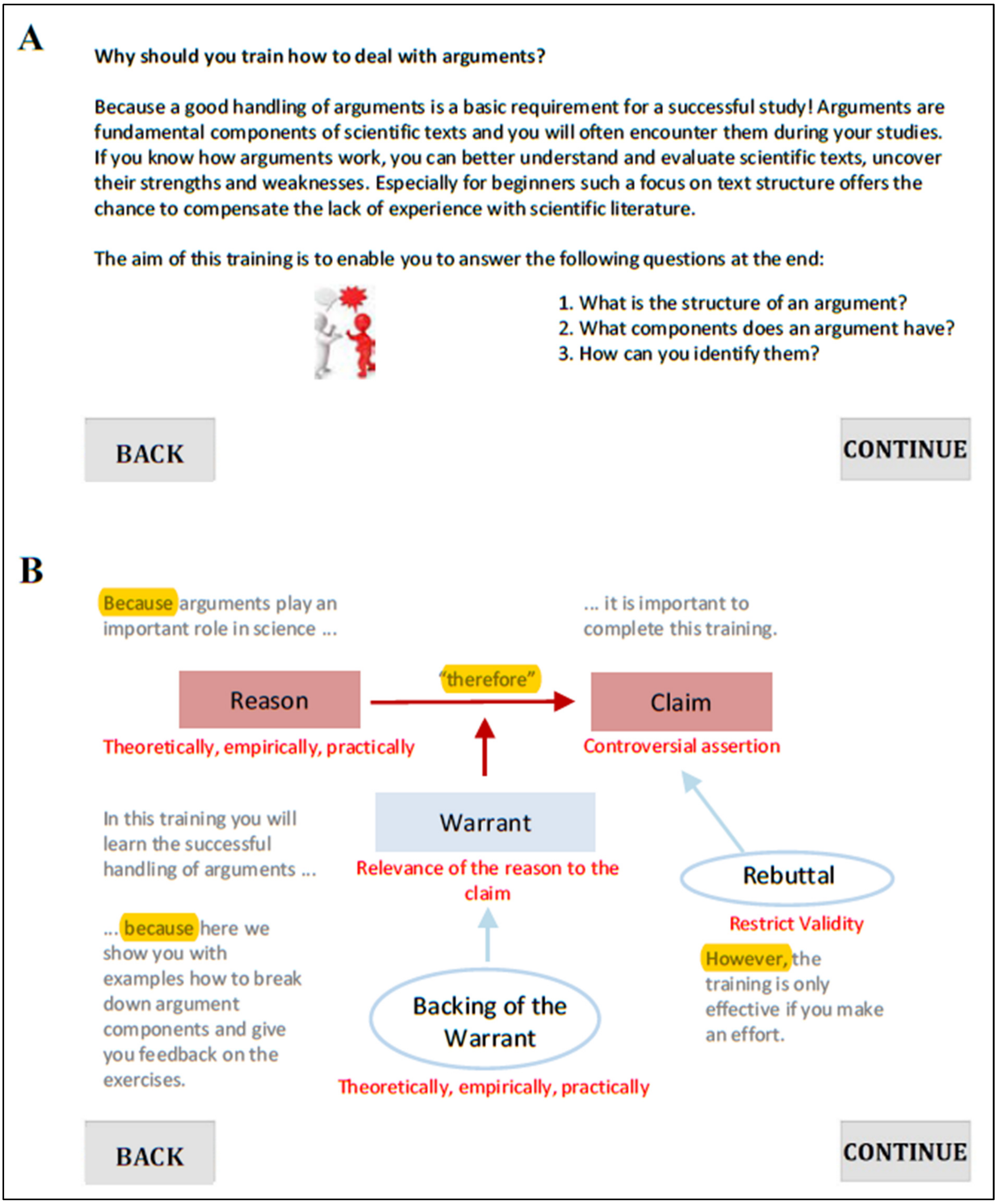


Figure 1.6 Example pages of the Argument Structure Training showing (A) the goals of the training and (B) the Toulmin model of argumentation (1958) and typical linguistic connectors.