Instructional Science (2019) 47:257–277
https://doi.org/10.1007/s11251-018-9478-9

The concreteness of titles affects metacognition and study motivation

Marie Lippmann[1] · Neil H. Schwartz[1] · Neil G. Jacobson[2] · Susanne Narciss[3]

Received: 22 May 2018 / Accepted: 13 November 2018 / Published online: 30 November 2018

Abstract
Two experiments investigated the extent to which the concreteness of titles affects metacognitive text expectations, study motivation, and comprehension test performance. Sixty-three American and 61 German students were presented with three titles (either concrete or abstract), based upon which the students estimated their expected ease-of-comprehension, and the expected interestingness, of three expository texts. Students also reported how motivated they were to study the texts. The students then studied the texts and completed comprehension tests. The results revealed that students expected texts with concrete (as opposed to abstract) titles to be easier to comprehend and more interesting, and were more motivated to study those texts. Structural Equation Modelling revealed that the effects of titles on reported study motivation were mediated by expected interestingness. In addition to that, expected interestingness and reported study motivation were partially mediated by expected ease-of-comprehension. Comprehension test performance was not affected. The results provide robust evidence for positive motivational effects of concrete titles. More specifically, the results indicate that concrete titles—which are specific and easy to imagine—promote students' motivation to study expository texts by encouraging the students to expect that they will find the texts interesting, and that they will be able to understand the texts.

Keywords Titles · Concreteness · Metacognition · Motivation · Text comprehension

✉ Marie Lippmann
mlippmann@csuchico.edu

Neil H. Schwartz
nschwartz@csuchico.edu

Neil G. Jacobson
ngjacobs@usc.edu

Susanne Narciss
Susanne.Narciss@tu-dresden.de

[1] Department of Psychology, California State University Chico, 100 West First Street, Chico, CA 95929, USA

[2] Rossier School of Education, University of Southern California, 3470 Trousdale Pkwy, Los Angeles, CA 90089, USA

[3] Psychology of Learning and Instruction, Technische Universität Dresden, Zellescher Weg 17, 01062 Dresden, Germany

Introduction

Two experiments investigated the extent to which the concreteness of titles affects metacognitive and motivational aspects of learning from expository texts. Expository texts are central components of most educational settings, and highly relevant to children's learning, adult education, vocational education, self-regulated learning, and life-long learning (Wiley et al. 2005; Coiro 2003; Lin and Zabrucky 1998). However, it is challenging for students to understand expository texts (Graesser 2007; Leopold and Mayer 2015; Best et al. 2005), and students often have difficulties monitoring their own learning process (Weaver 1990; Glenberg et al. 1982; Maki and Berry 1984). To empower students to effectively learn from expository texts, it is important to identify and understand text components that promote (or inhibit) text comprehension. The first text component that students encounter when they approach an expository text is its title—and titles can be more or less concrete. Imagine you were a student and your task was to study a text about airplanes, to learn how they function and work. Would it matter to you if the text you were about to study was entitled 'Getting Airplanes off the Ground' or 'Aerodynamic Laws of Ascension'? Would the concreteness of the title affect your metacognitive estimate of how easily you will understand the text? Would it affect how interesting you expect the text to be, and how motivated you are to study it? And would your expectations and your motivation affect your learning? To our knowledge, there are no investigations that have addressed the metacognitive and motivational effects of titles by examining their level of concreteness relative to students' text expectations, study motivation, and learning outcomes. The aim of the present investigation is thus to determine how the concreteness of titles affects expected ease-of-comprehension and interestingness, reported study motivation, and comprehension test performance—with the overarching goal of deriving practical implications for the design of titles across a variety of educational contexts.

Cognitive functions of titles

Traditionally, titles have been investigated with regard to their effects on the cognitive processing of text information (for an overview see Filippatou and Pumfrey 1995). Titles were typically varied either in terms of their absence or presence (Ausubel 1968; Bransford and Johnson 1972; Dooling and Mullet 1973; Schallert 1976; Schwartz and Flammer 1981), their location (before vs. after a text; Bransford and Johnson 1972; Dooling and Mullet 1973), and/or their level of correspondence with information in a text (integrative vs. non-integrative; Arnold and Brooks 1976; Schallert 1976; Kozminsky 1977; Schwartz and Flammer 1981; Yuill and Joscelyn 1988; Shepherd 1990; Filippatou and Pumfrey 1995; Lorch and Lorch 1996; Ritchey et al. 2008). Typical dependent measures included the recall and/or comprehension of entire passages or particular information within a passage. The results of these studies showed that titles which are presented prior to a text (a) activate relevant context information (Bransford and Johnson 1972), (b) direct attention and affect which information learners focus on (Lorch and Lorch 1996; Ritchey et al. 2008), (c) bias comprehension towards certain themes in a text (Schallert 1976; Kozminsky 1977), and (d) promote overall comprehension—particularly if the titles refer to the main ideas of their corresponding texts (Arnold and Brooks 1976; Yuill and Joscelyn 1988).

The concreteness of titles

The underlying rationale for the cognitive functions of titles is schema-theoretic in nature and proposes that titles activate relevant prior knowledge in the context of which the following text information is interpreted and remembered (for an overview see Filippatou and Pumfrey 1995). Whenever that activated context information elicits mental imagery, titles can be viewed not only as advance organizers but also as conceptual pegs, which conjure up mental images to which the following information then gets attached (Anderson et al. 1977; Sadoski et al. 2000). This view on titles stems from theories of text comprehension rooted in dual coding theory (Paivio et al. 1968; Sadoski and Paivio 2004, 2013; Schnotz and Bannert 2003; Schnotz 2005). Dual-coding oriented theories of text comprehension suggest that learners process written text in two code systems: the verbal/descriptive, and the non-verbal/depictive system (Sadoski and Paivio 2004, 2013; Schnotz and Bannert 2003; Schnotz 2005). Concrete language, as opposed to abstract, refers to tangible entities that can be perceived with the senses (Glanzer and Bowles 1976; Gernsbacher 1984; Hoffman 2016). Because concrete language relates to sensory, non-verbal experience, it is richer in referential connections between the verbal and non-verbal systems, and therefore more likely to elicit mental imagery than abstract language (Sadoski and Paivio 2004, 2013; Sadoski et al. 1993, 2000). As an example, conjuring up a mental image for the concrete phrase "seeing a green apple" is much easier than for the abstract phrase "knowing a common law".

From this point of view, Sadoski et al. (2000) proposed that concrete titles would promote text comprehension and recall by encouraging learners to encode the following information both verbally and non-verbally. The researchers tested this hypothesis in a set of experiments in which learners studied twenty-four short texts which were accompanied with either concrete or abstract titles. The results showed that concrete titles fostered the recall of literary stories and narratives, but not that of persuasive or expository texts (Sadoski et al. 2000). From our perspective, these mixed findings are difficult to interpret because the experiments focused solely on the cognitive effects of titles—as was the case in previous studies (Filippatou and Pumfrey 1995; Lorch and Lorch 1996; Ritchey et al. 2008).

Metacognitive and motivational functions of titles

More recent research has paid attention not only to the cognitive but also to the metacognitive and motivational functions of titles. Shimada (2016), for example, showed that titles which are perceived as easy to comprehend promote students' motivation to study expository text materials on disaster prevention. Because concrete titles are easy to envision, they are typically perceived to be easier to comprehend than abstract titles– even if both types of titles are equally familiar. The title pairs "How a Plane Flies" (concrete) versus "The Laws of Lift" (abstract), "Buffeting Jolts" (concrete) versus "Force Transfer" (abstract), and "Countertop Gadgets (concrete) versus "Domestic Devices" (abstract), for example, are perceived as equally familiar by students, but the concrete titles are perceived to be easier to comprehend (Sadoski et al. 2000).

Concrete titles are also perceived as more interesting than abstract titles, and this effect is partially mediated by ease-of-comprehension (Sadoski et al. 2000). On the one hand, concrete titles are interesting because the mental imagery they elicit is engaging

and involving (Sadoski et al. 2000). On the other hand, concrete titles are easier to comprehend than abstract titles, and comprehensible information is typically perceived as more interesting than information which is difficult to understand (Schraw et al. 1995; Schraw and Lehman 2001). The perceived interestingness of learning material is an important source of study motivation—particularly during difficult learning activities, such as studying expository texts (Fulmer et al. 2015; Krapp 1999; Hidi et al. 2004).

Concreteness in linguistic and cultural context

Concrete titles are likely to affect learners from varying linguistic and cultural backgrounds in similar ways because the processing of language in verbal and non-verbal systems is considered to be a fundamental and universal aspect of human cognition (Sadoski and Paivio 2013). Binder et al. (2009), for example, demonstrated in a large meta-analysis that the processing of concrete words includes sensory non-verbal experience (i.e. mental imagery), while the processing of abstract words depends primarily on verbal information. These results were supported by neuropsychological findings from different countries, including the US, the UK, and Germany (Binder et al. 2009; Wang et al. 2010; Hoffman 2016). Following up on these notions, the present investigation was conducted first in the US, and then again in Germany, with the intent to explore the generalizability of potential concreteness-effects of titles across two linguistically and culturally similar contexts. The linguistic and cultural contexts of the US and Germany were selected for two reasons. First, American English and German both belong to the Anglo-Frisian subgroup of West Germanic languages (Durrell 2006). As such, they share similar syntactic and semantic features (König and van der Auwera 1994). The second reason pertains to the close relationship between the academic cultures in the US and Germany. In both academic cultures, expository texts are a main medium for conveying information within educational contexts, and students are used to studying expository texts (Clyne 1987).

Research questions

Despite the ample empirical evidence for the potential metacognitive and motivational benefits of concrete titles, there are no studies that have systematically investigated the extent to which concrete versus abstract titles affect expected ease-of-comprehension, interestingness, study motivation, and learning outcomes. Based on the theoretical and empirical background outlined above, we raise the following research questions: (1) How does the concreteness of titles (Concrete vs. Abstract) affect learners' expectations of how easily they will understand expository texts?; (2) How does the concreteness of titles (Concrete vs. Abstract) affect learners' expectations of how interesting expository texts will be?, (3) How does the concreteness of titles (Concrete vs. Abstract) affect learners' motivation to study expository texts?; (4) How does the concreteness of titles (Concrete vs. Abstract) affect learning outcomes?; (5) What are the relationships between expected ease-of-comprehension, interestingness, reported study motivation, and learning outcomes? (6) Are effects of titles comparable in US American and German learners?

Hypotheses

Concrete titles are perceived as easier to comprehend and more interesting than abstract titles (Sadoski et al. 2000). We therefore hypothesize that learners will expect expository texts with concrete (as opposed to abstract) titles to be easier to comprehend (Hypothesis 1) and more interesting (Hypothesis 2). In line with dual-coding oriented theories of text comprehension (e.g. Schnotz 2005; Schnotz and Bannert 2003; Sadoski and Paivio 2013) and the empirical findings by Sadoski et al. (2000), we expect that expected ease-of-comprehension partially mediates the effects of titles on expected interestingness (Hypothesis 3). Ease-of-comprehension and interestingness are typically positively correlated to study motivation (e.g. Shimada 2016; Fulmer et al. 2015; Hidi et al. 2004; Krapp 1999), and concrete language is engaging and involving (e.g. Sadoski et al. 2000). We therefore expect learners to be more motivated to study expository texts with concrete (as opposed to abstract) titles (Hypothesis 4), partially mediated by expected interestingness (Hypothesis 5), and ease-of-comprehension—directly (Hypothesis 6a) and indirectly (Hypothesis 6b). Study motivation is a predictor for learning outcomes (e.g. Krapp 1999; Hidi et al. 2004). In addition, concrete titles foster text comprehension by functioning as conceptual pegs (Anderson et al. 1977; Sadoski et al. 2000). We therefore expect learners who study expository texts with concrete (as opposed to abstract) titles to achieve higher comprehension test scores (Hypothesis 7), partially mediated by study motivation (Hypothesis 8).

Methods

The hypotheses were investigated in two separate experiments—one conducted in the United States, the other one in Germany. The materials for the American experiment were produced first, and then systematically translated from English into German. Previous to the investigations, all materials were reviewed and approved by the Boards of Ethics at the universities at which the experiments were conducted. All data was collected anonymously, under consideration of the participants' privacy. All participants were informed about the terms of the data collection, and provided informed consent. After the pilot studies and the experiment, the participants received a detailed debriefing.

Participants

American participants

Sixty-three undergraduate students of a midsized university in the Western United States participated. Forty-three of the students were female, 19 male, and 1 student chose not to declare gender. The participants' ages ranged between 18 and 40 years ($M = 21.7$ years; $SD = 4.42$). All participants were native speakers of English, with a mean GPA of 3.0 ($SD = 0.5$). Forty-two percent of the participants studied psychology, 33% were enrolled in social studies, and 25% reported other majors, such as business, kinesiology, and theatre. The ethnic composition of the participant sample was: 1.6% Pacific Islander, 4.8% African American; 12.7% Asian American; 12.7% Hispanic, 61.9% Caucasian, and 3.2% of the participants chose not to declare ethnicity.

German participants

Sixty-one undergraduate psychology students of a large university in the East of Germany participated. Forty-two students were female, 19 male. The participants' ages ranged between 18 and 41 years (M = 23.05; SD = 4.96). All participants were native speakers of German, with a mean German GPA of 1.76 (SD = 0.66), which resembles an American GPA of 3.6. Sixty-five percent of the participants studied psychology, the remaining 35 percent were education majors.

Design

One factor (Concreteness of Titles) was varied on two levels (Concrete vs. Abstract) in a between-subjects-design.

Materials

The experimental materials were comprised of three expository texts (plus one practice text), each accompanied with one of two titles (Concrete vs. Abstract); a set of Likert-type rating scales; and three separate comprehension tests, each consisting of eight Meaning-Identification-Technique test sentences (MIT; Marchant III et al. 1988). All experimental materials were incorporated into a computer-based learning environment that was designed with the Study-2000-technology (Narciss et al. 2013).

Experimental texts

Four expository texts on scientific topics were derived from wiki pages and modified to suit the purpose of this study. The texts were examined by two science experts (i.e. one university professor of learning and instruction in science education and one university professor of physics) who approved of the text contents, and confirmed that domain-specific terminology was used correctly. Each text explained the design and functions of a different technological device used to overcome human boundaries (i.e. scuba regulators/airplanes/thermal imaging cameras/hearing aids). Each text was constructed using the revision techniques by Britton and Gülgöz (1991) as guiding principles to control for text coherence. Topics were chosen from different domains to avoid argument overlap, and texts were equated for word count (range 220–284) and readability (Flesh-Kincaid-grade level range 7.2–10.1). The readability range was chosen so that all texts could be well-understood by university undergraduates (i.e. Flesh-Kincaid-grade level $\leq$ 13).

The texts (without titles) were piloted on 18 undergraduate American students and 22 undergraduate German students of the universities at which the main studies were going to be conducted. Volunteers that took part in one of the pilot studies could neither take part in another pilot study, nor the main study. Participants rated each text for mental imagery, comprehensibility, interestingness, and content familiarity (as an indicator of prior knowledge) using a separate Likert-type rating scale for each rating, with scales ranging from 1 (very hard to form a mental image of this; very hard for me to understand; not interesting to me; not familiar to me in content) to 7 (very easy to form a mental image of this; very easy for me to understand; very interesting to me; very familiar to me in content). The rating scales were the same as used by Sadoski et al. (1993, 2000) and Sadoski (2001). The order of text presentation was randomized for each participant to prevent sequence effects.

Cronbach's α was at 0.5 for all ratings, which was low, but considered sufficient considering the small participant sample sizes (N = 18; 22) and the small number of texts (N = 4). Because of the small sample sizes in the pilot studies, the texts were compared with non-parametric Friedman tests of differences among repeated measures.

In both samples (American and German), the Friedman tests revealed no differences between the texts on scuba regulators, airplanes, and thermal imaging cameras with regard to the ratings (all $p > 0.05$). The texts elicited medium levels of mental imagery (M: 4.50–5.41; SD: 1.05–2.01), comprehensibility (M: 4.64–5.36; SD: 1.28–1.78), perceived interestingness (M: 3.73–4.18; SD: 1.47–2.13), and evoked low content familiarity (M: 2.50–3.45; SD: 1.14–1.89).

Low content familiarity was important to ensure that participants had the opportunity to learn something from the texts, and thus, to prevent ceiling effects in terms of text comprehension. In both samples, the text on hearing aids was rated slightly more familiar in content than the other three texts, and was therefore selected to function as the practice text.

Titles

Each of the three experimental texts was accompanied by a semantically concrete versus abstract title (*Getting Airplanes off the Ground* vs. *Aerodynamic Laws of Ascension*; *Breathing Underwater* vs. *Subaquatic Regulator Systems*; *Seeing Objects in the Dark* vs. *Thermal Variability Imaging*). All of the titles were integrative in that they referred to the main ideas, concepts, or themes of their corresponding texts (Filippatou and Pumfrey 1995). The concrete titles were based on phrases that directly referred to specific sensory experience (i.e. *breathing, seeing*, and *getting off the ground*). The abstract titles were based on intangible terms (*systems, laws, and variability*) in combination with general technical terms (*subaquatic*, *regulator, aerodynamics,* and *thermal*), and were mostly unrelated to sensory experience.

The practice text was presented without a specific title (i.e. the text was simply called 'Practice Passage') to avoid priming effects. According to Glanzer and Bowles (1976) and Paivio et al. (1968), concrete words have a higher probability to occur in printed text (i.e. a higher word frequency) than abstract words, and high-frequent words tend to be shorter than low-frequent words (Ozuru et al. 2012). The abstract titles are therefore on average three syllables longer than the concrete titles, while having an almost identical word count. Both title versions for each text were anchored to the first text sentence by a two- or three-content-word overlap in order to establish direct verbal associations between text and titles (Sadoski et al. 2000). This was also important because learners expect the most important information to be mentioned first in a text (Kieras 1981), and initially-mentioned information influences learners' perceptions of whole texts (Kieras 1980). A content word that was used in one title was not used in any other title. Pilot studies were conducted to determine whether the manipulation of semantic concreteness produced differences in mental imagery, ease-of-comprehension, and interestingness, as proposed by Sadoski et al. (1993, 2000) and Sadoski and Paivio (2004, 2013).

The titles (without texts) were piloted on 24 other undergraduate American and 22 other undergraduate German students from the same universities as before, using the same 7-point Likert-type rating scales that were used for piloting the texts. The order of title presentation was randomized for each participant. Cronbach's α ranged between 0.6 and 0.7 for the ratings in both samples (American and German). In order to determine whether the title-pairs (Concrete vs. Abstract) for each text differed significantly,

we compared the titles with non-parametric Wilcoxon tests because of the small sample sizes (N = 24; 22). The results showed significant differences between all of the title pairs (all $p \leq 0.01$). The concrete titles were easy to imagine (M: 6.21–6.42; Mdn: 7.00; SD: 0.97–1.56), highly comprehensible (M: 6.63–6.71; Mdn: 7.00; SD: 0.46–1.02), interesting (M: 5.04–6.00; Mdn: 5.00–7.00; SD: 1.10–1.53), and rather familiar in content (M: 5.41–6.37; Mdn: 6.00–7.00; SD: 0.96–1.87). The abstract titles ranged within a medium level of mental imagery (M: 3.92–4.21; Mdn: 4.00–4.50; SD: 1.74–2.19), comprehensibility (M: 3.92–4.58; Mdn: 4.00–5.00; SD: 1.69–1.84); interestingness (M: 4.13–4.46; Mdn: 4.00–5.00; SD: 1.84–1.92); and were not as familiar in content (M: 2.37–2.83; Mdn: 2.00; SD: 1.48–1.78).

Self-report measures

For the main experiments, ratings on expected ease-of-comprehension (EOC), interestingness (INT), and motivation to engage in a text (MOT) were assessed using separate 7-point Likert-type rating scales for each rating, and each text. The scales for each of the ratings ranged from 1 (= very hard to understand/not at all interesting/not at all motivated) to 7 (= very easy to understand/highly interesting/highly motivated). Participants were specifically asked: *How easy do you think it will be for you to understand a passage titled…? How interesting do you expect a passage titled…to be? How motivated are you to study a passage titled…?* We determined the reliabilities of the ratings, and obtained the following values (Cronbach's α): EOC $\alpha = 0.84$; INT $\alpha = 0.71$; MOT $\alpha = 0.77$ (American sample), and EOC $\alpha = 0.91$; INT $\alpha = 0.71$; MOT $\alpha = 0.84$ (German sample).

Comprehension tests

To assess conceptual text understanding we constructed three separate comprehension tests (one for each text), using the Meaning-Identification-Technique (MIT; Marchant III et al. 1988). In MIT, comprehension is assessed at the level of the propositional text base (for levels of text comprehension see Kintsch 2004) by exposing participants to test sentences to which they respond with either Yes or No, based on whether they think the sentences accurately describe information they encountered in a previously studied text. There are two types of test sentences in MIT: Paraphrases and meaning changes of paraphrases. Paraphrases carry the same meaning as statements in the text, and the correct response to these test sentences would therefore be YES. Meaning changes of paraphrases do not carry the same meaning as statements in the text, and should therefore be responded to with NO. To construct a comprehension test with the MIT, all original sentences from a text are first paraphrased by changing as many words as possible, without making the sentences artificially difficult or incoherent (Marchant III et al. 1988). Then, half of these paraphrased sentences are slightly changed in their meaning so that they no longer corresponded with the original statements from the texts (Marchant III et al. 1988). These meaning changes typically include the substitution of concepts from the text with other, seemingly similar concepts, as shown in the following example from the text on thermal imaging cameras:

> *Original sentence from text* Infrared detector elements create a detailed temperature pattern.
> *Paraphrase* The generated heat pattern is produced by infrared light sensors.
> *Meaning change* The generated heat pattern is based on ultraviolet emissions.

Based on the item-analysis of a pilot study in which another 38 American undergraduate students took part, we selected eight test-sentences per text—half of them paraphrases, half of them meaning changes. The resulting three comprehension tests (Cronbach's alphas were 0.7 for each of the tests) only included items that ranged within a medium level of item difficulty (range: 0.55–0.84), and that contributed to the internal consistency of the respective comprehension test. We further ensured that all major concepts which were stated in the texts were also addressed in the comprehension tests, and that each comprehension test addressed a comparable amount of information stated in the first and the second half of the text on which the comprehension test was based. The finalized American MIT test was then systematically translated into German.

Procedure

After providing informed consent for the investigation, the participants in both samples (American and German) were randomly assigned to one of the title-conditions (Concrete vs. Abstract), and completed one practice trial to become familiar with the computer-based learning environment, and the demands of the learning tasks (i.e. the participants studied the practice text and completed a practice MIT-test pertaining to that text). For the three experimental texts, participants first only saw the title of the text they were about to encounter and, based on that title, rated their expected ease-of-comprehension (EOC), and the expected interestingness (INT) of the following text, as well as the motivation to study that text (MOT). Participants studied the text which followed the title, and then completed the MIT-test pertaining to that text. This procedure was repeated for the remaining two experimental texts. All tasks were self-paced. The order of text appearance was counterbalanced. In a last step, participants completed a demographics questionnaire, and then received a detailed debriefing about the experiment.

Results

For each participant, the means of the ratings and comprehension test outcomes were computed. The data were analyzed using SPSS 24. The ANOVAs were Bonferroni-corrected at $p < 0.0125$ for testing 4 main-effect-hypotheses on one sample. The SEM was conducted using AMOS for SPSS, Version 24.

Descriptive statistics

Table 1 shows the means, standard deviations, skewness, and kurtosis for the ratings of ease-of-comprehension (EOC), interestingness (INT), motivation (MOT), and for comprehension test performance (COMP), for both samples (American and German).

All skewness and kurtosis values were less than an absolute value of 3, indicating that normality for subsequent analysis could be assumed (Tabachnick and Fidell 1996). No outliers were identified in the data.

Table 1 Descriptive statistics

Sample	Measure	N	Minimum	Maximum	Mean	SD	Skewness		Kurtosis	
							Statistic	SE	Statistic	SE
American	EOC	63	1	7	3.86	1.49	0.08	0.30	−0.93	0.59
	INT	63	1	7	4.17	1.38	−0.09	0.30	−0.40	0.59
	MOT	63	1	7	3.67	1.37	−0.20	0.30	−0.80	0.59
	COMP	63	3.33	8.00	5.71	1.17	−0.17	0.30	−0.84	0.59
German	EOC	61	1	7	3.84	1.43	0.10	0.30	−1.20	0.60
	INT	61	1	7	3.74	1.41	−0.21	0.30	−0.79	0.60
	MOT	61	1	7	4.06	1.47	−0.21	0.30	−0.94	0.60
	COMP	61	5.33	7.67	6.66	0.63	−0.46	0.30	−0.48	0.60

EOC expected ease-of-comprehension, *INT* expected interestingness, *MOT* reported study motivation, *COMP* comprehension test performance

Correlations

Zero-order correlations were conducted in both samples (American and German) to determine whether the proposed relationships between the variables were present (see Table 2). Except for the relationships between comprehension test performance and most other variables, all proposed relationships were found (see Table 2).

Comparability of US and German data sets

The descriptive statistics (Table 1) and correlational patterns (Table 2) for the American and German data sets indicate a high degree of comparability between the two samples. To statistically test for potential differences, the two data sets were compared with independent

Table 2 Pearson correlations main experiments (American and German sample)

Sample	Variable	Title	EOC	INT	MOT	COMP
American	Title	1	0.73**	0.58**	0.47**	0.12
	EOC	0.73**	1	0.57**	0.42**	0.21
	INT	0.58**	0.57**	1	0.81**	0.07
	MOT	0.47**	0.42**	0.81**	1	0.01
	COMP	0.12	0.21	0.07	0.01	1
German	Title	1	0.50**	0.49**	0.36**	0.11
	EOC	0.50**	1	0.84**	0.63**	0.27*
	INT	0.49**	0.84**	1	0.76**	0.23
	MOT	0.36**	0.63**	0.76**	1	0.27*
	COMP	0.11	0.27*	0.23	0.27*	1

EOC expected ease-of-comprehension, *INT* expected interestingness, *MOT* reported study motivation, *COMP* comprehension test performance

*Significant at $p < 0.05$ (2-tailed)

**Significant at $p < 0.01$ (2-tailed)

samples t-tests for each of the four outcome measures. The confidence interval was adjusted to 99.98% for testing multiple hypotheses. There were no significant differences between the US American and the German sample in terms of expected ease-of-comprehension [t (122) = 0.05, p = 0.90], expected interestingness [t (122) = 1.69, p = 0.09], and motivation [t (122) = − 1.53, p = 0.12). The samples only differed with regard to comprehension test performance [t (122) = − 5.56, p < 0.01), with German learners solving on average one more test item correctly (M = 6.66; SD = 0.63) than American learners (M = 5.71; SD = 1.17). Because of the theoretically proposed and empirically evidenced comparability of both samples, the data sets were combined to increase power for further analysis. Follow-up analyses for performance data was conducted separately for the two samples.

Main effects

Expected ease-of-comprehension (EOC)

The ANOVA to test Hypothesis 1 revealed a large effect of Titles on EOC [F (1, 122) = 78.25; MS_{error} = 1.30; p < 0.001; partial η^2 = 0.39]. Students who were presented with concrete titles expected the subsequent texts to be easier to comprehend (M = 4.73; SD = 1.13) than students who were presented with abstract titles (M = 2.92; SD = 1.15).

Expected interestingness (INT)

The ANOVA to test Hypothesis 2 revealed a large effect of Titles on INT [F (1, 122) = 47.12; MS_{error} = 1.44; p < 0.001; partial η^2 = 0.27]: Students who were presented with concrete titles expected the following texts to be more interesting (M = 4.68; SD = 1.12) than students who were presented with abstract titles (M = 3.20; SD = 1.27).

Reported study motivation (MOT)

The ANOVAs to test Hypothesis 4 revealed a large effect of Titles on MOT [F (1, 122) = 25.14; MS_{error} = 1.72; p < 0.001; partial η^2 = 0.17]: Students who were presented with concrete titles reported to be more motivated to study the following texts (M = 4.43; SD = 1.21) than students presented with abstract titles (M = 3.25; SD = 1.41).

Comprehension test performance (COMP)

The ANOVAs to test Hypothesis 7 revealed no effect of Titles on COMP [F (1, 122) = 1.48; MS_{error} = 1.11; p = 0.22; partial η^2 = 0.12]. Learners in both title groups solved about the same amount of test sentences correctly ($M_{concrete}$ = 6.26; $SD_{concrete}$ = 1.03; $M_{abstract}$ = 6.06; $SD_{abstract}$ = 1.07). Because the two samples originally differed in comprehension test performance, the analysis was repeated separately for each sample, with the same main result: There was no effects of titles in either sample (American: [F (1, 61) = 0.92; MS_{error} = 1.38; p = 0.34; partial η^2 = 0.02]; German: [F (1, 59) = 0.77; MSerror = 0.40; p = 0.38; partial η^2 = 0.01]). Learners in both title groups solved about the same amount of test sentences correctly in either sample (American: $M_{concrete}$ = 5.85; $SD_{concrete}$ = 1.22; $M_{abstract}$ = 5.57; $SD_{abstract}$ = 1.13; German: $M_{concrete}$ = 6.72; $SD_{concrete}$ = 0.53; $M_{abstract}$ = 6.58; $SD_{abstract}$ = 0.72).

Mediation analyses with structural equation modelling (SEM)

The more complex mediating relationships between titles, ease-of-comprehension (EOC), interestingness (INT), motivation (MOT), and comprehension test performance (COMP)—as proposed in *Hypotheses 3, 5, 6*, and *8*—were investigated with Structural Equation Modelling (SEM). SEM was used because it allows to create and confirm specific path models with more flexibility than traditional methods, such as path analyses via regressions (Kline 2011).

While regression analyses imply statistical relationships between variables based on conditional expected values, SEM assumes functional relationships grounded in conceptual models (Bollen and Pearl 2012). The paths for the conceptual model in this investigation were defined based on the theoretically meaningful relationships described in the introduction. The resulting empirical model is displayed in Fig. 1. For ease of interpretation, all paths are displayed as standardized. Paths marked with ** are significant at $p < 0.01$, paths marked with * are significant at $p < 0.05$. The overall model fits the data well: χ^2 ($N = 124$) = 9.18; $p = 0.01$; GFI = 0.97; NFI = 0.97; RMSEA = 0.17). In line with the ANOVAs, the paths of the proposed main effects of titles on EOC and INT were significant at $p < 0.01$.

Partial mediation of expected interestingness (INT) by (title) through expected ease-of-comprehension (EOC)

To test Hypothesis 3, we determined the standardized indirect effect of Titles on INT at (0.63) (0.61) = 0.38. We tested the significance of this indirect effect using bootstrapping procedures. Unstandardized indirect effects were computed for each of 2000 bootstrapped samples, and the 95% confidence interval was computed by determining the indirect effects at the 2.5th and 97.5th percentiles. The bootstrapped unstandardized indirect effect was 1.06, and the 95% confidence interval ranged from 0.76, 1.38. The indirect effect of Titles on INT through EOC was statistically significant ($p < 0.01$), thereby supporting Hypothesis 3.

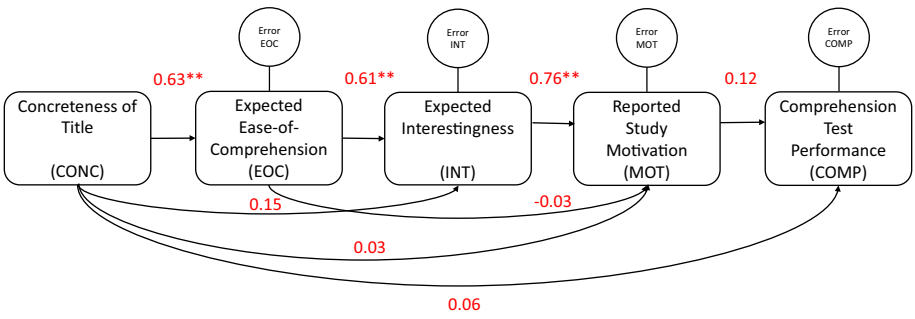

Fig. 1 Mediation analyses with structural equation modelling (SEM)

Partial mediation of study motivation (MOT) by (Title) through expected interestingness (INT) and expected ease-of-comprehension (EOC)

To test Hypotheses 5 and 6, we first determined the standardized indirect effect of Titles on MOT through INT and EOC at (0.63)(0.61)(−0.03)(0.76) = 0.38. Unstandardized indirect effects were computed for each of 2000 bootstrapped samples at a 95% confidence interval. The bootstrapped unstandardized effect was 1.09, and the 95% confidence interval ranged from 0.75, 1.40. The indirect effect of Titles on MOT was significant at $p < 0.01$. In a next step, we examined the relative contributions of EOC and INT on the mediation effect by determining the direct effect of INT on MOT, and the direct effect of EOC on MOT. The standardized direct effect of INT on MOT was 0.76. The bootstrapped unstandardized direct effect was 0.77, the 95% confidence interval ranged from 0.59, 0.90, and was significant at $p < 0.01$, thereby supporting Hypothesis 5. The standardized direct effect of EOC on MOT was −0.03. The bootstrapped unstandardized direct effect was −0.03, the 95% confidence interval ranged from −0.21, 0.15, and was not significant at $p = 0.76$, thereby rejecting Hypothesis 6a. For Hypothesis 6b, we determined the standardized indirect effect of EOC on MOT at (0.61)(0.76) = 0.46. The bootstrapped unstandardized effect was 0.45, at an interval range of 0.31–0.60. The indirect effect of EOC on MOT was significant at $p < 0.01$, thereby supporting Hypothesis 6 b.

Partial mediation of comprehension test performance (COMP) by title through study motivation (MOT)

The American and German samples differed significantly in comprehension test performance. In addition, comprehension test performance (COMP) did not correlate with any of the other variables in the American sample, and did not yield significant differences between the experimental groups in neither the American nor the German sample. We therefore computed the mediation analysis for Hypothesis 8 only for explorative purposes, and separately for the two samples. In the American sample, the standardized indirect effect of Titles on COMP was determined at (0.06)(0.15) = −0.02. The bootstrapped unstandardized effect was 0.15, at an interval range of −0.37 to 0.28. The indirect effect was not significant at $p = 0.64$, thereby rejecting Hypothesis 8. In the German sample, COMP correlated with MOT and EOC, and the path from MOT to COMP in the empirical SEM model reached significance at $p < 0.05$. The standardized indirect effect of Titles on COMP through MOT in the German sample was determined at (−0.01)(0.26) = 0.09. The bootstrapped unstandardized effect was 0.11, at an interval range of −0.01 to 0.29. The indirect effect of Titles on MOT did not reach significance ($p = 0.07$), thereby rejecting Hypothesis 8.

Discussion

The cognitive functions of titles are well-researched (e.g. Bransford and Johnson 1972; Filippatou and Pumfrey 1995; Lorch and Lorch 1996; Ritchey et al. 2008), yet little is known about the metacognitive and motivational effects of titles in instructional contexts (Shimada 2016). The primary goals of the present two experiments were thus to a) determine the extent to which concrete versus abstract titles affect metacognitive and

motivational expectations students generate with regard to expository texts, with a focus on expected ease-of-comprehension and expected interestingness; b) gain insight into the relationships between expected ease-of-comprehension, interestingness, and study motivation, c) determine if, and how, text expectations and study motivation are related to comprehension test performance; d) explore the comparability and generalizability of effects across two similar linguistic and cultural contexts (American and German), and d) derive practical implications for the design of titles for expository texts.

In line with our main hypotheses, the results of revealed that both American and German students expected texts with concrete (as opposed to abstract) titles to be easier to comprehend and more interesting, and reported to be more motivated to study those texts. Contrary to our hypotheses, the increased motivation to study texts with concrete titles did not result in increased comprehension test performance for these texts. Structural Equation Modelling (SEM) was applied to test a conceptual model of the relationships between expected ease-of-comprehension, interestingness, reported study motivation, and comprehension test performance. In line with our hypotheses, the concreteness of titles affected students' motivation to study expository texts, mediated by how interesting the students expected the texts to be. Expected interestingness and study motivation were partially mediated by the ease with which students expected to understand the texts (see Fig. 1).

Theoretical considerations, limitations, and implications

Dual-coding oriented theories of text comprehension and neuropsychological evidence (e.g. Sadoski and Paivio 2013; Schnotz 2005; and Bannert 2003; Binder et al. 2009; Hoffman 2016) suggest that concrete language is richer in referential connections between the verbal and non-verbal system, and therefore easier to mentally picture and comprehend, and perceived as more interesting, than abstract language (Sadoski and Paivio 2004, 2013; Sadoski et al. 2000). Theoretical approaches to motivation in text comprehension (e.g. Krapp 1999; Hidi et al. 2004) propose that the interestingness of learning materials is a driving force for study motivation (Hidi et al. 2004; Krapp 1999; Fulmer et al. 2015). In line with both theoretical perspectives, the students in our pilot studies found concrete titles relating to sensory experiences (such as *Breathing Underwater* or *Seeing Objects in the Dark)* easier to mentally picture and comprehend, and more interesting, than abstract titles based on intangible, rather technical terms (such as *Aerodynamic Laws of Ascension* or *Subaquatic Regulator Systems*). In line with empirical evidence for the universal role of language concreteness across linguistic contexts (Binder et al. 2009; Wang et al. 2010; Hoffman 2016), the same results were obtained for American and German participants.

For the main experiments, we proposed a conceptual model for the relationships between the concreteness of titles and students' expected ease-of comprehension, interestingness, study motivation, and comprehension test performance. Overall, this conceptual theoretical model was strongly supported by the data (Fig. 1)—particularly with regard to the mediating relationship between the concreteness of titles, expected text interestingness, and study motivation. This finding aligns with previous research showing that concrete information is often perceived as more interesting and motivating than information that is difficult to envision and understand (e.g. Sadoski et al. 2000; Schraw et al. 1995; Schraw and Lehman 2001).

With regard to the role of prior knowledge, it is important to note that effects may vary for learners with different levels of expertise. The pilot studies for the present experiments revealed low levels of prior knowledge in American and German

participants, and there were no outliers identified in terms of comprehension test performance in the main experiments. However, it is certainly possible that learners with high levels of prior knowledge might experience a trade-off for title concreteness and interestingness/motivation. That is, learners with a lot of prior knowledge might expect texts with concrete titles to be too easy and therefore not at all interesting, which could result in a decreased motivation to study texts with concrete as compared to abstract titles; an interesting hypothesis for future research.

A rather puzzling finding was the lack of association between comprehension test performance and the other variables. We attribute this lack of effects primarily to the employed comprehension measure, which was likely not sensitive enough to detect potential differences in learning outcomes between the experimental groups. In the Meaning-Identification-Technique (MIT; Marchant III et al. 1988), students identify whether a presented statement aligns with a statement they had encountered in a previously studied text. The presented statement is either a paraphrase (i.e. a correct statement) or a meaning-change of an original statement from the text (i.e. an incorrect statement). Even though we altered as many words as possible to create the paraphrases and meaning changes from the original statements, the resulting items still provided contexts which could be utilized by students to access their mental model of the texts (Bransford and Johnson 1972; Anderson et al. 1977; 2005). In free recall tasks—which are more typically employed as outcome measures in studies on titles (e.g. Dooling and Mullet 1973; Kozminsky 1977; Schwartz and Flammer 1981; Yuill and Joscelyn 1988), students have fewer cues available to access their mental text models. Free recall tasks may thus be more suitable to assess differences in text comprehension, based on titles—particularly when titles serve as the retrieval cues for the free recall tasks. A second consideration regarding the lack of effects on comprehension test performance pertains to the experimental setting and procedure. In both title-groups (Concrete vs. Abstract), students were aware that they were going to be tested on their knowledge after reading. In addition, the experiments were conducted in highly controlled lab settings, in presence of the experimenters. These two factors may have induced high levels of socially desirable study behavior in the participants, resulting in proper study behavior despite low levels of motivation in the abstract title-group (Richman et al. 1999). An interesting way to approach this issue would be an online replication of the present studies in which the participants are removed from the highly controlled lab environment, and less inclined to behave in socially desirable ways (Richman et al. 1999). A third consideration pertains to the experimental texts, which were comparatively short (220–284 words) and of medium concreteness, comprehensibility, and interestingness (see Tables 1 and 2). Because the texts did not fully align with either title condition (i.e. they were less concrete, comprehensible, and interesting than their titles in the concrete-title-condition, and more concrete, comprehensible, and interesting than their titles in the abstract-title-condition), students in either title-group may have adjusted their motivation levels in either direction during reading, based on the text characteristics. A way to address this issue in future research could include a slightly more complex 2×2-design in which both, titles and texts, are varied systematically in terms of their concreteness. However, in spite of the limitations in terms of comprehension test performance, the results of the present investigation substantially extend the current literature by helping us understand the processes by which titles affect student's text expectations in terms of ease-of-comprehension and interestingness, and how these expectations relate to study motivation.

Practical implications

Based on the mediating relationships between the concreteness of titles and students' text expectations and study motivation, the present investigation offer a number of practical implications for the design of titles in instructional contexts. The first implication pertains to the relationship between the concreteness of titles and expected ease-of-comprehension. In both samples (American and German), students expected texts with concrete titles to be easier to comprehend than texts with abstract titles. This is important because many students are in fact afraid of approaching expository texts, particularly if they know they are going to be tested on the text contents (Lin et al. 2000). Employing concrete as opposed to abstract titles will help to counteract that initial fear by increasing the students' expected ease-of-comprehension.

The second practical implication pertains to the relationships between the concreteness of titles and expected interestingness and study motivation. Students were more motivated to study expository texts with concrete (as opposed to abstract) titles, and this was in part due to the expectation that those texts would be more interesting. Studies over the past three decades have shown that students' academic motivation declines over the course of their educational paths (e.g. Hidi and Harackiewicz 2000; Anderman and Maehr 1994; Harter 1981). More specifically, students' interests and attitudes towards school decline over time, particularly with regard to scientific topics (e.g. Hidi and Harackiewicz 2000; Eccles and Wigfield 1992; Eccles et al. 1998). It is therefore imperative for educators, text book editors, publishers, and designers of online learning environments to employ instructional strategies that promote and cultivate study motivation—particularly with regard to topics that rely heavily on learning from expository texts as is the case in science education (e.g. Coiro 2003; Leopold and Mayer 2015). As the results of the present two investigations show, employing concrete as opposed to abstract titles has a strong potential to promote and maintain study motivation by increasing the expected interestingness of expository texts.

References

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research, 64,* 287–309.

Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. (1977). Two faces of the conceptual peg hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 3*(2), 142–149. https://doi.org/10.1037/0278-7393.3.2.142.

Arnold, D. J., & Brooks, P. H. (1976). Influence of contextual organizing material on childrens' listening comprehension. *Journal of Educational Psychology, 68,* 711–716.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston. https://doi.org/10.1080/00220671.1988.10885871.

Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzoni, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 287–313). Cambridge, MA: MIT.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28*(5), 610–632. https://doi.org/10.1016/0749-596X(89)90016-8.

Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders, 25*(1), 65–83.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.

Bollen, K. A., & Pearl, J. (2012). *Handbook of causal analysis for social research*. New York: Springer.

Bouffard, T., & Narciss, S. (2011). Benefits and risks of positive biases in self-evaluation of academic competence: Introduction. *International Journal of Educational Research., 50,* 205–208. https://doi.org/10.1016/j.ijer.2011.08.001.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 717–726. https://doi.org/10.1016/S0022-5371(72)80006-9.

Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 83*(3), 329–345. https://doi.org/10.1037/0022-0663.83.3.329.

Brown, J. D., & Marshall, M. A. (2001). Self-esteem and emotion: Some thoughts about feelings. *Personality and Social Psychology Bulletin, 27*(5), 575–584. https://doi.org/10.1177/0146167201275006.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–282. https://doi.org/10.1016/S0742-051X(97)80002-9.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253.

Clyne, M. (1987). Cultural differences in the organization of academic texts: English and German. *Journal of Pragmatics, 11*(2), 211–241. https://doi.org/10.1016/0378-2166(87)90196-2.

Coiro, J. (2003). Exploring literacy on the internet: Reading comprehension on the internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher, 56*(5), 458–464.

Corkill, A. J., Bruning, R. H., & Glover, J. A. (1988). Advance organizers: Concrete versus abstract. *The Journal of Educational Research, 82*(2), 76–81. https://doi.org/10.1080/00220671.1988.10885871.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology, 32*(3), 332–347.

Dooling, D. J, & Mullet, R. L. (1973). Locus of thematic effects in retention of prose. *Journal of Experimental Psychology, 97*(3), 404–406.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension a brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69–106.

Durrell, M. (2006). Germanic Languages. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 53–55). Amsterdam: Elsevier (North Holland Publishing Co.). https://doi.org/10.1016/B0-08-044854-2/02189-1.

Eccles, J. S., & Wigfield, A. (1992). The development of achievement-task values: A theoretical analysis. *Developmental Review, 12,* 265–310.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Social, emotional, and personality development in Handbook of Child Psychology* (Vol. 3, pp. 1017–1096). New York: Wiley.

Filippatou, D. (1995). *Effects of pictures and titles on reading accuracy and reading comprehension of primary school children, including children with specific developmental dyslexia (SpDD)*. Unpublished doctoral dissertation, University of Manchester.

Filippatou, D., & Pumfrey, P. D. (1995). Pictures, titles, reading accuracy and reading comprehension: A research review (1973–95). *Educational Research, 38*(3), 259–291.

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology, 41,* 98–110. https://doi.org/10.1016/j.cedpsych.2014.12.005.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences, 7*(3), 385–400. https://doi.org/10.1016/0191-8869(86)90014-0.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*(2), 256–281. https://doi.org/10.1037/0096-3445.113.2.256.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2*(1), 21–31. https://doi.org/10.1037/0278-7393.2.1.21.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition, 10*(6), 597–602. https://doi.org/10.3758/BF03202442.

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3–26). Mahwah, NJ: Lawrence Erlbaum Associates.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*(2), 193–202. https://doi.org/10.3758/BF03195564.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 334–372. https://doi.org/10.3102/003465430303953.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in educational theory and practice*. New York: Routledge.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology, 17,* 300–312.

Harter, S. (1985). Competence as a dimension of self-evaluation: Toward a comprehensive model of self-worth. In R. L. Leahy (Ed.), *The development of the self* (pp. 55–121). Orlando, FL: Academic Press.

Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository texts. *Reading Research Quarterly, 23*(4), 465–483. https://doi.org/10.2307/747644.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151–179.

Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Yun Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 89–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Hoffman, P. (2016). The meaning of 'life' and other abstract words: Insights from neuropsychology. *Journal of Neuropsychology, 10*(2), 317–343. https://doi.org/10.1111/jnp.12065.

Kieras, D. E. (1980). Initial mention as a signal to thematic content in technical passages. *Memory & Cognition, 8*(4), 345–353. https://doi.org/10.3758/BF03198274.

Kieras, D. E. (1981). Topicalization effects in cued recall of technical prose. *Memory & Cognition, 9*(6), 541–549. https://doi.org/10.3758/BF03202348.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270–1328). Newark, DE: International Reading Association.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.

König, E., & van der Auwera, J. (Eds.). (1994). *The Germanic languages*. New York: Routledge.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*(2), 149–171. https://doi.org/10.1006/ccog.2000.0433.

Kozminsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition, 5*(4), 482–490. https://doi.org/10.3758/BF03197390.

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education, 14*(1), 23–40.

Leopold, C., & Mayer, R. E. (2015). An imagination effect in learning from scientific text. *Journal of Educational Psychology, 107*(1), 47–63. https://doi.org/10.1037/a0037142.

Levin, J. R. (1989). A transfer-appropriate processing perspective of pictures in prose. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Educational applications*. New York: Springer.

Lin, L.-M., Moore, D. W., & Zabrucky, K. M. (2000). Metacomprehension knowledge and comprehension of expository and narrative texts among younger and older adults. *Educational Gerontology, 26*(8), 737–749.

Lin, L. M., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*(4), 345–391. https://doi.org/10.1006/ceps.1998.0972.

Long, S. A., Winograd, P. N., & Bridge, C. A. (1989). The effects of reader and text characteristics on imagery reported during and after reading. *Reading Research Quarterly, 24*(3), 353–372. https://doi.org/10.2307/747774.

Lorch, R. F., Jr. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*(3), 209–234. https://doi.org/10.1007/BF01320135.

Lorch, R. F., Jr., & Lorch, E. P. (1996). Effects of headings on text recall and summarization. *Contemporary Educational Psychology, 21*(3), 261–278. https://doi.org/10.1006/ceps.1996.0022.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280–296. https://doi.org/10.1037/0021-9010.67.3.280.

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 10*(4), 663–679. https://doi.org/10.1037/0278-7393.10.4.663.

Marchant, H. G., III, Royer, J. M., & Greene, B. A. (1988). Superior reliability and validity for a new form of the Sentence Verification Technique for measuring comprehension. *Educational and Psychological Measurement, 48*(3), 827–834. https://doi.org/10.1177/0013164488483032.

Masson, M. E., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 35*(2), 509. https://doi.org/10.1037/a0014876.

Mazzoni, G., & Cornoldi, C. (1993). Strategies in study time allocation: Why is study time sometimes not effective? *Journal of Experimental Psychology: General, 122*(1), 47–60. https://doi.org/10.1037/0096-3445.122.1.47.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1–43. https://doi.org/10.1207/s1532690xci1401_1.

Mengelkamp, C., & Bannert, M. (2010). Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for learning outcome. *Memory & Cognition, 38*(4), 441–451. https://doi.org/10.3758/MC.38.4.441.

Meyer, B. J. (1975). *The organization of prose and its effects on memory* (Vol. 1). Amsterdam: Elsevier (North-Holland Publishing Co.).

Narciss, S., Koerndle, H., & Dresel, M. (2011). Self-evaluation accuracy and satisfaction with performance: Are there affective costs or benefits of positive self-evaluation bias? *International Journal of Educational Research, 50*(4), 230–240. https://doi.org/10.1016/j.ijer.2011.08.004.

Narciss, S., Koerndle, H., & Proske, A. (2013). Challenges of investigating metacognitive tool use and effects in (rich) web-based learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 243–260). New York: Springer. https://doi.org/10.1007/978-1-4419-5546-3_17.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*(2), 102–116. https://doi.org/10.1037/0003-066X.51.2.102.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26,* 125–173. https://doi.org/10.1016/S0079-7421(08)60053-5.

Nietfeld, J. L., Enders, C. K., & Schraw, G. (2006). A Monte Carlo comparison of measures of relative and absolute monitoring accuracy. *Educational and Psychological Measurement, 66*(2), 258–271. https://doi.org/10.1177/0013164404273945.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003.

Ozuru, Y., Kurby, C. A., & McNamara, D. S. (2012). The effect of metacomprehension judgment task on comprehension monitoring and metacognitive accuracy. *Metacognition and Learning, 7*(2), 113–131. https://doi.org/10.1007/s11409-012-9087-y.

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

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1), 1–25. https://doi.org/10.1037/h0025327.

Patall, E. A. (2013). Constructing motivation through choice, interest, and interestingness. *Journal of Educational Psychology, 105*(2), 522. https://doi.org/10.1037/a0030307.

Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*(1), 69–80. https://doi.org/10.1037/0278-7393.28.1.69.

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84*(5), 754–775.

Ritchey, K., Schuster, J., & Allen, J. (2008). How the relationship between text and headings influences readers' memory. *Contemporary Educational Psychology, 33*(4), 859–874. https://doi.org/10.1016/j.cedpsych.2007.11.001.

Royer, J. M., & Cable, G. W. (1976). Illustrations, analogies, and facilitative transfer in prose learning. *Journal of Educational Psychology, 68*(2), 205. https://doi.org/10.1037/0022-0663.68.2.205.

Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educational Psychology Review, 13*(3), 263–281. https://doi.org/10.1023/A:1016675822931.

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology, 85*(2), 291. https://doi.org/10.1037/0022-0663.85.2.291.

Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology, 92*(1), 85. https://doi.org/10.1037/0022-0663.92.1.85.

Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1329–1362). Newark, NJ: International Reading Association.

Sadoski, M., & Paivio, A. (2013). *Imagery and text—A dual-coding theory of reading and writing* (2nd ed.). New York: Routledge.

Schallert, D. L. (1976). Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior, 15*(6), 621–632.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. https://doi.org/10.1016/S0959-4752(02)00017-8.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*(1), 33–45.

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Literacy Research, 27*(1), 1–17. https://doi.org/10.1080/10862969509547866.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review, 13*(1), 23–52. https://doi.org/10.1023/A:1009004801455.

Schüler, A., Arndt, J., & Scheiter, K. (2015). Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations? *Learning and Instruction, 35,* 62–72. https://doi.org/10.1016/j.learninstruc.2014.09.005.

Schwartz, M. N. K., & Flammer, A. (1981). Text structure and title-effects on comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 20,* 61–66.

Shepherd, S. (1990). *Some effects of pictures and titles on 7–8 year olds' recall and comprehension of aurally presented stories*. Unpublished M.Ed. dissertation, University of Manchester.

Shimada, H. (2016). Effects of components of educational materials on motivation for reading. *Japanese Journal of Educational Psychology, 64*(3), 296–306. https://doi.org/10.5926/jjep.64.296.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science, 17*(1), 57–60. https://doi.org/10.1111/j.1467-8721.2008.00548.x.

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*(4), 437–475. https://doi.org/10.1023/A:1009084430926.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. Boston: Allyn & Bacon.

Thiede, K. W., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66.

Thiede, K. W., Wiley, J., & Griffin, T. D. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology, 81*(2), 264–273. https://doi.org/10.1348/135910710X510494.

Wade, S. E., Schraw, G., Buxton, W. M., & Hayes, M. T. (1993). Seduction of the strategic reader: Effects of interest on strategies and recall. *Reading Research Quarterly, 28*(2), 93–114. https://doi.org/10.2307/747885.

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkavera, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping, 31*(10), 1459–1468.

Weaver, C. A. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(2), 214–222. https://doi.org/10.1037/0278-7393.16.2.214.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology, 132*(4), 408–428. https://doi.org/10.3200/GENP.132.4.408-428.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 27–30). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Yuill, N., & Joscelyn, T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology, 80*(2), 152–158.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). New York: Routledge.