17 Learning Tech
Journal of learning resources, didactics and technology.i

LÆRE
MIDDEL
DK

Analysis model for the subject-specific use of text-image constructions in instructional texts

Analysis model that captures the subject-specific potential in writing-image constructs

By Stig Toke Gissel

Correct citation of this article according to the APA system
(American Psychological Association System, 7th Edition):
Gissel, S. T. (2026). Analysis model for the subject-specific use of text-image constructions in instructional texts. *Learning Tech*, *17*, 47-78. DOI: 10.7146/lt.v10i17.158795

Abstract

When a multimodal construction that combines image and text appears in an instructional material, it makes a difference when analyzing the image as a motivated sign. In instructional materials, such a sign will basically have a learning and teaching purpose. Therefore, it is essential to adopt a subject-disciplinary and subject-didactic focus in the analysis of the use of multimodal elements in instructional texts. This article presents a new method for analyzing how text-image compositions are to a greater or lesser extent and for different instructional purposes used to support learning and acquisition in instructional texts. Methodologically, theories and concepts from multimodal social semiotics, theory of disciplinary specificity in subject-related texts and learning materials, as well as concepts from instructional material analysis are used. The article's contribution lies in its methodological development and has relevance for student teachers, teachers, and producers of teaching aids.

Når en multimodal konstruktion, der kombinerer billede og verbaltekst, optræder i en undervisningstekst, gør det en forskel, når billedet analyseres som motiveret tegn. I læremidlet vil et sådant tegn grundlæggende have et lærings- og undervisningsformål. Derfor er det væsentligt at anlægge et fagligt og fagdidaktisk fokus i analysen af brugen af multimodale elementer i undervisningstekster. Denne artikel præsenterer en ny metode til at analysere, hvordan tekst-billede-kompositioner - i større eller mindre grad og med forskellige didaktiske formål - bruges til at understøtte læring og tilegnelse i undervisningstekster. Metodisk anvendes teorier og begreber fra den socialsemiotiske multimodale teori, teori om fagspecificitet i fagrelaterede tekster og læremidler samt begreber fra læremiddelanalyse. Artiklens bidrag er metodologisk udvikling og har relevans for lærerstuderende, lærere og producenter af undervisningsmidler.

Analysis model for the subject-specific use of text-image constructions in instructional texts

Analysis model that captures the subject -specific potential in writing-image con-structs

Stig Toke Gissel, UCL Erhvervsakademi og Professionshøjskole

1. Indledning

PISA 2022 has confirmed the discouraging fact that more and more students are leaving elementary school without functional reading competencies (OECD, 2023). Thus, as many as 19% of Danish 15-year-olds perform below the minimum level that enables them to use reading to achieve their own goals and participate actively in social life. For previous PISA rounds, this figure was 15-16%. Furthermore, Denmark stands out negatively by having a significantly smaller proportion of high-achieving readers than comparable countries (Gissel, 2023).

The concept of 'reading literacy' not only describes the ability of a student to quickly and correctly decode a written text. Recognizing that the text landscape is complex, multimodal and very often also digital, PISA's definition of text reading now – unlike before – includes texts in a broader sense. They can be both digital, dynamic and non-sequential, as is the case for web, mobile and other digital texts (OECD, 2019). In other words, the perception of reading as an activity that deals exclusively with written words is outdated (OECD, 2019).

1.1 Multimodality in instructional texts

Social semiotics focuses on the multimodal nature of texts and the

necessity that, in the analysis of texts, we not only look at the written word but also include the contribution of other semiotic systems to the text's overall meaning potential. The theory of multimodality and social semiotics has shown that the different modalities constitute semiotic resources, each of which has its own advantages and challenges in terms of creating meaning. The digital and analogue instructional materials used in primary schools are, to a large extent, multimodal. It is obvious that instructional texts, like other contemporary texts, place more emphasis on the use of visual resources for communication than previously (Kress & van Leeuwen, 2006).

This applies not least to instructional texts. *Instructional texts* are characterized by being designed for specific teaching and learning purposes (Gissel, 2026a) and contain a degree of didactic intention. This is typically seen in the form of explicated learning objectives, student tasks, texts with an appropriate level of difficulty and support for the student's learning process, and evaluation methods. A typical example of an instructional text is a textbook for a specific subject and grade level or a course within a specific subject area in school. In recent decades, the emphasis on images in relation to written text has increased in instructional texts, so that images play a more important and prominent role in teaching materials (Brugar & Kress, 2008; Lee, 2010; Janko & Knecht, 2014).

1.2 Multimodality as a resource to support learning in instructional materials

Cognitively oriented research on the use of multimedia in relation to learning and its effects (e.g., Mayer, 2014b; Butcher, 2006; Mayer & Gallini, 1990) shows that significant learning benefits can be achieved through the integration of text and graphics in instructional texts. However, the positive effects of the combination of graphics and words depend on whether the texts are designed appropriately (Mayer, 2014b; Molina et al., 2018). Some research has been concerned with assessing the appropriateness of various visual representations in instructional materials and how these elements are linked to the written parts of the text (e.g. King-Sears et al., 2018; Cheng et al., 2015; Slough et al., 2010).

Furthermore, the positive effects of multimodal presentations depend on whether students are able to use these multimodal resources appropriately. Several studies have shown that students pay much more attention to the written text in instructional materials than to the visual elements (Liu & Chuang, 2011; Pozzer & Roth, 2003; Shin, 2017). Other research has shown that students may have problems linking complex graphic-verbal text combinations and processing in-

formation from graphic elements in instructional materials (Mason et al. 2015; Schnotz et al., 2014). Hung (2014) used eye-tracking to investigate how students utilize information from written and visual representations when attempting to understand a science instructional text. Hung found that students paid less attention to illustrations than to written parts of the text, but that students who utilized the visual parts of the text to a greater extent had better comprehension.

This calls for students to be familiarized with reading strategies and reading skills to handle and benefit from multimodal instructional materials (Danielsson & Selander, 2016). Thus, Lim and Toh (2020) point to the necessity of teaching both generic reading strategies and the importance of different multimodal sign systems. However, research indicates that teaching how to read visual elements in instructional materials is often neglected by teachers (Coleman, McTigue, & Smolkin, 2011).

Other research has shown that it can yield positive benefits to carry out teaching in reading the visual. Alvermann and Wilson (2011) investigated how teaching reading strategies specifically aimed at handling semiotic resources in science texts can contribute to students participating more actively in meaning-making processes. Brugar and Roberts (2017) show how training efforts for teachers aimed at the targeted use of graphic elements in teaching materials for social studies can have positive effects in relation to students' understanding of these elements. Ohle-Peters et al. (2023) also emphasize the importance of the teacher as a mediator, finding that teachers' professional competence in teaching students about combinations of graphics and written text and the quality of teachers' teaching were related to how skilled students became at integrating meaning from combinations of words and images.

Danielsson and Selander (2021) argue that there is a challenge in that both teachers and students seem to assume that images in teaching materials support the content and that images themselves do not pose a challenge to students' meaning-making processes (p. 27).

On the one hand, the use of visual semiotic resources can support students' understanding of content in instructional texts, but on the other hand, the increased use of visual resources also potentially increases the complexity of the instructional materials and can challenge students' acquisition.

The instructional materials that students use in school help shape students' expectations, knowledge, and reading behavior in relation to utilizing graphic elements in the educational resources for learning. Therefore, it is important to investigate which multimodal constructions students encounter in instructional materials, what didactic

purposes the graphic elements have, and how they are presented in the instructional material in relation to the student's reading and learning process.

As mentioned, research shows that teachers may neglect to teach visual literacy. However, it is an open but important question whether instructional materials themselves neglect to strengthen students' visual literacy, that is, whether instructional materials provide explanations of how a disciplinary graphic element is constructed and should be read, or if the material scaffolds learning processes that aim to strengthen students' visual literacy. Fterniati (2009) analyses how multimodal texts are presented in Greek instructional materials, what types of activities support multimodal text comprehension, analysis, and production, and whether the teaching materials support teaching in different forms of communication. Fterniati (2009) found a significant presence of multimodal elements in the teaching materials studied. However, the teaching materials did not provide for an exploration of their multimodal aspects, which suggests that more could be done to use these texts to develop students' multimodal reading skills. Whether the instructional material supports the strengthening of students' visual literacy will be analyzed using the model presented in this article.

1.3 The image as a didactically meaningful sign

Karlsson (2012) argues that the analysis of multimodal texts should be based on an analysis of the context that has a broader scope than the purely text-based context analysis that is typical of social semiotic work. When we analyze instructional materials, text-image constructions are motivated signs that, firstly, can be assumed to aim to support the students' acquisition processes, and secondly, are embedded in a professional discourse, tradition and curriculum, i.e. an understanding of what the goals, methods, and content of the subject.

Kress & van Leeuwen (2006) write that reading images is a sign-creating process. Students must thus actively create a meaningful sign from, for example, an instructional material author embedding a picture in an instructional-material text. Kress and van Leeuwen are primarily concerned with how the relationship between people in the picture and the recipient is established (the interactive meaning internal to the picture). In this study, however, the focus is instead on the academic motives for using pictures in instructional materials. It is argued that this requires attention to the discipline-specific and subject-didactic context in which the picture - with all its means of creating meaning potential - is embedded.

The image in an instructional text is a motivated sign (Kress & van

Leeuwen, 2006) - on potentially several levels. An example can illustrate the possible complexity. If the author of a textbook for teaching L1 includes a painting in the instructional material, the image takes on a new, special function in the context of the educational text. The artist who made the painting obviously has his or her motives for producing the sign. In the context of the instructional material, it may be the student's task to interpret what the sender's motive for the artwork. Likewise, the instructional material can draw the student's attention to (parts of) the image's meaning-making elements (the textual metafunction) and the painting's motive. But the image can serve a number of additional functions in the context, e.g. an ideational function as an illustration of the verbal text's statements, or a more interpersonal function by simply being decorative.

I propose the concept of subject-didactic salience to describe what it is in the image that the subject-didactic context makes central. The subject-didactic context can be understood here as the (presumed) motivation of the instructional- material author for embedding the image in the instructional material.

In the teacher's use of the instructional material in the teaching situation, we can add an extra layer of motivation, as the teacher has his or her own motives for using the instructional-material text. In his or her adaptation, the teacher can choose to take other paths than those intended by the instructional-material author (Gissel et al., 2025). Since our study does not examine the use of instructional materials but only their potential didactic potential (Bundsgaard & Hansen, 2011), this layer of meaning is absent.

1.4 Multimodality in a subject-disciplinary and subject-didactic light

Different subjects use language in different ways (e.g. Halliday & Martin, 1993; Schleppegrell, 2004). Shanahan and Shanahan (2008, 2012) have demonstrated that the subject-specific goals in the subjects are realized through specific disciplinary approaches to academic reading, called "disciplinary literacy". Each subject has its own approaches to cognition and understanding, and subjects use language in different ways (e.g. Halliday & Martin, 1993; Schleppegrell, 2004). The texts of the disciplines have distinctive features that reflect each subject's approach to cognition and understanding and thus require subject-specific reading skills. In history, for example, being a critical examiner of historical sources is a special subject-specific reading style, while in the L1 subject, it is important to acquire a fictional reading form when encountering literary texts.

Against this background, it is noteworthy that what we can call

discipline-specific visual literacy has received relatively little attention from researchers (Guo, Wright & McTigue, 2018). Theories of disciplinary literacy do not originally have a multimodal focus, but the hypothesis guiding this article is that, within individual subjects, there are special multimodally constituted ways of representing academic content - just as each subject promotes special ways of working with multimodal constructions. As mentioned, research has indicated that teachers often overlook or fail to teach students about the specific demands or challenges that graphic elements in the teaching materials of the different subjects can entail (Danielsson, 2010). This makes it crucial to point out the specific demands imposed by the instructional materials and academic texts of the different subjects and to understand the extent to which there are commonalities across subjects.

A recent review of research on the relationship between multimodality and disciplinary literacy in instructional materials (Gissel et al., 2026b) showed that, out of the 50 included articles, about half draw no conclusions related to the subjects involved in the analysis. In these studies, the school subjects function as rather arbitrary examples rather than as disciplines with distinctive ways of understanding and thus representing the world.

Studies that did draw conclusions from the disciplinary specificity of multimodal instructional materials typically addressed only a single school subject (Gissel et al., 2026b), and the research design thus did not allow for comparative approaches. Typically, the disciplinary focus is on the relationship between multimodality and a specific content area within the school subject or a very narrow range of representations - for example, artistic representations in textbooks (McKean, 2002) or graphs (Lee, 2010) in relation to a particular subject.

Lee (2010) provides a qualified proposal for a subject-specific and discipline-specific study of multimodality in instructional materials. Lee (2010) conducted a historical survey of the subject-specific use of various representations in science textbooks, in which educational purposes were one of the themes analyzed. Lee classified the purposes of the representations as 1) explaining scientific ideas, 2) illustrating their use in everyday life, or 3) demonstrating how the student can replicate a scientific phenomenon. Lee also analyzed whether the ideational or interpersonal metafunction was dominant in the use of scientific representations to determine whether the representations primarily served a professional or a social purpose. Lee found that the use of decorative illustrations with social purposes had been present in teaching aids for the past 60 years before the time of the study, but that the use of photographs has increased – probably to connect with the students' everyday experiences.

Guo, Wright and McTigue (2018) and Fingeret (2012) are exceptions to a mono-subject and narrow focus. They analyzed the diverse use of graphics in science and social studies teaching materials and highlight the subject differences. The categories used to characterize the application of graphics in the two subjects describe the academic goals achieved through different uses of graphics. Guo, Wright and McTigue (2018) found that science teaching materials contained more diagrams and photographs functioning representatively, i.e., the graphics illustrate parts of the literal meaning of the written text or specify the abstract. In social studies teaching materials, on the other hand, a wider range of graphics was used, which were often more challenging and functioned interpretively in relation to the written text.

But Guo, Wright and McTigue (2018) and Fingeret (2012) also conducted a deductive analyses, understood in such a way that the instructional materials were analyzed based on broad and interdisciplinary multimodal basic types, which could actually be used in all subjects. The analyses can thus show how these categories find different degrees of application in teaching materials for different subjects. However, the purpose of using a basic type is transverse, not subject-specific.

There is thus a need for new analytical approaches that enable a description of the discipline-specific use of multimodality. I propose to do this by proceeding inductively. In this article, I present a new analytical approach to writing-image constellations in instructional materials, which analyze each verbal-lingustic-image construction based on a number of given parameters and then, for each occurrence, conceptualizes the specific subject-didactic purpose that the multimodal construction in the instructional material can serve.

This approach makes it possible to inductively build an understanding of which multimodal constructions characterize the teaching materials of the individual subjects, to formulate disciplinary super categories for patterns in these occurrences, and to compare across subjects and provide a proposal for a transdisciplinary typology that future teaching material analyses can nuance, expand, and enrich.

This leads me to the following research question, which I will attempt to elucidate in the article:

How can text-image constructions in instructional texts be analyzed and characterized in a way that allows us to illuminate the academic motives for using the text-image constructions?

2. Method: Analysis model for the educational use of text-image constructions in instructional materials

Since this article is methodologically developmental, it will consist of a review and discussion of the justifications for the choices made in relation to analyzing instructional materials from a multimodal perspective, as well as examples drawn from teaching materials for teaching in various subjects in the school's subject range.

Figure 1.
Model of layers and elements in the analysis of subject-specific use of text-image constructions in instructional materials.

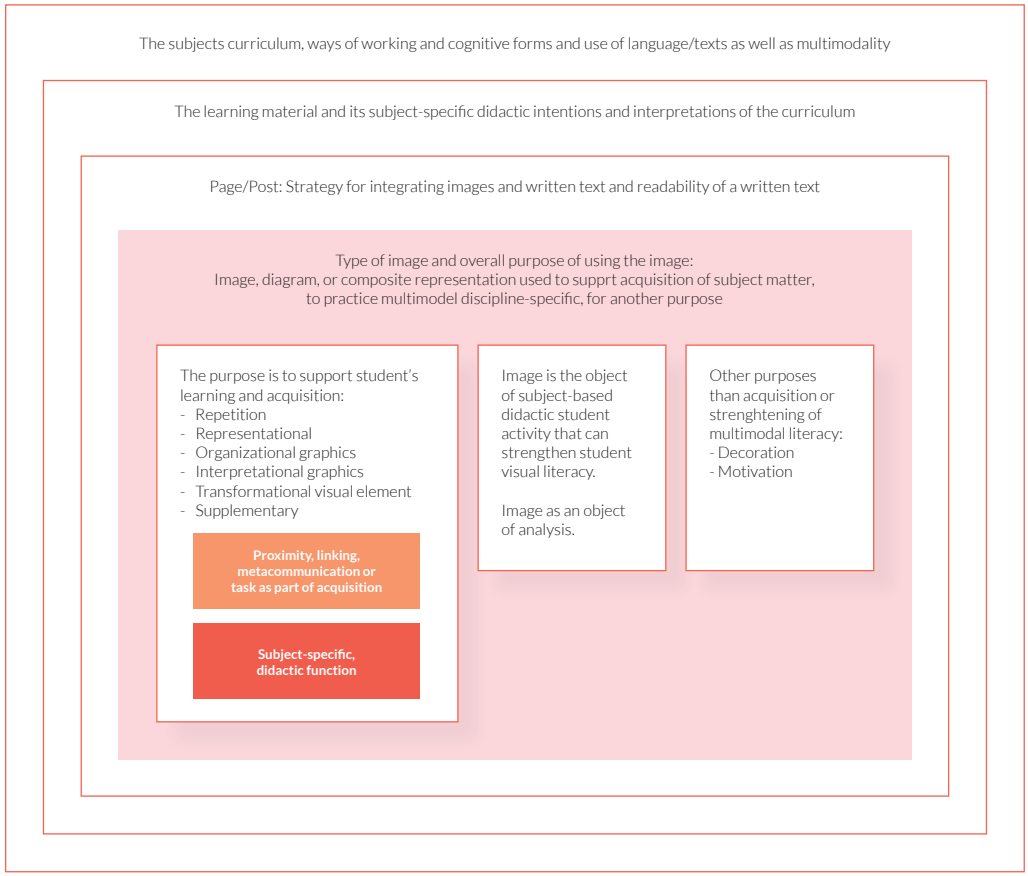


Figure 1 shows the components of the analysis model, which will be presented step by step in the following. The figure shows that the analysis begins by explaining the part of the context for the instructional material that consists of the subject's curriculum, typical forms of work and cognition, and the subject's disciplinary-specific use of texts, language, and multimodality. The next step is a description of the context that the instructional material itself, as a whole, constitutes for the choices made - right down to the level of individual images in the instructional material.

In the definition of instructional text above, it was clear that the instructional texts are characterized by didactic intentionality. The didactic intentions ideally condition all choices made in the design of the instructional text and therefore constitute an important context for understanding the function of each individual image. The next layer in the analysis consist of the individual page in the instructional text -the spread in an analog instructional text, or the web page in the digital one. Here, I am interested in how images and written text are generally integrated into the layout of the page and the accessibility of the academic content in the written text.

The next step (represented by another layer) is to go down to the level of the individual image and determine its type and overall function in relation to supporting the student's acquisition of academic content (if any). Furthermore, the analysis should consider what measures have been taken to strengthen the student's learning gains from exposure to the image, whether the image is the object of an academic analysis, or whether the image fulfills another function. Each of these layers will be explained in the following sections. In the gray boxes, the considerations are distilled into analytical questions.

But first, some meta-considerations are presented regarding which levels are relevant in analyzing the subject-didactic function of images in instructional texts.

2.1 Foci in the analysis of complex, multimodal instructional texts

When analyzing instructional texts, you deal with relatively complex texts that are often composed of several visual and verbal texts that relate to each other in different ways and work together to create meaning. This means that different semiotic resource systems interact. In the social semiotic analysis of such complex texts, however, it can be problematic to rank the many details and potential meaning-bearing units in relation to each other (Boeriis, 2012). What is most important, and what is less central?

In this model, I focus on how visual elements, together with, for example, captions or other visual elements, form local text units

within the framework of a more general whole - namely, the overall page in the instructional material, which in turn forms part of the overall instructional text. Our goal is to understand how different subjects use certain combinations of meaning-creating resources in the service of subject didactics.

Therefore, I do not analyze the subcomponents of these local text units (e.g., the elements in a photograph or a video) but instead attempt to describe how the local text units create meaning together with the overall written text and other elements on the pages, thus contributing to the creation of subject-didactic meaning. In the present model, I focus on text-image relations and thus delimit, for example, layout and use of typography.

2.1.1 *The disciplinary context - subject characteristics*

The categorization of the subject-didactic function of the graphic-verbal text combinations is based on each subject's curriculum, typical forms of work and cognition, and the disciplinary use of and interest in multimodality. Each subject has its own definitions of what an academic inclusion of graphic elements and has special semiotic approaches.

> Create an overall description of what the subject's curriculum says about the use of verbal language and multimodality, why this is interesting from the perspective of the subject, and how these are typically worked with.

2.1.2 *The instructional text and its subject-specific didactic intentions and interpretations of the curriculum*

An instructional text constitutes a translation or interpretation of the curriculum and a proposal for how the teaching of the subject can be carried out. In the instructional text, subject-didactic intentions are ideally formulated in the instructional material, including a view on how to teach the subject in question and a theory of learning in the teacher's guide (Gissel et al., 2023). This is an important context for understanding the choices made in the student part of the instructional material, including the use of multimodality.

> What are the academic intentions of the instructional material, including the subject-didactic view and the theory of learning?

2.1.3 Overall strategy for integrating images and written text on the page and the difficulty level of the written text

At this point in the analysis, I focus on the context within the instructional text for the use of visual material.

An important context for understanding how images and text work together on a given page in an instructional text is the overall strategy that is generally adopted on the page to integrate the two forms of representation. Peterson (2016) describes and tests three ways of integrating written text and images, which I will use as the basis for this part of the analysis. The three forms of integration range from images being marginal in relation to the written text, through an intermediate form, to a model with full integration between written text and images, where the reader largely determines the reading path:

— *The written text is the primary mode of representation*, and images are marginal in relation to the text. Readers must either figure out for themselves when to interrupt the reading of the written text in order to read the image – unless there is an explicit link to the graphic element in the body of the text (e.g., "see figure 1").

— *Images are the primary form of representation*, with the written text supporting the images. In Peterson's experiment, there were a series of boxes where the image was placed at the top, and the short written text was placed under each image as a caption. Here, the reader gets the feeling of being able to determine the reading path and manner to a certain extent.

— *Full integration between images and written text*. This can be seen, for example, in a diagram where graphics and words are integrated, or a picture of a lake where small pieces of text are placed on the image. Entire pages can integrate written text and image in this way. Here, there is no clear reading path; the reader largely decides for themselves where to start and end.

Peterson's results indicated that full integration provided benefits in terms of both student understanding and situational interest.

Overall strategy for integrating images and written text on the page:
On the overall page/post, is the written text the primary form of representation, are the images the primary form of representation, or is there full (or a high degree of) integration between the two?

Another important contextual factor in assessing the appropriateness of graphic elements in the instructional material is the difficulty of the written text. The more difficult, the more abstract, and the less readable the verbal text is, the greater the need may be to support the student's understanding through visual aids.

When assessing the difficulty of the verbal text for the chapter or course being analyzed, it is useful to focus on:

- The degree of technicality of the texts, i.e., the use of terms that are far from everyday language.
- Whether a connection is made between the technical language of the text and the student's everyday language (e.g., through explanations and examples).
- Whether the verbal text is characterized by nominalizations, passive constructions, and academic language use - all of which are factors that can reduce the readability of the texts (Gissel, 2026b).

The assessment of the readability of the written text is important in order to assess the need to utilize multimodality to support the student's acquisition, as well as whether the potential for allowing multimodality to support the student's academic acquisition is utilized in the instructional material.

2.1.4 *The type of representation form*

It is obvious that the potential for meaning in a visual element in an instructional text depends on the type of representation being dealt with. With the term image, I include both what is traditionally called graphics and pictures. According to Bertin (1983), graphics (e.g., diagrams, maps, graphs and network diagrams) are defined as belonging to monosemic sign systems, i.e., signs that have a clear and unique meaning and whose design depend on predefined conventions. Conversely, pictorial representations (e.g., paintings, photographs and drawings) are polysemic, because they can be interpreted subjectively and ambiguously.

You can distinguish between the following visual representations:

— Images: For example, photos, films, paintings, or drawings.

— Graphics: For example, graphs, maps, diagrams, or models.

— Composite texts: Where several images or diagrams are combined, possibly with written text, within the framework of the same unit, for example, a timeline with images or a video with speech and images.

This is somewhat of a simplification compared to previous studies, which operate with up to eight types with 59 associated subtypes (Fingeret, 2012; see also Guo, Wright & McTigue, 2018). Roberts et al. (2013) operate with eight types: graphics with captions, diagrams, flowcharts, graphs, 'insets' (e.g., picture in picture), maps, tables, and timelines. In my view, the specific type of graphic representation is not so important to quantify. Rather, it is useful to identify instances of and form a typology for subject-didactic functions of visual material and to be aware that these intentions can be realized in different ways with different types of graphics. These typologies are yet to be developed, but a tentative exploration for six subjects can be found in Gissel et al. (2026a).

As the analysis model first identifies the type of representation form, it makes it possible to identify which types of representations dominate a given learning resource, as well as whether there are certain types of representations are particularly linked to the subject, i.e., cross-curricular and between different learning resources, or whether there are rather local patterns within a given learning resource. This analysis of representation type is also based on a number of previous studies of the use of multimodality in instructional texts (e.g. Alkhateeb, 2019; Devetak, Vogrinc, & Glažar, 2010; Guo, Wright & McTigue, 2018; Lee, 2010; Oruç, Uşurlu & Tokcan, 2010; Šimik, 2021).

The model does not include icons that metacommunicate about, for example, organization (could be an icon with two heads symbolizing pair work) or markers for textual relationships (e.g., a red box with a white arrow symbolizing that something is a link). These icons and textual markers are used across subjects - for example, when in the digital subject portals - and can be seen as general didactic markers that are not relevant in isolation when understanding the discipline-specific use of visual elements.

2.2 The functions of the visual element

In the next steps of the analysis, the function of the visual element will be analyzed. I propose three categories of functions (see Figure 1):

— Use of visual elements in instructional materials that have the purpose of *supporting the student's learning and acquisition*, i.e. that the visual element serves to support or expand the student's understanding of the written text. This use of image-written text construction has characteristics that cross subjects, but it is necessary to move a step further - from a general characteristic of the purpose of using a visual element - to formulate a specific subject didactic function. The impact of such images can be enhanced through compliance with the principle of proximity, explicit linking, metacommunication about multimodal literacy, or by formulating a task as part of the student's acquisition.

— The image is the *object of subject-based didactic student activity*, which can directly or indirectly strengthen multimodal literacy.

— A *decorative use of visual elements*, which serves purposes other than academic learning and the acquisition or strengthening of multimodal literacy.

2.3 Interdisciplinary typologies for the use of visual elements

As mentioned in the article's state of the art, it is atypical in the research literature to analyze multimodality from a perspective that broadly asks what visual elements are found in the learning resources for the subject, and how these are connected to the subject's curriculum (Gissel et al., 2026b). The vast majority of studies base their analyses of the learning resources on general analytical categories. In this context, various researchers have contributed in particular and in different ways to describing how a general, interdisciplinary use of visual elements in instructional texts should be designed and characterized.

The first approach I will highlight is Mayer and colleagues (e.g. Mayer, 2014a; Mayer, 2014b; Mayer & Gallini, 1990), who have formulated a number of principles for learning through so-called multimedia, i.e., texts that contain both words (spoken and printed) and graphics (e.g. images, video, animation or diagrams). This tradition is based on cognitive load theory. The principles constitute concrete guidelines for the appropriate design of multimedia and, as clear hypotheses, have been subjected to empirical testing by various researchers. The principles thus relate to learning resources, but neither the principles

themselves nor the studies that empirically test them adopt a disciplinary perspective (Gissel et al., 2026b). Mayer & Co's principles will be included on an ongoing basis where relevant.

The second tradition on which the present approach stands, consists of various attempts to develop typologies for the possible relationships between images and written text, as well as the functions of these relationships. This is described specifically for texts that are read for learning, including instructional texts, by Levin (1981), and more generally for all text types by van Leeuwen (2005). In the following, I will briefly review and compare Levin´s and van Leeuwen's types and provide a rough assessment of how much variation I would expect to find between different subjects' use of each function. First, combinations of writing and images that aim to support students' learning are discussed. Then, I review the use of images where students must analyze the visual element, and lastly, a decorative, non-academic use of visual elements.

2.4 When image and written text are combined to support student learning

Levin et al. (1987) identified eight functions that visual elements in instructional texts can have. Levin ranked them according to how beneficial an effect they could be attributed in relation to students' learning outcomes. The first three of the eight are thus not specifically academic functions: When images function as pure decoration, are intended to boost the sale of the book, or have a primarily interpersonal function in relation to creating motivation in the student. The

Figure 2.
Relationships between image and word. Source: Theo van Leeuwen Introducing Social Semiotics (2005).

Image-text relations

Elaboration	Specification	The image makes the text more specific (illustration)
		The text makes the image more specific (anchorage)
	Explantion	The text paraphrases the image (or vice versa)
Extension	Similarity	The content of the text is similar to that of the image
	Contrast	The content of the text contrasts to that of the image
	Complement	The content of the image adds further information to that of the text, and vice versa ('relay')

function of decoration will be discussed separately below. The decorative function is important to include because studies show that there are a great many decorative images in contemporary teaching materials (Gissel et al., 2026a). It is therefore crucial to discuss whether this use of images typically has a beneficial effect or perhaps a downright harmful effect on the student's learning outcomes.

Van Leeuwen (2005) describes five relationships between image and words, two of which generally deepen meaning, and three that expand meaning (Figure 2).

Van Leeuwen's relationships partially overlap with Levin's, but van Leeuwen is concerned with what the two forms of representation do to each other's meaning potential when combined. Van Leeuwen thus does not take a didactic view of the use of images and verbal text, although I expect that the categories are somewhat applicable in relation to instructional texts. Therefore, van Leeuwen's typology will be compared with Levin's in the following, and an attempt will be made to translate it into a educational, instructional material context.

In relation to Levin's typology, I have now reached the functions with visual elements that have an actual learning purpose. Levin's fourth function is *repetition*, i.e., when the image simply repeats information from the verbal text, so that the image serves as a repetition of the verbal text´s content. This function thus implies, in van Leeuwen's terminology, a similarity between written text and image. Here, social semiotics would point out that the two essentially different modalities, written words and images, have different affordances and thus will have different potential for meaning - despite their similarities and the intention of repetition.

This use of images, that broadly show what is described in the verbal text, can be seen in basal readers, i.e., textbooks for beginning reading instruction, where the image can function as a support for the student's understanding of the written text - but where the image will also concretize empty spaces (Iser, 1971) in the verbal text. Thus, the image can function as support for a conversation about the text that has been read and, in relation to the often sparse verbal text, as an expansion of the available meaning potential - which van Leeuwen also states as the overall purpose of the similarity relation. This exemplified use of graphics that function as repetition is very special for beginner reading materials, where the student may need a scaffolding around decoding and text comprehension attempts. It is interesting to investigate the prevalence of this use of images in other disciplinary contexts.

Levin's *representational* function corresponds to van Leeuwen's specification. The representational function involves images that are

Figure 3.
Image from a digital instructional text for teaching history, more specifically WW2 (from the course "Anden Verdenskrig", Knudsen & Buttenschøn, n.d.).

relevant to the content of the verbal text and that present the content of the verbal text in a different way, namely visually. The image represents the content of the verbal text in a more specific, concrete way.

An example of the representational function is shown in Figure 3. The body of the history teaching material states that Poland was invaded and quickly surrendered; the Germans and Russians each took half. The photograph specifies the verbal text, as it shows a specific group of soldiers in a specific combat action as part of the invasion. With van Leeuwen's typology in mind, we can also state that written words can, however, also specify the meaning of the image within its context. The caption in Figure 3 thus specifies (or anchors) the meaning potential of the image by providing the image with a year for the event and by pointing out that these are German soldiers, not, for example, Russian ones. I expect this specified use of images and diagrams to play different specific roles in different subjects. In the history subject, images are most often regarded as sources whose indexical (Peirce, 1997) connection to the historical past is central.

A similar, overall representational, specifying, and concretizing function can be fulfilled by a photograph of a dandelion in relation to a written text about herbs in a science textbook. However, the speci-

fication here serves a different subject didactic purpose. In this case, the picture of the dandelion specifies a supercategory - herbs - with an example: the dandelion. It is not so important where or when the picture was taken, or who took it: the specific dandelion is simply a representative of the overall category, dandelions.

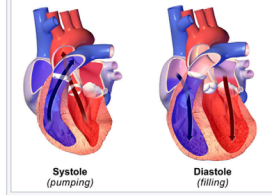
I expect that all subjects' instructional texts will include images that have a representational, specifying and concretizing function in relation to verbal text. Carney and Levin (2002) point to this function as the most widespread. However, I also expect to find subject-specific variation in how this type of illustration is typically used, as exemplified above for history and science. Thus, a discipline-specific approach will require that we try to be as precise as possible in our descriptions of what kind of text is specified through a particular use of images (or vice versa), and how we can characterize this subject-specific use of specification. In Gissel et al. (2026a), an attempt is made at an inductive, subject-specific analysis of writing-image relations, resulting in a tentative subject-didactic typology for the purposes of the images in six subjects.

Organizational graphics provide students with an overview or structure of the academic content in the written text (Carney & Levin, 2002). Illustrations can help the student to make the information in the written text coherent (Levin, 1981). These can be graphics that organize information in a clear way, such as a timeline or table (an example is shown in Figure 2 with van Leeuwen's typology), an illustration showing the phases of photosynthesis, or the steps in a procedure in craft subjects. What van Leeuwen (2005) calls explanatory and Carney and Levin (2002) call *interpretational* graphics aim to support the

Figure 4.
Article from Wikipedia about blood pressure.
Source: https://en.wikipedia.org/wiki/Blood_pressure

Physiology [edit]

During each heartbeat, blood pressure varies between a maximum (systolic) and a minimum (diastolic) pressure.[64][*unreliable medical source*] The blood pressure in the circulation is principally due to the pumping action of the heart.[65] However, blood pressure is also regulated by neural regulation from the brain (see Hypertension and the brain), as well as osmotic regulation from the kidney. Differences in mean blood pressure drive the flow of blood around the circulation. The rate of mean blood flow depends on both blood pressure and the resistance to flow presented by the blood vessels. In the absence of hydrostatic effects (e.g. standing), mean blood pressure decreases as the circulating blood moves away from the heart through arteries and capillaries due to viscous losses of energy. Mean blood pressure



Cardiac systole and diastole

students´ understanding of difficult text passages. An example is the illustration from the Wikipedia article on blood pressure (Figure 4), in which arrows on the two images of the heart show the heart as a pump that alternates between filling (arrows in) and pumping blood (arrows out). Interpretive graphics can have elements of both representational and organizing images, but will go beyond both, by presenting the content in a way that helps the reader's understanding (Guo, Wright & McTigue, 2018).

Van Leeuwen points out that the direction can go both ways: the written text can explain the image, and the image can explain the written text.

Transformational visuals support the learners´ memory by presenting something in a way that they can more easily relate to (Levin, 1981; Carney & Levin, 2002). Examples include a narrative template designeded as a face or a bridge. Transformational illustrations are intended to function as mnemonics.

Also worth mentioning is the type of interaction that van Leeuwen calls *supplementary*. This is not mentioned by Levin (Levin, 1981; Carney & Levin, 2002), but refers to cases where the image contains related information that is not addressed in the written text (or vice versa).

Hopefully, it has become clear that I consistently emphasize the need to analyze the academic, specific content represented through visual elements. Here I am in line with Peterson (2016), who precisely formulates the necessity of focusing on student prerequisites, and not least, on content when assessing the appropriateness of a visual element:

> *The suitability of a text–image integration scheme must depend upon the information represented. (...) But it seems likely both audience (from grade school to higher education learners) and subject area (from cell theory to Earth science) determines scheme relevance at any given time.* (Peterson, 2016, p. 1369)

Therefore, for each occurrence of a visual element, a suggestion must be given as to which subject-specific function the juxtaposition of image and written text serves. The analyst is thus forced to formulate their own suggestions for concepts, which - when several similar subject-specific uses have been observed - can be synthesized into comprehensive overarching concepts.

Does the juxtaposition of image and written text in an instructional text function as repetition, representation, organizing, interpretational, transformational or supplementary?
How can the subject-specific use of this function be described?

2.5 Potential strengthening of the communicative power of visuals

Images used to support students' learning and acquisition can potentially be made more potent by adhering to the principle of spatial coherence, by creating explicit links between images and written text, through metacommunication about how, for example, a diagram should be read, or by the author of the instructional text activating the student in relation to processing the image, for example through a task. These possible measures will be reviewed below.

According to Mayer's (2014b) principles regarding the use of visual elements in learning resources and their connection with printed words, the principle of spatial contiguity is central. The principle states that printed words and graphics that belong together should be placed close to each other; this supports the recipient in creating the intended connection between words and graphics. The principle of spatial proximity probably applies to all of the possible relationships between graphics and written text mentioned.

However, *explicit linking* between written words and images can make explicit what spatial proximity only indirectly supports. With explicit linking, the reader is guided toward the illustration. Graphics and written words can be linked so that images are provided with annotations clarifying which parts of the academic content in the body text the diagram or image illustrates. This integration between visual elements and verbal text is not always handled appropriately in instructional texts (Kapıcı & Açıkalın, 2015; Slough et al., 2010; Bliss, 1990).

Research has found that if images in instructional texts are not explicitly linked to the written text or are insufficiently integrated, then students' comprehension suffers (e.g., Jaeger & Wiley, 2014; Wright, McTigue, Eslami, & Reynolds, 2014). Mayer et al. (1995) found that a combination of spatial coherence between prose and images and relevant annotation (where relevant parts of the verbal text are repeated in the annotation) in relation to the images significantly enhanced recipients' meaningful learning.

Image and verbal text can also be linked through signals (e.g. an icon or a reference, e.g. see figure 2) in the body text, which shows the reader when a particular figure is relevant for the reader to process. General instructions to the student à la "Read the text and look at the

pictures" in the introduction to an instructional text should not be considered explicit linking. This must be linked to each individual image instance in order to serve its purpose.

Implicit linking occurs when there is a degree of correspondence between the content of the body text and, for example, the motif of an image, but where the readers must infer the connection themselves.

Metacommunication is where the verbal text *explains*, for example, a diagram, so that the student is explained how different forms of representation create meaning. We must also assume that this can strengthen both the student's understanding of the text-image construction - but also the student's general multimodal literacy. In the example above (Figure 4), where the image explains the verbal text, annotation by the image could explain the meaning of the colors in the drawing or the meaning of the arrows. In geography, metacommunication can be a caption that accompanies a map.

A final group of measures that can enhance students' benefit from visual elements in instructional materials is when the instructional text contains instructions on how students should process the illustration. This, too, is a form of metacommunication to the student, but here it is about *activating* the student rather than the student gaining knowledge about multimodal literacy through the caption. Postigo and López-Manjón (2015) investigated how graphic elements are academically integrated in instructional materials by analyzing whether the use of graphics in instructional materials is connected to activities that guide students' subject-specific didactic interaction with the graphic elements - or whether there is an academic framework or connection between the written text and the graphics in some other way. The authors found that there was generally no support for students' processing of graphics in the analyzed instructional materials.

Adler (1993) has examined instructions for students' processing in relation to images with a representational function, but there is no reason to believe that it cannot be usefully applied to images with other functions. Adler (1993) conducted an experiment in which college students were assigned one of four processing instructions: a condition with no instruction, one that simply required them to observe things in the image (e.g., 'How many plants are seen in the image?'), a question that elaborated on the meaning of the image in context (e.g., 'How does the image relate to the text? Write your answer below.'), and a question that asked for perspective, or an open elaboration of the image (e.g., 'What other examples of this type of conflict do you know of?'). Adler found that the latter had a significant effect on what students could remember from the text.

— Is there an explicit link between written text and image?

— Is there metacommunication so that the student is explained how different forms of representation create meaning?

— Is there explicit instruction for the student on how to process the image?

2.6 Image as an object for academic student activity that strengthens multimodal literacy

In the previous section, the relationship between written text and image was discussed in relation to supporting students´ acquisition of academic content. Among other things, it was described how student activities can be linked to an illustration, or how the image can be provided with metacommunication.

However, it is also important to analyze further when the image is the object of analysis, as an attempt to strengthen the student's subject-disciplinary, multimodal literacy. This means that the visual element becomes subject content in itself and does not have the function of making other subject content more understandable. The image is a primary text. Examples are when a painting is to be analyzed in the L1 subject, when students analyze a bar graph in mathematics and learn how to read it, or when students learn to read a map in geography. Here, the student uses the subject's methods and concepts to analyze the visual object.

Is the image the subject of academic analysis? Describe, as precisely as possible, the disciplinary perspective underlying the analysis.

2.7 Decorative visual elements

Decorative visuals (Levin et al., 1987) are defined as being decidedly unacademic, as they do not meaningfully support the written text in the instructional material. Mayer formulated the principle of coherence, which states that we learn better from a multimedia resource if superfluous material is omitted, as we can better focus on the essentials when there are no distracting elements. In line with this, Jaeger and Wiley (2014) showed that decorative images impaired students' accuracy in relation to meta-understanding of science subject text, i.e., how good they are at assessing their own understanding. How-

ever, other research indicates that decorative images may not be so detrimental to learning outcomes.

Scherer, Verkühlen & Dutke (2023) thus found that so-called related *decorative images* positively influenced knowledge acquisition and metacognition. Related decorative images are images that do not function, for example, in a representational or explanatory way in relation to the written text: they do not convey relevant information specific to the text, but they remain completely general within the topic of the text (e.g., a picture of a tank accompanying a text about World War II).

Lenzner, Schnotz, and Müller (2013) used Mayer's theory of integrating multimedia into learning resources as a starting point to examine the effects of decorative images versus instructional images. They found that decorative images attracted less attention and had minimal distracting effects on students. Furthermore, decorative images caused better mood, alertness, and calmness in 7th-8th grade students, but did not significantly affect students' situational interest or their assessment of the difficulty of the material. Decorative images neither directly harmed nor directly benefited learning outcomes but moderated the beneficial effects of instructional images (i.e., the multimedia effect, Mayer, 2014b), especially for students with lower prior knowledge.

Children's picture books also show different variants of interaction between written language and images, such as complementarity and ironic counterpoint (Nikolajeva & Scott, 2000), the purpose of which is aesthetic effects rather than communicative.

Does the image have a purely decorative function?
Is it a related decorative image?

3. Discussion

I hope that the presented analysis model can be relevant for both teaching material producers, teachers, and student teachers. It is of great importance that we understand and map:

— the demands that students encounter through the teaching materials' use of, among other things, images,

— how students are best supported through images in their acquisition of the subject matter,

— how teaching materials and teaching can contribute to strengthening students' multimodal, subject-disciplinary literacy,

— and how teaching materials in individual subjects and across subjects can best be designed to contribute to the above-mentioned objectives.

The design of this analysis approach is the first step towards gaining the necessary insight into the teaching materials' use of images. Using the model in analytical work that spans across subjects and teaching materials is the next logical step. An attempt at this will be described in Gissel et al. (2026a).

It is an open question whether disciplinary differences in the use of images in instructional texts actually exist. This is the assumption in this article. However, the literature review indicates that the disciplinary-specific focus, not least in relation to research on multimodality in instructional materials, has been limited. I attempt to change this with this article's proposal for an analytical approach.

References

Adler, C. (1993). *Directed picture processing: The effects for learners on recall of related text* (Dissertation). University of South Florida.

Alkhateeb, M. (2019). Multiple Representations in 8th Grade Mathematics Textbook and the Extent to which Teachers Implement Them. *International Electronic Journal of Mathematics Education, 14*(1), 137-145. https://doi.org/10.12973/iejme/3982

Alvermann, D. E., & Wilson, A. A. (2011). Comprehension Strategy Instruction for Multimodal Texts in Science. *Theory Into Practice, 50*(2), 116–124. https://doi.org/10.1080/00405841.2011.558436

Bertin, J. (1983). *Semiology of Graphics*. The University of Wisconsin Press.

Bliss, T. (1990). Visuals in perspective: An analysis of U.S. history textbooks. *Social Studies, 81*, 10–10. doi:10.1080/00377996.1990.9957485

Boeriis, M. (2012). Tekstzoom: Om en dynamisk funktionel rangstruktur i visuelle tekster. In M. Boeriis & T. H. Andersen (Eds.), *Nordisk Socialsemiotik: pædagogiske, multimodale og sprogvidenskabelige landvindinger* (pp. 131-153). Syddansk Universitetsforlag.

Brugar, K. A., & Roberts, K. L. (2017). Seeing is believing: Promoting visual literacy in elementary social studies. *Journal of Teacher Education, 68*(3), 262–279. https://doi.org/10.1177/0022487117696280

Bundsgaard, J., & Hansen, T. I. (2011). Holistic evaluations of learning materials. In Rodríguez, J. R., Horsley, M., & S. V. Knudsen (Eds.), *Local, National and Transnational Identities in Textbooks and Educational Media: Ten International Conference on Research on Textbooks and Educational Media September 2009 Santiago de Compostela – Spain* (pp. 502-520). Santiago: IARTEM.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197. DOI: 10.1037/0022-0663.98.1.182.

Carney, R., & Levin, J. R. (2002). Pictorial Illustrations Still Improve Students' Learning from Text. *Educational Psychology Review, 14*(1). 10.1023/A:1013176309260

Cheng, M.-C., Chou, P.-I., Wang, Y.-T., & Lin, C.-H. (2015). Learning Effects of a Science Textbook Designed with Adapted Cognitive Process Principles on Grade 5 Students. *International Journal of Science and Mathematics Education, 13*, 467–488. https://doi.org/10.1007/s10763-013-9492-8

Coleman, J. M., McTigue, E. M., & Smolkin, L. B. (2011). Elementary teachers' use of graphical representations in science teaching. *Journal of Science Teacher Education,* 22, 613–643.

Danielsson, K. (2010). Learning Chemistry. Text Use and Text Talk in a Finland-Swedish Chemistry Classroom. *IARTEM e-journal, 3*(2), 1–28.

Danielsson, K., & Selander, S. (2016). Reading Multimodal Texts for Learning – a Model for Cultivating Multimodal Literacy. *Designs for Learning, 8*(1), 25–36. http://dx.doi.org/10.16993/dfl.72

Danielsson, K. & Selander, S. (2021). *Multimodal Texts in Disciplinary Education: A Comprehensive Framework.* Springer. https://doi.org/10.1007/978-3-030-63960-0

Devetak, I., Vogrinc, J., & Glažar, S. A. (2010). States of Matter Explanations in Slovenian Textbooks for Students Aged 6 to 14. *The International Journal of Environmental and Science Education, 5*(2), 217-235.

Fingeret, L. (2012). *Visuals in children's informational texts: A content analysis* (Unpublished doctoral dissertation). Michigan State University. https://www.proquest.com/docview/1039317370?parentSessionId=URjPHmFSg4p17D-TEpRuAaX6QbyQlf9KGIvcmZbCgGf8%3D&sourcetype=Dissertations%20&%20Theses

Fterniati, A. (2009). Teaching Multimodality in Greek Elementary School Language Arts. *The International Journal of Learning, 16*(4), 1-15.

Gissel, S. T. (2023). *PISA 2022 Læsning: Delrapport [PISA 2022 Reading: interim report].* (1 udg.) VIVE Det Nationale Forsknings- og Analysecenter for Velfærd. https://www.uvm.dk/-/media/filer/uvm/int/231204-pisa-2022-laesning-pdf-ua.pdf

Gissel, S. T. (2026a). Typologi for læremidler og ressourcer, der kan bruges til læring eller undervisning [Typology for instructional resources and resources, that can be used for purposes of learning or teaching]. *MONA - Matematik- Og Naturfagsdidaktik, 26*(1). https://doi.org/10.7146/mona.v26i1.166713

Gissel, S. T. (2026b). *Mediedidaktik i teori og praksis [Media didactics in theory and practice]*. Copenhagen: Academica/Hans Reitzels forlag. 3rd edition.

Gissel, S. T., Buch, B., Oksbjerg, M., Kjeldsen, K., & Lytje, M. (2023). Lærervejledningslandskabet i Danmark [The teacher guide landscape in Denmark]. *Learning Tech – Tidsskrift for læremidler, didaktik og teknologi*, (12), 41-74. https://doi.org/10.7146/lt.v7i12.132029

Gissel, S., Carlsen, D., & Buch, B. (2025). *Teachers' Use of Learning Materials*. In: Bagoly-Simó, P., Khaneboubi, M., Knight, B. (Eds.). Educational Media: International Perspectives. Oxford, United Kingdom: Peter Lang Verlag. Retrieved Dec 1, 2025, from https://doi.org/10.3726/b21556

Gissel, S. T., Carlsen, D., Kinnerup, L.B., Kjeldsen, K., Møller, M. & Oksbjerg, M. (2026a). Fagdidaktisk anvendelse af multimodalitet på danske digitale fagportaler. Signalement af samspillet mellem skrevet tekst og visuelle elementer i udvalgte fag. *Learning Tech – Tidsskrift for læremidler, didaktik og teknologi*, 17, 79-114.

Gissel, S. T., Kjeldsen, K., Carlsen, D., Oksbjerg, M., Møller, M., Nielsen, T., Kinnerup, L.B., & Johannesen, C. (2026b). A scoping review of research on disciplinary use of graphics-text-constructs in didactic learning materials. L*earning Tech – Tidsskrift for læremidler, didaktik og teknologi*, 17, 17-46.

Guo, D., Wright, K. L., & McTigue, E. M. (2018). A Content Analysis of Visuals in Elementary School Textbooks. *The Elementary School Journal, 119*(2), 245-265. https://doi.org/10.1086/700266

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science. Literacy and discursive power.* University of Pittsburgh Press.

Hansen, J. J. (2006). *Mellem design og didaktik: Om digitale læremidler i skolen.* (Ph.d.-afhandling). Syddansk Universitet.

Hung, Y.-N. (2014). "What Are You Looking At?" An Eye Movement Exploration in Science Text Reading. *International Journal of Science and Mathematics Education, 12*, 241–260. DOI: 10.1007/s10763-013-9433-6

Iser, W. (1971). Indeterminacy and the Reader's Response. In J. Hillis Miller (ed.), *Aspects of Narrative: Selected Papers from the English Institute* (pp. 1-45). Columbia University Press.

Jaeger, A. J., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction, 34*(6), 58–73. https://doi.org/10.1016/j.learninstruc.2014.08.002

Janko, T., & Knecht, P. (2014). Visuals in Geography Textbooks and Increasing the Reliability of a Research Instrument. In: P. Knecht, E. Matthes, S. Schütze, & B. Aamotsbakken (Eds.), *Methodologie und Methoden der Schulbuch- und Lehrmittelforschung* (pp. 227–239). Klinkhardt.

Kapıcı, H. Ö., & Açıkalın, F. S. (2015). Examination of visuals about the particulate nature of matter in Turkish middle school science textbooks. *Chemistry Education Research and Practice, 16*, 518-536. https://doi.org/10.1039/C5RP00032G

Karlsson, A.-M. (2012). Multimodaliteten och språket: Om lexikogrammatikens styrkor och begränsningar i förhållande till några arbetslivsdiskurser. In M. Boeriis & T. H. Andersen (Eds.), *Nordisk Socialsemiotik: pædagogiske, multimodale og sprogvidenskabelige landvindinger* (pp. 17–38). Syddansk Universitetsforlag.

Knudsen, B., & Buttenschøn, M. (n.d.). *Anden Verdenskrig [Second World War]*. https://historie.alinea.dk/course/AA03-anden-verdenskrig

Kress, G., & van Leeuwen, T. (2006). *Reading Images. The Grammar of Visual Design*. Routledge.

Lee, V. R. (2010). Adaptations and Continuities in the Use and Design of Visual Representations in US Middle School Science Textbooks. *International Journal of Science Education, 32*(8), 1099-1126.

Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41*, 811-831. https://doi.org/10.1007/s11251-012-9256-z

Levin, J. R. (1981). On functions of pictures in prose. In F. J. Pirozzolo, and M. C. Wittrock (Eds.), *Neuropsychological and Cognitive Processes in Reading* (pp. 203–228). Academic Press.

Levin, J. R., Anglin, G. J., and Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows and H. A. Houghton (Eds.), *The Psychology of Illustration*: I. Basic Research (pp. 51–85). Springer.

Lim, F. V., & Toh, W. (2020). How to teach digital reading? *Journal of Information Literacy, 14*(2), pp. 24-43.

Liu, H. C., & Chuang, H. H. (2011). An examination of cognitive processing of multimedia information based on viewers' eye movements. *Interactive Learning Environments, 19*(5), 503-517. https://doi.org/10.1080/10494820903520123

Mason, L., Pluchino, P. & Tornatora, M.C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology, 41,* pp.172–187. doi:10.1016/j.cedpsych.2015.01.004.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*(1), 31-43.

Mayer, R. E. (Ed). (2014a). *Multimedia learning* (2nd ed). Cambridge University Press.

Mayer, R. E. (2014b). Research Based Principles for Designing Multimedia Instruction. In Benassi, V. A., Overson, C., & Hakala, C. M. (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 59-70). http://teachpsych.org/ebooks/asle2014/index.php

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715–726. https://doi.org/10.1037/0022-0663.82.4.715

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*(1), 31–43. https://doi.org/10.1007/BF02300480

McKean, B. (2002). Artistic Representations in Three Fifth-Grade Social Studies Textbooks. *The Elementary School Journal, 103*(2), 187-197. The University of Chicago Press. https://www.jstor.org/stable/1002234

Molina, A. I., Navarro, Ó., Ortega, M., & Lacruz, M. (2018). Evaluating multimedia learning materials in primary education using eye tracking. *Computer Standards & Interfaces, 59*, 45-60. https://doi.org/10.1016/j.csi.2018.02.004

Nikolajeva, M., Scott, C. (2000). The Dynamics of Picturebook Communication. *Children's Literature in Education, 31*(4), 225–239. https://doi.org/10.1023/A:1026426902123

OECD. (2019). *PISA 2018 Reading Framework*. OECD Publishing.

OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Ohle-Peters, A., McElvany, N., & Ullrich, M. (2023). Students Reading Text-Picture-Material: The Role of Teacher Competence and Instructional Quality. *International Journal of Educational Psychology, 12*(3), 251-268. https://doi.org/10.17583/ijep.11558

Oruç, Ş., Uşurlu, N. B., & Tokcan, H. (2010). Using graphic illustrations with social studies textbooks. *Procedia Social and Behavioral Sciences,* 2, 1037–1042. https://doi.org/10.1016/j.sbspro.2010.03.146

Peirce, C. S. (1997). *Collected Papers of Charles Sanders Peirce* (C. Hartshorne & P. Weiss, Eds.). Thoemmes Press.

Peterson, M. (2016). Schemes for Integrating Text and Image in the Science Textbook: Effects on Comprehension and Situational Interest. *International Journal of Environmental and Science Education, 11*(6), 1365–1385.

Postigo, Y., & López-Manjón, A. (2015). Graphicacy in biology textbooks: Analysis of activities with images. *Infancia y Aprendizaje*, 1-33. DOI: 10.1080/02103702.2015.1054667

Pozzer, L. L, & Roth, W. M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. *Journal of Research in Science Teaching, 40*(10), 1089-1114. https://doi.org/10.1002/tea.10122

Roberts, K. L., Norman, R. R., Duke, N. K., Morsink, P., Martin, N. M., & Knight, J. A. (2013). Diagrams, timelines, and tables—Oh, my! Fostering graphical literacy. *Reading Teacher, 67*, 12–24.

Scherer, D., Verkühlen, A. & Dutke, S. (2023). Effects of related decorative pictures on learning and metacognition. *Instructional Science, 51*, 571–594. https://doi.org/10.1007/s11251-023-09618-8

Schleppegrell, M. (2004). *The Language of Schooling. A Functional Linguistics Perspective.* Lawrence Erlbaum.

Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology, 106*(4), 974–989. https://doi.org/10.1037/a0037054

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*(1), 40–59.

Shanahan, T., & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in Language Disorders, 32*(1), 7–18.

Shin, W. (2017). Analysis of elementary school students' visual attention on the editorial design of 'structure and function of our body' in the 2007-2009 revised elementary science textbook. *Journal of Korean Elementary Science Education, 36*(4), 428-438.

Šimik, O. (2021). Analysis of Visual Components in Czech History Textbooks for Lower Grades of Elementary Schools. In P. Bagoly-Simó & Z. Sikorová (Eds.), *IARTEM 2015 Textbooks and Educational Media: Perspectives from Subject Education* (pp. 119–133). Springer. https://doi.org/10.1007/978-3-030-80346-9_10

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science Textbooks' Use of Graphical Representation: A Descriptive Analysis of Four Sixth Grade Science Texts. *Reading Psychology, 31*(3), 301-325. https://doi.org/10.1080/02702710903256502

van Leeuwen, T. (2005). *Introducing Social Semiotics*. Routledge.

Wright, K.L., McTigue, E.M., Eslami, Z.R., University, T.A., & Reynolds, D. (2014). More than just eye-catching: Evaluating graphic quality in middle school English language learners' science textbooks. *Journal of Curriculum and Instruction, 8*, 89–109. https://doi.org/:10.3776/joc-i.2014.v8n2p89-109

Keywords

Instructional materials; Educational resources; Multimodality; Text-image-relations; Disciplinary literacy; Content area literacy