Visual Representations in Economic Education From an Interdisciplinary Perspective

Dissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

Malte Ring

aus Herrenberg

Tübingen

2020

1. Betreuerin:	Prof. Dr. Taiga Brahm
2. Betreuer:	Prof. Dr. Christoph Randler

Tag der mündlichen Prüfung:	02.11.2020
Dekan:	Prof. Dr. Josef Schmid
1. Gutachterin:	Prof. Dr. Taiga Brahm
2. Gutachter:	Prof. Dr. Andreas Lachner

Danksagung

An erster Stelle möchte ich meiner Betreuerin Prof. Dr. Taiga Brahm und meinem Betreuer Prof. Dr. Christoph Randler für die vielseitige konstruktive Unterstützung meiner gesamten Promotion danken. Ein besonderer Dank geht dabei an Taiga: Danke für die vielen spannenden (lustigen!) Gespräche, die maßgeblich zum Gelingen dieser Arbeit beigetragen haben. Danke für die guten Antworten auf meine kleinen und großen Fragen und die Möglichkeit, in einem konstruktiven und sympathischen akademischen Umfeld „aufzuwachsen“ und meinen Horizont zu erweitern.

Auch meinen Koautorinnen, Prof. Dr. Katharina Scheiter und Dr. Juliane Richter, möchte ich für die spannende Zusammenarbeit und die Unterstützung bei der Erstellung des Artikels danken. Ein herzliches Dankeschön möchte ich auch an Prof. Dr. Andreas Lachner aussprechen, nicht nur für die Bereitschaft, meine Dissertation zu begutachten, sondern auch für die kleinen Ratschläge während meiner Promotion.

Ein großer Dank gebührt meinen Teamkolleginnen. Allen voran Marina Pumptow: Danke für die schöne Zeit, die vielen spannenden Gespräche und die Unterstützung bei allen statistischen und inhaltlichen Fragen. Vielen Dank an Dr. Michelle Rudeloff, Dr.(!) Ute Grewe, Markus Trämer und Luisa Scherzinger für die moralische und inhaltliche Unterstützung, den kreativen Austausch von Ideen, die kleinen Tür-und-Angel-Gespräche und die schöne gemeinsame Zeit.

Ganz herzlich möchte ich mich auch bei allen Kolleginnen und Kollegen der Tübingen School of Education sowie der LEAD Graduiertenschule & Forschungsnetzwerk bedanken, die mich auf dem Weg begleitet, unterstützt und positiv beeinflusst haben. Danke an Ulrich Ludewig für die fachlichen Ratschläge in der Anfangszeit und danke an Sarah Bez für den konstruktiven Austausch und dein Feedback auf den letzten Metern meiner Dissertation. Ein besonderer Dank geht an Patrizia Breil: Danke für dein offenes Ohr, die vielen anregenden (fachlichen) Diskussionen und deine Unterstützung in allen Belangen!

Nur durch die Beteiligung der Praxis kann praxisnahe Forschung gelingen. Daher möchte ich auch den Studierenden, den Schülerinnen und Schülern sowie Lehrpersonen danken, die mich in meiner Arbeit unterstützt und an meinen Studien teilgenommen haben.

Vielen Dank an meine Familie, die mir stets zur Seite gestanden ist. Zum Abschluss möchte ich mich bei meiner Partnerin Sina Bahlke bedanken: Danke, dass du immer da bist, danke für das Verständnis für meine Arbeit und danke für die großartige Begleitung auf meinem Weg!

Summary

Visual representations such as graphs and charts are important tools to make data and models more understandable. In our daily life, we are confronted with graphs that visualize information, such as election results or the development of stock prices. Across disciplines, graphs and charts are also used in research and teaching to explain domain concepts, for example when the relationship between price, demand and supply is modeled in economics or when biologists analyze how prey and predator populations influence each other over time. The ability to understand these visual representations is therefore not only necessary in daily life but also part of domain expertise. Although graphs and charts are omnipresent in the 21st century, prior research has demonstrated that all students cannot be assumed to intuitively understand these visualizations. In contrast, large-scale and in-depth studies have identified various difficulties of learners, for example when they are unable to read and interpret data graphs or cannot connect visual representations to the underlying domain principle. Different research communities have thus modeled and analyzed the ability of learners to work with visual representations (mostly in the science domains) and investigated the effect of instructional support, which helps learners to understand graphical representations and relate them to their context. Bringing together these different research disciplines, this dissertation investigates, in three separate studies, how learners can read data graphs; how visual representations are used in secondary economic education; and lastly, how learners can be supported in integrating graphs and text.

In the first study, eighth graders' ability to read graphs was investigated. Focusing on graphs related to sustainable development, which students could encounter in their everyday life, the study measured how well they could read single data points and trends or perform small extrapolations. The instrument was used with 198 students from four different schools, all with the highest school track (Gymnasium). To test whether tasks with increasing complexity (from points to extrapolations) would also be more difficult for learners, the data was analyzed with item response theory. Furthermore, the relationship between graph reading and learner characteristics such as academic performance, motivation, interest and domain knowledge was examined. The results revealed that eighth graders were able to read data graphs rather well, and no systematic relationship was found between what an item asked for and the item difficulty. The ability to read graphs correlated with academic language performance, academic math performance, as well as content knowledge and prior engagement with sustainable development.

In the second study, a textbook analysis was combined with teacher interviews to investigate the use of graphical representations in learning material and instruction. To gain an overview of the graphical representations in textbooks, they were categorized according to their form (graph/chart) and the extent to which they visualize a domain principle. In 10 semi-structured interviews, teachers were asked how and why they use graphical representations, what they expect of their students and what typical mistakes their students make when they work with visual representations. The teacher interviews revealed that graphs and charts are used regularly in teaching, not only to visualize economic models but also to display data related to economic variables (e.g., development of growth domestic product [GDP]) and to train students in critically analyzing graphical representations. The following are among the challenges for learners: math- and data-related issues (e.g., when learners are unable to differentiate between absolute and relative numbers) and a lack of integration of representation and domain (e.g., when learners cannot identify the relevant information for a domain question or are unable to connect graphical information to other external representations such as texts).

Finally, through a quasi-experimental design, the third study tested how learners can be supported in learning with text and graphs. For this purpose, students received learning material from two domains in one of three conditions: Either the correspondences between text and graph were already highlighted or they were asked to highlight the relevant connections themselves while studying the material (active signals) or, lastly, without alterations to design or learner-task (control). After the study phase, the learning outcome was tested with recall and comprehension questions. Overall, students who studied already signaled material performed equally well compared to the control group. On average, students in the active-signal group achieved significantly fewer points in the biology posttest compared to the control group. When learners had high prior knowledge, however, they could profit from actively integrating both representations. Furthermore, in economics, the relationship between prior knowledge and learning outcome was partially mediated via the quality of learner-generated signals; that is, learners with high prior knowledge were better in connecting graphs and text, which in turn was associated with higher learning outcomes.

In this dissertation, the findings of these three studies are summarized and discussed against the background of the research context in different disciplines (economic education, science education and educational psychology). At the end, implications for future research and educational policy and practice are derived.

Zusammenfassung

Visuelle Repräsentationen wie Diagramme und Graphen werden verwendet, um Modelle und Daten zu veranschaulichen. Datendarstellung, in denen beispielsweise Wahlergebnisse zusammengefasst oder die Entwicklung der Aktienkurse gezeigt werden, begegnen uns im Alltag daher immer wieder. Auch in Wissenschaft und Lehre werden visuelle Repräsentationen eingesetzt, um zentrale Konzepte anschaulich darzustellen. Das Preis-Mengen-Diagramm in der Wirtschaftswissenschaft oder die Populationsdynamik von Räuber-Beute-Beziehungen in der Biologie sind dafür typische Beispiele. Die Fähigkeit, mit solchen Visualisierungen umzugehen, ist daher nicht nur relevant für den Alltag, sondern spielt auch für die Entwicklung von Fachkompetenz eine zentrale Rolle. Obwohl Diagramme und Graphen im 21. Jahrhundert allgegenwärtig sind, kann nicht davon ausgegangen werden, dass sie für Lernende ohne Unterstützung und Übung verständlich sind. Vielmehr haben unterschiedliche Studien gezeigt, dass sich bei der Arbeit mit Diagrammen verschiedene Schwierigkeiten ergeben, dazu zählen beispielsweise Fehler beim Lesen der Diagramme oder bei der Verknüpfung von visuellen Repräsentationen mit Fachinhalten. Unterschiedliche Forschungsrichtungen haben daher Modelle für Diagrammkompetenz entwickelt und untersucht, wie Lernende beim Lesen von Diagrammen und bei der Verbindung von Diagramm und Fachinhalt unterstütz werden können. Die vorliegende Dissertation untersucht auf Basis dieser Befunde in drei unterschiedlichen Studien, wie gut Schülerinnen und Schüler Diagramme lesen, wie visuelle Repräsentationen im Wirtschaftsunterricht genutzt werden und wie Lernende bei der Verknüpfung von visueller Repräsentation und dazugehörigem Text unterstützt werden können.

Im Rahmen der ersten Studie wurde die Diagrammlesekompetenz von 198 Achtklässlerinnen und Achtklässlern an vier verschiedenen Gymnasien untersucht. Dafür wurden typische Datendarstellungen aus dem Themenkomplex Nachhaltige Entwicklung verwendet, denen Schülerinnen und Schüler in ihrem Alltag begegnen könnten. Für diese wurde jeweils untersucht, wie gut die Schülerinnen und Schüler einzelne Datenpunkte und Trends lesen, aber auch, ob sie auf Basis dessen extrapolieren können. Um zu untersuchen, ob die Schwierigkeit mit steigender Komplexität der Aufgaben zunimmt, wurden die Daten mithilfe der *Item Response Theory* ausgewertet. Zusätzlich wurde der Zusammenhang der Diagrammlesekompetenz mit anderen Variablen untersucht, darunter beispielsweise Deutsch- und Mathematiknote, Motivation, Interesse und Wissen über nachhaltige Entwicklung. Insgesamt war die Diagrammlesekompetenz der Schülerinnen und Schüler gut ausgeprägt, ein Zusammenhang zwischen Aufgabenanforderung (Datenpunkt, Trend, Extrapolation) und

Schwierigkeit hat sich nicht gezeigt. Weiterhin konnte nachgewiesen werden, dass die Diagrammlesekompetenz positiv mit der Leistung in Deutsch und Mathematik sowie mit dem Wissen über nachhaltige Entwicklung zusammenhängt. Auch wenn sich Schülerinnen und Schüler in den Wochen zuvor mit dem Thema auseinandergesetzt haben, konnten positive Auswirkungen auf die Diagrammlesekompetenz festgestellt werden.

Mithilfe einer Schulbuchanalyse und Interviews mit Lehrpersonen wurde in der zweiten Studie untersucht, wie visuelle Repräsentationen im wirtschaftlichen Schulfächern eingesetzt werden. Hierfür wurden zunächst die visuellen Repräsentationen in Wirtschaftsschulbüchern im Hinblick auf ihre Form und Domänenspezifität analysiert um typische Visualisierungen des Fachs zu identifizieren. Darüber hinaus wurden in zehn leitfadengestützten Interviews Lehrpersonen gefragt, welche Diagramme sie (warum) in ihrem Unterricht einsetzen, was sie dabei von ihren Schülerinnen und Schülern erwarten und welche typischen Schwierigkeiten sie bei ihren Schülerinnen und Schülern im Umgang mit Diagrammen feststellen. Dabei zeigte sich, dass Lehrende Diagramme und Graphen zu unterschiedlichen Zwecken einsetzen. Dazu gehören neben der Illustration von typischen Fachkonzepten auch die Orientierung an aktuellen Statistiken wirtschaftlicher Indikatoren, sowie eine gezielte Förderung von Diagrammkompetenz. Schwierigkeiten auf Seiten der Lernenden ergeben sich dabei nicht nur beim Umgang mit Zahlen (zum Beispiel bei der Unterscheidung von absoluten und relativen Zahlen), sondern auch im Hinblick auf die Identifikation zentraler Konzepte und deren Bezüge zu anderen Repräsentationsformen wie Fachtexten.

Basierend auf einem quasi-experimentellen Design wurde in der dritten Studie untersucht, wie Studierende bei der Verknüpfung von Diagramm und Fachkontext unterstützt werden können. Im Rahmen der Studie wurden Lernende zunächst gebeten, biologische und wirtschaftliche Fachinhalte mithilfe von ausgehändigtem Material zu lernen. Das Material lag dabei in einer von drei Bedingungen vor: (1) zentrale äquivalente Informationen in Text und Diagramm waren in der gleichen Farbe hervorgehoben, (2) zentralen Informationen in Text und Diagramm sollten während des Lernens von den Lernenden selbst hervorgehoben werden oder (3) ohne Hervorhebungen (Kontrollgruppe). Im Anschluss wurde die Erinnerungs- und Verstehensleistung erfasst. Die Ergebnisse zeigen, dass bereits bestehende Hervorhebungen in Text und Diagramm keinen signifikanten Einfluss auf die Lernleistung haben. Wenn Lernende die Hervorhebungen während des Lernens selbst übernehmen mussten, wirkte sich das in Biologie im Durchschnitt negativ auf die Lernleistung aus. Lediglich Lernende mit hohem Vorwissen konnten von der dieser Bedingung in Biologie profitieren. Die Beziehung zwischen Vorwissen und Lernleistung wurde dabei teilweise über die Qualität der Hervorhebung mediiert, das heißt, Lernende mit hohem Vorwissen waren besser darin, die zentralen äquivalenten Informationen hervorzuheben, was wiederum positiv mit der Lernleistung zusammenhing.

Im Rahmen der vorliegenden Dissertation werden die Ergebnisse der drei Studien zusammengefasst und vor dem Hintergrund des breiten Forschungshintergrunds (Wirtschaftsdidaktik, Naturwissenschaftsdidaktik, Pädagogische Psychologie) diskutiert. Abschließend werden aus den Ergebnissen Implikationen für zukünftige Forschung und Praxis abgeleitet.

Content

List of figures

Chapter 1 Introduction and theoretical framework

Chapter 2 Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs

Chapter 3 Logical pictures in secondary economic education: textbook analysis and teacher perception

Chapter 4 How to support text-graph integration: Comparing the effects of passive and active signaling on learning outcomes

List of tables

List of abbreviations

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CLT	Cognitive Load Theory
CTML	Cognitive Theory of Multimedia Learning
DEFT	Design, Functions, Tasks
FAO	Food and Agriculture Organization
GDP	Growth Domestic Product
ICAP	Interactive, Constructive, Active, Passive
IRT	Item Response Theory
ITPC	Integrated model of Text and Picture Comprehension
LGVT	Lesegeschwindigkeits-und -Verständnistest
MMLE	Marginal Maximum Likelihood Estimates
PISA	Programme for International Student Assessment
STEM	Science, Technology, Engineering, Mathematics
TIMSS	Trends in International Mathematics and Science Study
TOGS	Test of Graphing in Science
WLE	Weighted Likelihood Estimations

1 Introduction and theoretical framework

1.1 Relevance of external representations

"*The world cannot be understood without numbers. And the world cannot be understood with numbers alone*." is a quote from the book, Factfulness: Ten Reasons We're Wrong About the World – and Why Things Are Better Than You Think, by Hans Rosling, Ola Rossling and Anna Rosling-Rönnlund (2018, p. 225). In this book, the authors argue that the world is not as poor or underdeveloped as we generally tend to think. The book is highly praised (Brueck, 2018; Gates, 2018; Goerwitz, 2019; Law, 2019), not only because it makes a sound argument but also because it leads the reader along the arguments mainly by displaying data visualizations and contextualizing them with text information. It is one of many books that can be easily utilized to make the argument that visual displays of data ("numbers") as well as their context ("not numbers alone") in combination are important tools of communication in the 21st century. We live in a data-driven world: Not only in the scientific community, where graphs often depict the most important information, but also in our everyday life, data visualizations are used, for example, to convey information about election results in newspapers or to provide an overview of our account balances.

Graphical representations (i.e., not only data graphs but also other visualizations such as diagrams or flow charts) are generally used to represent important ideas and models in different domains across the (social) sciences. In biology, for instance, some of the most important ideas have been conveyed as graphs; the theory of island biogeography (Figure 1, upper-left panel, MacArthur & Wilson, 2001) or the prey–predator relationship (Figure 1, lower-left panel, Volterra, 1928) are prominent examples. The same is true for economics as a social science: Marshall's supply-and-demand graph, which depicts price as a function of demand and supply, is one of the most important models and is still taught in microeconomics in higher education today (Figure 1, upper-right panel, Marshall, 2009). Another example, a visualization of the model of the circular flow of income, is commonly used in economic education in high schools (Figure 1, lower-right panel, Marks & Kotula, 2009).

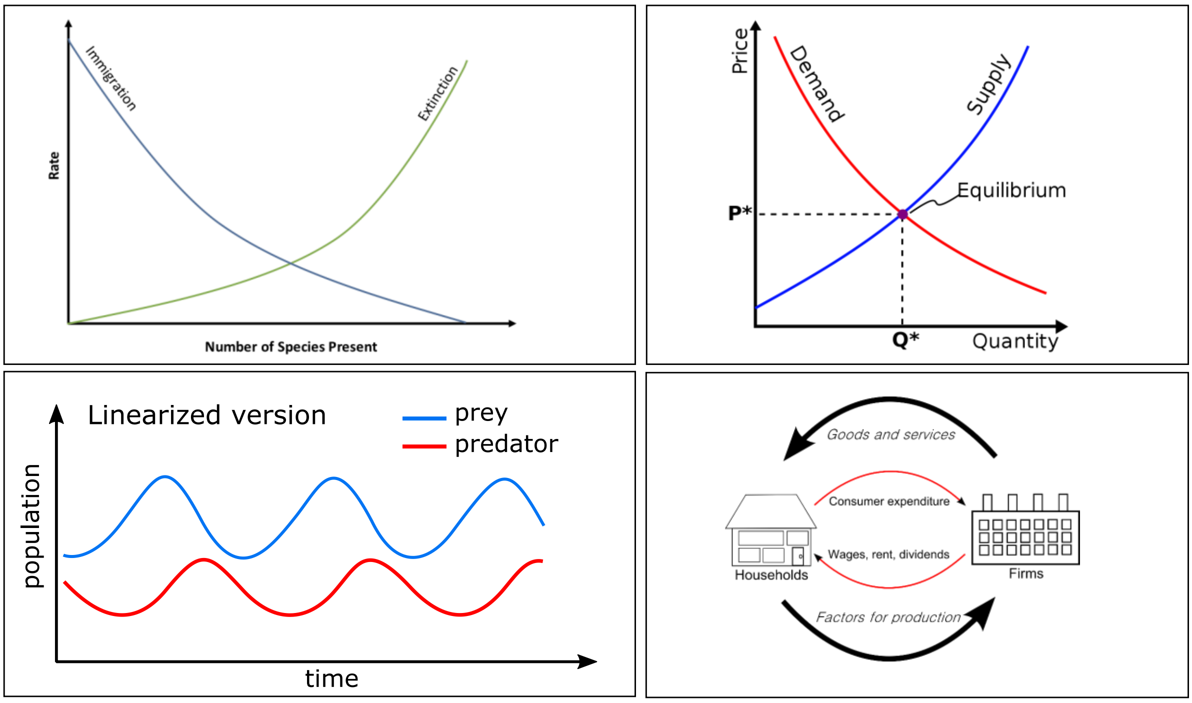


Figure 1. Examples of common graphical representations in biology and economics

Upper-left panel: User Marcus Lapeyrolerie (2018) [The theory of island biogeography], Wikipedia, https://commons.wikimedia.org/wiki/File:Equilibrium_Model_for_Number_of_Species_on_an_Island.png, CC BY-SA 4.0
Upper-right panel: User SilverStar (2015) [Supply and demand equilibrium], Wikipedia, https://commons.wikimedia.org/wiki/File:Supply-demand-equilibrium.svg, CC BY-SA 3.0
Lower-left panel: User ApsidistraK (2017) [Prey-Predator-Relationship], Wikipedia, https://commons.wikimedia.org/wiki/File:Lotka_Volterra_dynamics.svg, CC BY-SA 4.0
Lower-right panel: User Irconomics (2008) [Circular flow of income model], Wikipedia, https://commons.wikimedia.org/wiki/File:Circular_flow_of_goods_income.png, CC BY-SA 3.0

In light of the omnipresence of graphs and diagrams, learners should develop a thorough understanding of graphs, especially in secondary education (Roberts et al., 2013). In Germany, graphs and diagrams are part of different school curricula in multiple subjects. In economics, for example, the ability to illustrate and analyze graphical representations of economic learning content ("ökonomische Sachverhalte grafisch darstellen und auswerten", Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016b, p. 13) is expressed as one objective regarding general methodological competence; however, it also plays a role when working with specific graphs or diagrams, for instance when explaining the price-building mechanism with the help of a Marshall graph ("anhand eines Preis-Mengen-Diagramms die Preisbildung beim Polypol auf dem vollkommenen Markt und die Grenzen dieses Modells erklären", Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016b, p. 16). In biology, learners are expected to extract information from graphs, among other representations ("Informationen aus Texten, Bildern, Tabellen, Diagrammen oder Grafiken entnehmen," Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016a, p. 10), and they should be able to illustrate complex biological relationships in graphs, diagrams, charts and models ("komplexe biologische Sachverhalte mithilfe von Schemata, Grafiken, Modellen oder Diagrammen anschaulich darstellen”, Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016a, p. 10).

It is thus no surprise that items assessing graph comprehension are also integrated into large-scale studies such as the Trends in International Mathematics and Science Study (TIMSS; Baumert, Bos, & Watermann, 1998) and the Programme for International Student Assessment (PISA; Artelt et al., 2001). However, the ability to understand graphs and representations is tested not only as part of a “generic” data or scientific or statistical literacy construct in large-scale studies, but also in tests that intend to measure domain knowledge in economics (Beck, Krumm, & Dubs, 1998) or graph comprehension and domain knowledge in science (Lai et al., 2016; McKenzie & Padilla, 1986). The reason for this is that the ability to deal with representations of a domain is linked with domain knowledge (Ainsworth, 2006): We expect a scientist to understand a graph used by his or her peers, and we expect an economist to be able to understand the relationships depicted by Alfred Marshall (but not necessarily the other way around).

To answer a domain question with a graphical representation, it is necessary to (a) have basic reading and mathematical abilities, to (b) know the reading rules and relevant operators that should be used with the representation and to (c) have the relevant background or context knowledge to use the representation (together with other representations) to answer the question at hand (Friel, Curcio, & Bright, 2001).

For the Marshall graph, this means being able to (a) read the labels on the axis, (b) understand that the curves represent demand and supply and that price is modeled as the independent variable[1] and (c) use the graph to analyze how different factors such as tariffs and commodity prices might influence supply and demand. The latter, however, might not be concluded from the graph alone; for example, it would be necessary to at least have a basic understanding of how tariffs are related to supply and how it is represented in the model depicted in the graph. Integrating information from different sources (e.g., text, graphs, tables and formulas) and/or from prior knowledge thus seems to be required.

[1] In contrast to almost all graphs in (social) science, the independent variable (price) is depicted on the y-axis. The reason for this is that while Marshall thought of quantity as the independent variable, with prices adjusting to clear the market, later scholars though it more reasonable to model price as the independent variable that influences quantity. According to Humphrey (1992), modern economics agrees on the second interpretation, but further use Marshall´s graph to visualize the principle even though he was not the only one to visualize the relationship between supply, demand and price.

To better understand the ability to work with external representations, scholars in different research fields (see Figure 2) have categorized various graphical representations (e.g., Kosslyn, 1999; Levin, Anglin, & Carney, 1987; Lohse, Biolsi, Walker, & Rueter, 1994; Schnotz, 2001; Slough & McTigue, 2013; Winn, 1987), analyzed the ability of learners to work with graphical representations (e.g., Åberg-Bengtsson & Ottosson, 2006; Curcio, 1987; Friel et al., 2001; Galesic & Garcia-Retamero, 2011; Glazer, 2011; Gültepe, 2016; Kotzebue, Gerstl, & Nerdel, 2015; Lachmayer, 2008; Lai et al., 2016), discussed how different domains and disciplines use graphical representations in teaching and learning material (e.g., Aprea & Bayer, 2010; Cook, 2011; Jägerskog, 2020; Kozma, 2003; Reingewertz, 2013; Wu & Puntambekar, 2012) and tested methods that support learners to understand and connect multiple representations (e.g., Alpizar, Adesope, & Wong, 2020; Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Mautone & Mayer, 2001; Richter, Scheiter, & Eitel, 2016; Scheiter & Eitel, 2015; Schneider, Beege, Nebel, & Rey, 2018; Seufert, 2003, 2019; van Gog, 2014).

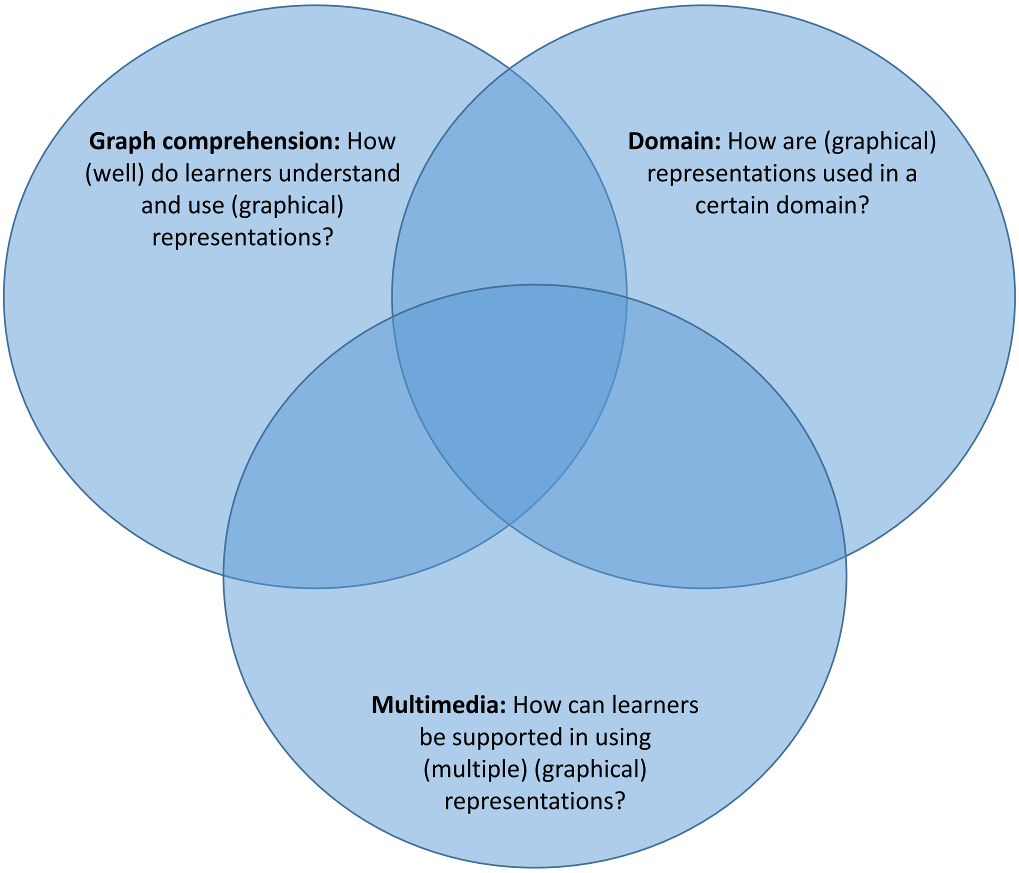


Figure 2. Research fields analyzing external representations from different perspectives (own illustration)

Since most studies in educational research have analyzed students’ ability to work with visual representations in the science domains, and because the ability to work with visual representations is at least partly domain-specific (Ainsworth, 2006), it remains unclear whether the results of graph comprehension and multimedia research in science and math education can be transferred to social science (Westelinck, Valcke, Craene, & Kirschner, 2005). Compared to a long research tradition in science and math education, research regarding graphical representation in economic education is scarce and has mostly focused on higher education (e.g., Cohn, Cohn, Balch, & Bradley, 2001). In line with findings from other domains, it could be demonstrated that in higher education, learners often fail to understand important visual representations (e.g., Strober & Cook, 1992). Regarding secondary economic education, studies have focused on the general benefits and pedagogical affordances of visual representations (e.g., Raso, 2018) and the influence of the representational format on learning outcomes (e.g., Jägerskog, 2020), or they have evaluated the design of illustrations in textbooks (e.g., Aprea & Bayer, 2010). Little is known, however, about how (well) learners work with graphical representations, how representations are used in learning material and teaching or how learners can be supported in integrating visual representations into other representations such as texts.

This lack of research is addressed in this dissertation, and the overarching goal of the present work is to connect different research fields (economic education, science education and educational psychology) to better understand how learners use (multiple) visual representations (in economics) and how they can be supported in that regard. To this end, multiple methodological approaches are employed to analyze how learners read authentic graphs in the context of sustainable development (Study 1), how visual representations are used in textbooks and teaching in secondary economics (Study 2) and how learners can be supported in the integration of graphs and text when learning with graph–text material in biology and economics (Study 3).

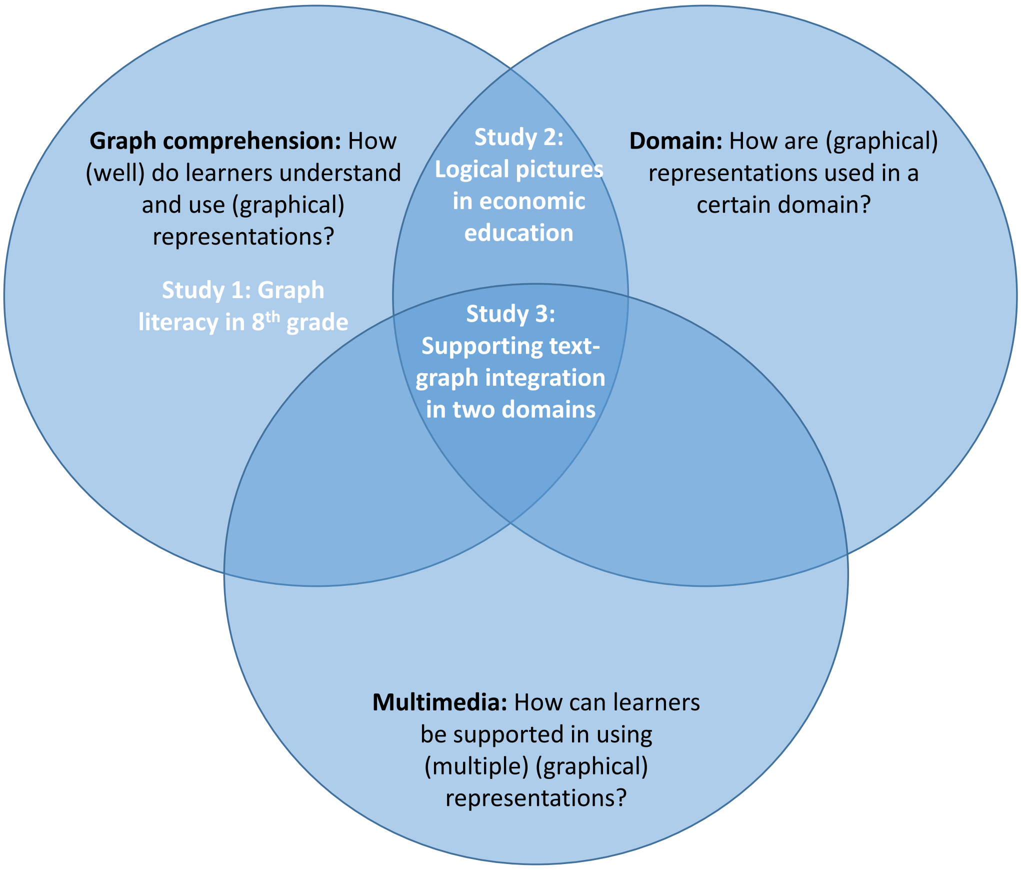


Figure 3. Placing the three studies in this thesis in the research fields (own illustration)

In the following sections, an overview of different external representations and the use of graphical representations in teaching is presented first; the focus here is mainly on the relationship between the visual representation and the domain as well as the research regarding graphical representations in economic education. Second, research in different fields – mainly science and math education – is summarized to describe the relevant abilities and challenges that learners face when they work with different visual representations. Third, the integration of multiple representations and the influence of instructional support on learning outcomes are discussed. The introduction closes with emerging research questions. Then, three studies follow that address these research questions. The first study analyzes the graph literacy of eighth graders in a German high school with the help of item response theory and relates their ability to read graphs to different school achievement factors. For the second study, teacher interviews were conducted and combined with an analysis of economic textbooks to discuss the use of representations in secondary economic education. In the third study, we analyze under which conditions (learner-generated) highlighting of equivalent information in text and graphs can be beneficial for learning outcomes in economics and biology. In the final chapter, the findings of all three studies are discussed regarding the theoretical, methodological and practical implications.

1.2 Use of graphical representation in teaching

1.2.1 Graphical representations in learning material

Graphical representations are generally discussed as part of external representations. External representations are "*structures in the environment that allow the learner to interact with some content domain*" (Vries, 2012, p. 2016); "external" specifies that (in comparison to internal representations) these "structures" are not cognitive constructions but actual structures in the material world. In other words, "external representation" is an umbrella term for different visualizations – not only graphs but also text, diagrams, figures, illustrations and pictures, among other things. Several authors have developed different taxonomies of external representations in learning material, based on various characteristics and for multiple purposes. Although some similarities exist, no classification is universally accepted yet (Ainsworth, 2006).

There are three major challenges when comparing and integrating the different classifications of external representations. First, the classifications have different scopes: While some classifications focus only on specific types of pictures (Winn (1987), for example differentiates, between graphs, charts and diagrams), others classify different external representations on a higher level and therefore differentiate between text, realistic pictures and logical pictures (e.g., Schnotz, 2001); newer classifications integrate digital formats as well (Ainsworth, 2014). Second, the various classifications use different terminology, a "chart" in one classification might be a "graph," a "diagram," an "infographic," or a "visual representation" in another framework (Harris, 2000; Kosslyn, 1999; Winn, 1987). Third, whereas most classifications differentiate between structural-form characteristics, other classifications distinguish pictorial elements in learning material according to the function they serve when they are combined with text. For example, Levin, Anglin and Carney (1987) distinguish between decorative, representational, organizational, interpretational and transformational pictures.

Decorative illustrations do not add information, but an affective component. In an economic textbook, for instance, a text explains how supply and demand are influenced by price using the strawberry market as an example (Bauer, Hamm-Reinöhl, Podes, & Riedel, 2012, p. 29). In this case, a generic picture of a strawberry is decorative in nature because it does not add information about the economic learning content (but might serve a different function, such as catching the reader's eye). Representational pictures directly depict what is described in the text (e.g., a drawing of a heart next to text about the physiology of the heart). Organizational pictures provide the reader with coherence by thematically organizing material; a typical example might be a timeline that organizes the historical events described in a text. Interpretational illustrations add new information to the text and serve as clarifiers for difficult material (e.g., an illustration of the activity at an axon terminal often serves as a clarifier to explain the signal transmission between neurons via synapses). In a transformational illustration, information is recoded to promote a reader's understanding and recall ability, Carney and Levin (2002) used a mnemonic image as an example, where information about a town called Belleview in a text was accompanied by an interactive image involving a bell (thus promoting recall of relevant information through the recoded picture). In summary, it is not the graphical representation alone, but the graphical representation and the text context that decide how a graphical representation is classified. A picture of a strawberry might serve a decorative function (e.g., when it is displayed next to text about the strawberry market), but it might also serve a representational function in a text describing different fruits.

In contrast, most frameworks classify according to the structural aspects of an external representation. For example, Schnotz (2001) works with a form classification where external representations are distinguished according to the similarity between the represented object and the representation. A descriptional representation is used to convey information through symbols, such as words or formulas, where there is no similarity between the represented object and the representation. The word "dog," for example, has no similarity to the represented object (it does not look like a dog), and unless one knows the meaning of the word, it is not possible to infer the object from the representation. In contrast, a depictional representation is used to visualize information to describe a real-world object with the help of a visual-graphical representation. For a representation to be depictional, some similarity must exist between the representation and the object, which can either be realistic or logical. The image of a dog would be a realistic depictional representation because it corresponds structurally to the object in the real world (it looks like a dog), whereas a bar chart representing the number of dogs and cats in a certain area is a logical representation since the heights of the different bars are equivalent to the quantity they describe. Regarding logical pictures such as graphs, it is important to note that the connection between logical pictures and the represented information is "*neither arbitrary, as is the relation between words and concepts, nor a first-order isomorphism, as is the relation between pictures and their referents*" (Shah, Freedman, & Vekiri, 2005, p. 429).

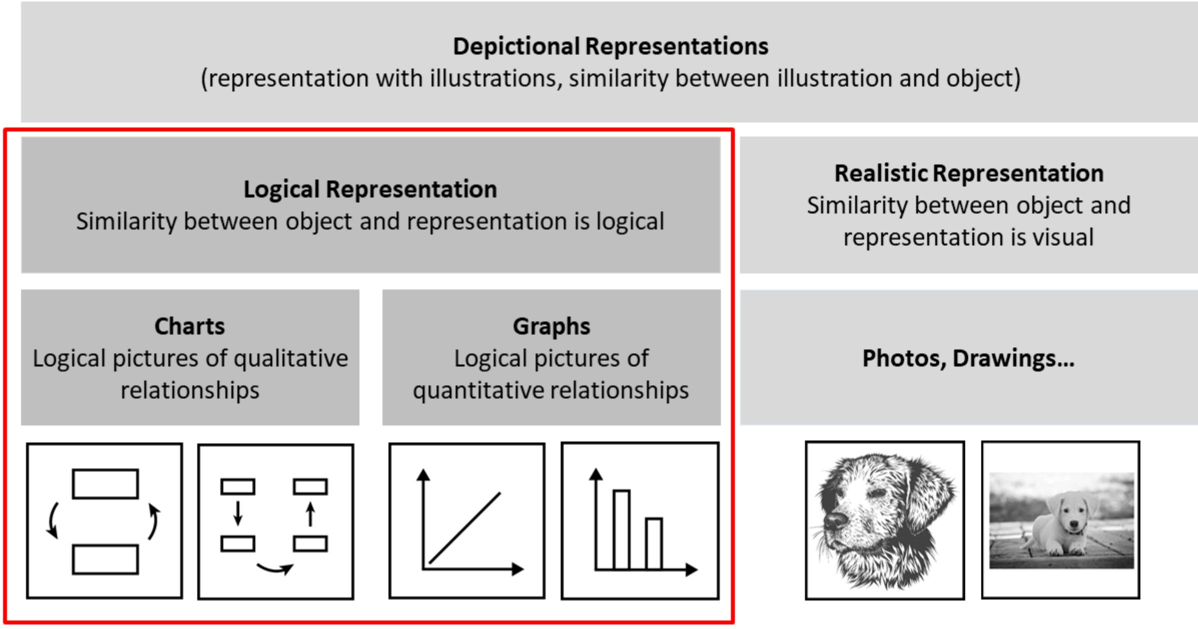


Figure 4. Classification of depictional representation

Own illustration based on the classification by Schnotz, 2001; pictures from pixabay were used as realistic representations: User Gorkhs (2020) [Vector drawing of a dog], pixabay, https://pixabay.com/de/vectors/hund-tier-haustier-h%C3%BCndchen-1728494/, User 3194556 [Picture of a dog], pixabay, https://pixabay.com/de/photos/hund-haustier-tier-niedlich-wei%C3%9F-1903313/

Within logical pictures, Schnotz (2001) further distinguishes between two forms: (1) charts that visualize qualitative relationships and do not have clear reading rules and (2) graphs that represent quantitative relationships and have clear reading rules (Figure 4). According to this terminology the flow of income model (Figure 1) would therefore be categorized as a chart, whereas the relationship between price, supply and demand (Marshall graph, Figure 1) is a graph.

Not only are the classifications useful to keep different visual representations apart, but they also allow for a differentiated analysis of their potential effectiveness in supporting the learning process. Levie and Lentz (1982) found that function is a relevant category in that regard: Adding decorative pictures to learning material does not benefit learning outcomes, whereas for representational pictures, this is generally the case (see Subsection 1.4.3). Furthermore, for form, Schnotz and Kürschner (2008) demonstrated that the structure of an external representation is connected to the internal mental model-building, and different representational formats thus might lead to different conceptions of domain principles (see Subsection 1.2.3.2)

Overall, multiple approaches can be employed to classify visual representations in teaching and learning. The different classifications are characterized according to function (e.g., Levin et al., 1987) or form (e.g., Schnotz, 2001) and differ not only in their scope but also in their terminology. In this dissertation, the terminology of Schnotz (2001) and the terms depictional, graphical and visual representations are thus used as umbrella terms for all representations with illustrations (realistic representations and logical representations). Moreover, as the focus of this dissertation is on logical representations, the classification is further used to distinguish between graphs and charts within logical pictures. This framework is mainly relevant in the first two sections, as the form category is more important when the ability to work with a single graphical representation is examined (e.g., when the effect of form on learning outcomes in economic education is discussed in Subsection 1.2.3.2). The function of a graphical representation and its relation to other representations become more relevant when learning environments with multiple representations are examined because the effect of graphical representations cannot be analyzed without considering their role in the context. Accordingly, the function is discussed in greater detail when the theoretical foundation for material consisting of text and pictures is presented (see Section 1.4). Since research on multimedia encompasses multiple representational forms (e.g., texts, graphs, charts and equations), the term “multiple representations” is used when referring to more than one representation (e.g., when the design of text–graph learning material is discussed in Section 1.4).

1.2.2 Graphical representations as visualizations of domain principles

Teaching and learning in (social) science are focused not only on facts but also on the understanding of complex and dynamic models and relationships that go beyond everyday concepts (P. Davies & Mangan, 2007). These relationships and phenomena are often invisible and cannot be experienced directly. Graphical representations, such as charts and graphs, can be used as visualizations of these abstract concepts and relationships, and they are thus used to communicate about models of the domain. In economics, this can be seen in research (Demir & Tollison, 2015) as well as education (Cohn & Cohn, 1994). The supply-and-demand graph, for example, is used as a visualization of the relationship between supply, demand and price. Not only does it allow readers to grasp the relationship quickly, but it is also used as a model to analyze how other factors influence supply and demand and thus influence quantity and price. For instance, poor weather in particular might lead to a decline in supply for seasonal goods, such as fruit and vegetables. As a consequence, fewer goods are traded at a higher price. In the graph, this is visualized by a parallel upwards shift of the supply line (S_1); the new price (P_1) and the new quantity (Q_1) are the coordinates of the intersection of the demand line and the new supply line (see Figure 5).

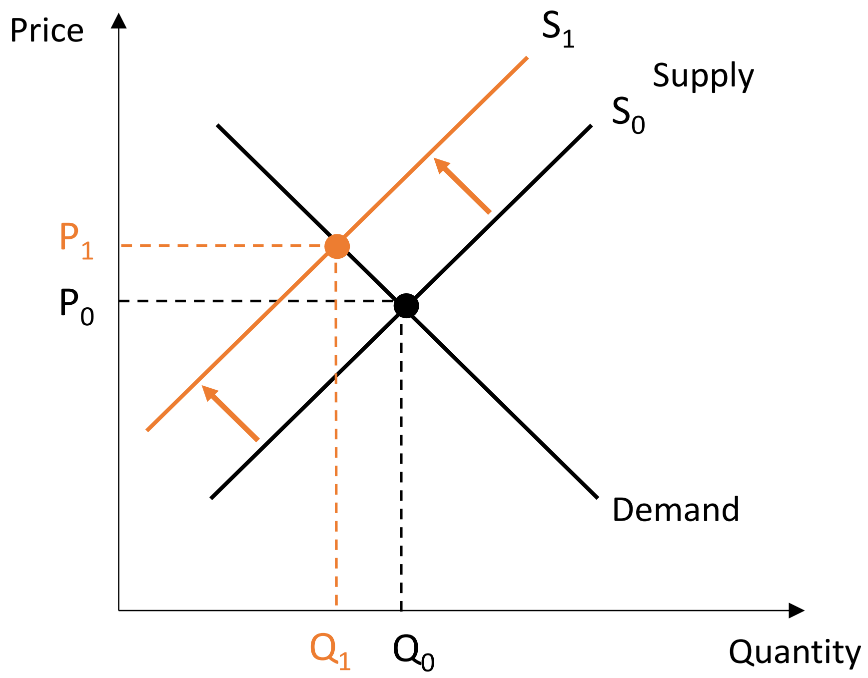


Figure 5. The effect of a decline in supply visualized in an equilibrium graph (own illustration)

In economics, graphs – especially equilibrium graphs – are such typical representations of the domain that according to P. Davies and Mangan (2007), the ability to understand the graphical representations defines an economist's way of thinking:

> *[…] economists use graphs to represent relationships between relationships (many lines on the same diagram), precisely in order to examine the operation of a system. Students face difficulties when they do not appreciate that this is the purpose of the graph, and do not realise that when they say 'I understand the concepts but not the graphs' they are revealing that they have not developed an economist's way of thinking and practising.* (pp. 720-721)

In a similar vein, Strober and Cook (1992) investigated the relationship between graphs and economic understanding by analyzing videotaped conversations of higher education economics students who discussed graphical solutions to typical economic labor market problems. The authors found that although students were confident in their understanding of economics, they often did not understand the graphical representations (e.g., in the conversations, they used the terms correctly but were not able to interpret and draw graphs accurately). Just as their colleagues Davis and Mangan, the authors also claim that a strong connection exists between understanding graphs and understanding economic principles:

> *"[…] students who say that they understand economics but do not understand graphs really do not understand the complexities of economic concepts and that their difficulties are*

probably more a result of problems with concept formation than of problems with spatial relation" (Strober & Cook, 1992, p. 126).

The ability to work with representations could be seen as part of domain knowledge not solely in economics: A study by Kozma and Russell (1997) found that the ability to fluently use different representations (such as graphs, animations and equations) to communicate about chemistry models separates experts from novices. Moreover, in biology, the ability to work with visual representations also correlates highly with domain knowledge (Nitz, Ainsworth, Nerdel, & Prechtl, 2014). In summary, the ability to work with representations is intertwined with domain knowledge – at least for "typical" representations of the domain. This is because learners are required not only to understand the form of representation (e.g., to be able to understand how the representation encodes and presents information), but also to know the relevant operators and therefore to understand the relationship between the domain and the representation (Ainsworth, 2006). For example, when working with a Marshall graph in economics, learners must learn how various factors influence supply and demand (and they should know that this influence is usually visualized as a parallel shift and not, for instance, a change in slope). When comparing different goods, however, the slope of the curve is indeed relevant as different goods react differently to changes in price; for example, essential goods might have a high demand regardless of price, whereas the demand for luxury goods might change drastically when the price is changed (Pindyck & Rubinfeld, 2016). Therefore, different operators are relevant, and learners must connect the graphical representation to the domain principle to understand what type of operator is used in which case.

In physics, when learners deal with line graphs depicting the distance and time of an object, they need to know that when they want to determine velocity, the relevant graphical feature is the gradient rather than the height of the line (McDermott, Rosenquist, & van Zee, 1987). In other words, to answer domain questions, learners must know the reading rules as well as the relevant operators when they work with graphical representations. In teaching, this results in what some scholars describe as "representational dilemma": teachers must enable learning about representations, such as how to work with a certain graph, as well as learning from representations, such as what the graph tell us for the domain question (Dreher & Kuntze, 2015; Rau, Aleven, & Rummel, 2017).

Westelinck et al. (2005) made the argument that since the use of visual representations differs in natural science compared to social science, the results of multimedia and representational competence research are not transferable between the two domain groups. In their study, the researchers tested the learning outcomes of 194 educational science students for material that consisted of either text only or text combined with graphical representation. They observed that studying text without external graphical representations sometimes resulted in higher performance. This is a somewhat surprising result, given the abundance of evidence in favor of the "multimedia effect" (e.g., Mayer, 2002, 2014b) – a term used to describe the finding that material consisting of text and graphical representations is more efficient for learning outcomes than text alone (see also Subsection 1.4.3). As an explanation for their finding, the authors argue that while in natural science, graphical representations are built on a consensual iconic sign system, and the reading rules of a visual representation are thus gradually mastered by learners of the domain, this is not the case for social science. Furthermore, they suggest that graphical representations in the field of the natural sciences are rather intuitive and depictive compared to social science. Therefore, learners profit from the additional graphical representation in science domains (where most studies regarding the multimedia effect were conducted) but might need more scaffolding to effectively use graphical representations in social science (Westelinck et al., 2005).

In summary, graphical representations are used in different domains, and because some representations visualize domain principles, their understanding is regarded as an aspect of domain expertise. Since different domains use different graphical representations (or the same representations differently), the relevant abilities that are needed to work with visual representations differ between domains. Therefore, when considering learning with representations in a certain domain, it is necessary to review how graphical representations are used by scholars and teachers in the respective domain.

1.2.3 Use of graphical representations in economic education

Analyzing the ability of students to work with (multiple) graphical representations is an established research tradition in science and math education as well as educational psychology (see, for instance, Åberg-Bengtsson & Ottosson, 2006; Bertin, 1983; Curcio, 1987; Friel et al., 2001; Kosslyn, 1999; Kozma & Russell, 1997; Lachmayer, Nerdel, & Prechtl, 2007; Leinhardt, Zaslavsky, & Stein, 1990; Mautone & Mayer, 2007; Nitz et al., 2014; Renkl & Scheiter, 2017; Stieff, Hegarty, & Deslongchamps, 2011; Winn, 1987). In comparison, in social science – especially in economic education – students' relevant abilities to work with graphical representations are not yet a major research interest, and the following two questions are mainly discussed:[2] (1) Under which conditions is adding logical pictures to verbal lectures and learning material beneficial for learning outcomes (e.g., Cohn et al., 2001; Cohn & Cohn, 1994; Kourilsky & Wittrock, 1987; Wheat, 2007), and (2) which graphical representations are used (and should be used) in economic teaching (e.g., Boatman, Courtney, & Lee, 2008; Chiou, 2009; M. Davies, 2011; Jägerskog, 2020; Marangos & Alley, 2007; Reingewertz, 2013; Wheat, 2007). Studies that discuss multiple visualizations, text–graph integration or translation between representations are rare (e.g., Stern, Aprea, & Ebner, 2003). Furthermore, graphical representations play a (minor) role in other discourses, such as in the discussion concerning effective explanations in economic education (e.g., Findeisen, 2017; Schopf, Raso, & Kahr, 2019). The majority of studies that discuss graphical representations in economic education focus on higher education; to date, little attention has been paid to secondary education (the exceptions are, e.g., Jägerskog, 2020; Kourilsky & Wittrock, 1987; Raso, 2018).

1.2.3.1 The effect of (additional) graphs in economic lectures

The effectiveness of learning with graphs as one form of graphical representation has been analyzed primarily regarding economic learning outcomes in micro- and macroeconomics. Kourilsky and Wittrock (1987) demonstrated that, compared to solely verbal presentation, high school students ($N = 83$) benefited from the combination of verbal and graphical representations only if the visual representations were presented *after* the verbal presentation. As a possible explanation, the authors argue that the familiar verbal mode allowed learners to focus on understanding the underlying principle, which could then be connected to the features of the unknown graphical representations in the second step.

In a study by Cohn and Cohn (1994), 78 higher education students were issued a handout and asked to take notes on a lecture about the equity and efficiency of a proportional wage tax. One-half of the students received a handout with the graphs used in the lecture, whereas the other half received the handout without the additional graphs and thus drew the graphs as part of their note-taking. Analyzing the notes, the authors observed that most students were unable to accurately reproduce the graphs from the lecture. Within the drawing group, students who scored higher regarding graph accuracy performed better than their counterparts in a posttest. Furthermore, the group with instructor-supplied graphs outperformed students with low

[2] These questions are closely related to research in graph comprehension and multiple representations and are therefore discussed from different viewpoints throughout this dissertation. Here the focus lies primarily on learning outcome in economics. Conceptual models that generally discuss the effect and interplay of learner, representational design and function of a representation can be found in 1.3.4 (for graph competence) and 1.4.3 (for text and picture learning environments).

accuracy scores (whereas students who accurately drew graphs scored higher than the group who did not need to draw graphs themselves).

In addition, in a study by Cohn et al. (2001), 208 university students were randomly assigned to a lecture with graphs or a lecture without graphs. Students who were presented graphs together with the verbal lecture had either the same or significantly *lower* scores than students in the lecture without graphs. The authors state that one reason why the graphs were not helpful (or even counterproductive) could be that the used graphs are complex, and the students did not spend enough time to fully understand the visualized principles.

In summary, findings regarding the effects of adding graphs to verbal instruction on learning economics are inconclusive. Based on the results thus far, it is reasonable to assume that for graphs to be beneficial, they must be understood by the learners. This, however, seems to be difficult for many economic students without additional support.

1.2.3.2 The effect of different graphical formats on learning outcomes

Partly based on arguments that graphs are not effective, different authors have argued that instead of or in addition to using graphs, other visual representations might be less challenging for learners and more effective for teaching. Especially charts, such as concept maps, flow charts and causal loop diagrams, have been suggested (Chiou, 2009; M. Davies, 2011; Jägerskog, 2020; Marangos & Alley, 2007; Reingewertz, 2013; Wheat, 2007). Those suggestions are sometimes based on the argument regarding whether students *like* to work with certain types of visual representations (e.g., Boatman et al., 2008; Marangos & Alley, 2007), whereas quasi-experimental studies that manipulate the form of the graphical representations and measure economic learning outcomes are still rare. Wheat (2007), for example, investigated the use of feedback loops, where the relationships between different variables were visualized as loop diagrams in interactive computer simulation models, rather than the typical supply-and-demand graph. One of his central theoretical arguments in favor of feedback loops (see Figure 6) is that they are better at visualizing dynamic processes (e.g., the shift of equilibrium in economic market models) in comparison to their static graph counterparts.

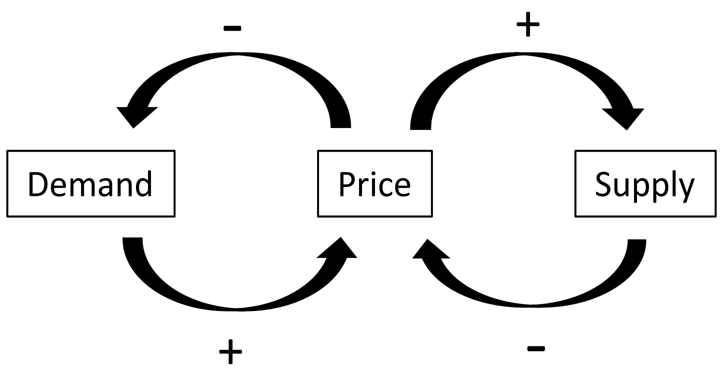


Figure 6. Feedback loop visualization of the relationship between demand, supply and price (own illustration)

In multiple experiments, the author demonstrated that students not only preferred the feedback loop compared to the graph, but also achieved higher posttest scores when studying loops instead of graphs. In a similar vein, Jägerskog (2020) studied the influence of a graphical format on students' understanding of price and confirmed Wheat's findings for novice learners. Based on quantitative and qualitative data, she found that upper secondary students develop more qualified conceptions of causal relationships in pricing when they study with feedback loops compared to using graphs. Based on these results, she concluded that "*the supply/demand graph may be very constructive for communicating within the discipline and among experts, but not for communicating the foundations of causality in pricing to novice learners in the field*." (Jägerskog, 2020, p. 69).

Based on these findings, it can be concluded that the graphical format is a relevant category for learning with visual representations not only in terms of surface features but also regarding the understanding of domain principles such as the dynamic mechanisms of price-building in economics.

1.2.3.3 Graphical representations in economic learning material and teaching

Regarding graphical representations used in secondary education learning material, little research has been conducted so far in the domain of economics (in Germany): Raso (2018) analyzed the use of visualizations in secondary economic education and discussed their function as well as the relevant prerequisites for their effectiveness. According to Raso, visual representations can attract learners' attention, help teachers to structure an explanation and lastly support a learner with both comprehension and recall of specific domain knowledge. Different factors are relevant prerequisites for graphical representations to be effective as learning aids, namely, the design of the representation, the learners' prerequisites (such as their prior knowledge) and the use of representations in the classroom (Raso, 2018).

Regarding the design of the representation, Aprea and Bayer (2010) analyzed the instructional quality of graphs and diagrams in business and economics textbooks for secondary education students based on design principles derived from multiple representations and graph comprehension research. They concluded that logical pictures in textbooks use too many unnecessary details and do not support the reader in integrating those pictures with other representations (e.g., by signaling – see Subsection 1.4.5.1). Furthermore, Niederhaus (2011) analyzed linguistic characteristics as well as illustrations in vocational training textbooks for electrical engineering as well as cosmetics and body care. Her results support the argument that the use of graphical representations differ between domains: On the one hand, textbooks for the occupational field of cosmetics and body care contain significantly more pictures with a lower degree of abstraction, namely, photos and drawings, than textbooks for the occupational field of electrical engineering. On the other hand, electrical engineering textbooks contain significantly more pictures with a high degree of abstraction, such as schematics, graphs and tables. Finally, in a study by Findeisen (2017), pre-service teachers explained economic content and constructed graphical representations to accompany their explanations. While the author pointed out that many of the prospective teachers made graphical representations of their own accord (whereas a few only used visualizations when prompted), the graphical representations lacked accuracy and clarity. For example, when a pre-service teacher drew a flowchart to explain sales tax in accounting, the arrows were sometimes pointing in the wrong direction. All in all, the author concluded that accurately visualizing economic learning content while explaining is a challenge for pre-service teachers, which, in turn, reduces students' understanding of economic phenomena.

In conclusion, graphical representations are used in textbooks and as part of explanations in economic education. To date, mainly the format has been analyzed, whereas other influencing factors, such as their actual use in the classroom and students' abilities and challenges, are understudied.

1.2.4 Section summary

Different graphical representations can be found in learning material. The classification used in this dissertation is based on Schnotz (2001) and distinguishes (a) between realistic representations and logical representations and (b) within logical representations, between graphs, describing quantitative relationships, and charts, describing qualitative relationships. The classification is used because logical pictures – especially graphs – are relevant representations in economic education (P. Davies & Mangan, 2007), and the distinction between graphs and charts is relevant regarding learning outcomes (Jägerskog, 2020).

In domains, graphical representations are used to visualize not only data but also domain principles (e.g., the relationship between price, demand and supply in economics). A student's ability to work with certain representations is thus partly domain-specific, as it is intertwined with domain knowledge (Ainsworth, 2006). Researchers in the domain of economic education have mainly focused on the questions of whether adding logical representations to lectures is beneficial for learning outcomes (e.g., Cohn et al., 2001) and what logical representation should be used in teaching (e.g., Wheat, 2007). The findings regarding the effect of adding graphical representations to verbal explanations and lectures is inconclusive, as adding graphical representations can only have an effect if they are understood by the learners. However, this is dependent on how graphical representations are embedded in teaching (e.g., the degree to which they are connected with a verbal or textual explanation), which graphical representation is used (e.g., graphs vs. feedback loops) and what learners already know (e.g., regarding the content as well as the reading rules of a representation).

In secondary economic education, the research regarding graphical representation is scarce. Most researchers focus on the design of representations in learning material (e.g., Aprea & Bayer, 2010), whereas their use in teaching and the relevant abilities and challenges that learners face are not yet main research interests. As economic education research develops from a domain perspective, research is often restricted to the effectiveness of representations when learning economics (and thus, for example, the question of which format is the best for learning outcomes), without modeling graph competence as an individual ability that might be influenced by different interacting factors (answering the questions of which format is the best for which question and which learner).

The ability to work with different visualizations and hence influencing factors has been studied in particular by scholars of math and science education and (educational) psychology. Therefore, in the next section, different models and empirical results regarding the abilities and challenges learners face when working with graphical representations are summarized. It is reasonable to assume that some representation-specific abilities and challenges might be transferable to economic education, following the argument of the domain specificity of representational competence; however, not all results might be relevant for the economic domain.

1.3 Learning with visual representations

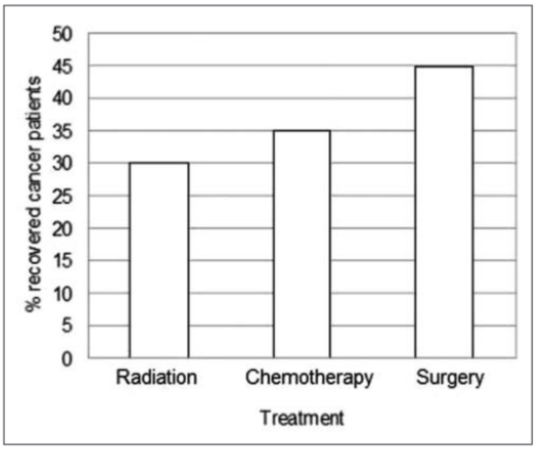
1.3.1 Terminology

One major challenge in the research field of learning with visual representations relates to connecting and integrating numerous research traditions that are concerned with the ability to work with graphs and other (multiple) representations. Varying terminology is used among different research communities. The most common terms are graph literacy or graphicacy (Åberg-Bengtsson & Ottosson, 2006; Galesic & Garcia-Retamero, 2011), graph comprehension or graphing ability (Berg & Smith, 1994; Curcio, 1987; Friel et al., 2001; McKenzie & Padilla, 1986), statistical or data literacy (Aoyama, 2007; Gal, 2002; Gould, 2017; Wallman, 1993; Watson & Callingham, 2003), and representational competence (Kozma & Russell, 2005; Nitz et al., 2014; Stieff, Scopelitis, Lira, & Desutter, 2016). With the exception of representational competence, all models discussed in this section focus on visual representations of quantitative relationships (i.e., graphs).

The terms are not used consistently across the different research fields; one of the central conceptual differences between the various models is what they expect of a graph-competent learner. Therefore, in the following subsections, the terms are defined to summarize the literature along a continuum from models that use specific task requirements (e.g., that ask learners to read single data points), to models that analyze a learner's ability to draw more general conclusions from graphs, and lastly, to models where the focus is on performing multiple operations (such as reading, construction and translation) and thus on the ability to represent domain principles.

In this dissertation, the term "graph literacy" or "graphicacy" is used for frameworks that focus on an individual's ability to read and interpret graphically presented information, such as line graphs, bar graphs and pie graphs (as well as maps). A typical item in a test to assess graphicacy shows a graph (that could be used in the real world) and asks participants to read off a point or to compare slopes. Although the representations are taken from a domain – for example, geography in Åberg-Bengtsson and Ottosson (2006), medical decision-making in Galesic and Garcia-Retamero (2011) or sustainable development, in Study 1 in Chapter 2 – the primary aim of these tests is not to assess participants' understanding of principles or models from the domain, but rather to evaluate their ability to extract information from graphs. The graphs therefore usually visualize a specific situation or data set rather than an abstract functional relationship. To read these data graphs, readers must have a general understanding of graph-reading rules as well as be able to do (simple) mathematical operations; intensive domain knowledge is not necessary. An example item can be seen in Figure 7, where participants are asked to read off a value of a bar chart as well as to compare two bars to each other (Galesic & Garcia-Retamero, 2011).

Here is some information about cancer therapies.



Q1. What percentage of patients recovered after chemotherapy?

_______ %

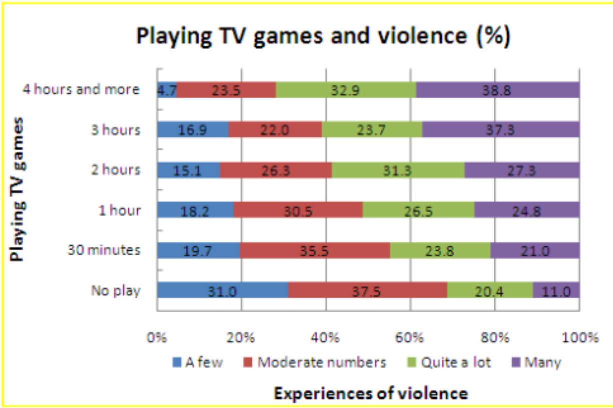
Q2. What is the difference between the percentage of patients who recovered after a surgery and the percentage of patients who recovered after radiation therapy?

_______ %

Figure 7. Item example of graph literacy (Galesic & Garcia-Retamero, 2011, p. 453)

In comparison, the term "statistical literacy" is used to describe research with a focus on a learner's ability to interpret, evaluate and communicate about statistical information (Aoyama, 2007; Gal, 2002). Statistical literacy is a broader concept, and the interpretation and evaluation of data-oriented graphs is part of statistical literacy, as graphs are typical representations of statistical data, relationships and concepts (Watson & Callingham, 2003). In addition to graph interpretation, statistical literacy encompasses the computation of relevant statistical-mathematical constructs such as median, average or correlation (Gal, 2002). Some authors include "data literacy" and therefore integrate knowledge and understanding of how data is collected, stored and used (Gould, 2017). In tests that are meant to measure statistical literacy, a small set of items might thus involve the interpretation and evaluation of graphs; for example, in a test by Aoyama (2007), an item included questions where learners not only needed to read the data but also needed to critically evaluate whether a statement could be concluded from the data (see Figure 8).

Item 3 (VLN). The following graph is the result of research that investigates how many hours per day elementary students play TV games at home and how many experiences of violence (e.g. hitting or pushing a classmate, or pulling someone's hair) they have.



3.1. What is the percentage of students experiencing "Quite a lot" of violence in the group playing a TV game for one hour?

3.2. As the hours of playing a TV game increase, is the proportion of people experimenting "a few" violence also increasing?

a:Increasing; b:Decreasing; c:Do not change; d:I don't know

3.3. As the hours of playing a TV game increase, is the proportion of people experimenting "many" violence also increasing?

a:Increasing; b:Decreasing; c:Do not change; d:I don't know

3.4. It is suggested that a long time spent playing TV games will cause violence. Do you agree with this opinion? Please choose and write your idea.

a:Agree; b:Disagree; c:Neither; d:I don't know

Figure 8. Item example of statistical literacy (Aoyama, 2007, p. 304)

The term "graph comprehension" is used in this dissertation to describe how learners understand graphs regarding the visualized relationship and the connection to domain principles. In comparison to graph literacy, items in these tests include items that focus on general relationships between different parameters that are depicted in (line) graphs as well as the connection between graph features and domain concepts (Lai et al., 2016). Furthermore, graph comprehension is often assessed as part of a larger graph competence, which encompasses other graph-related abilities, such as graph construction and graph evaluation (Hattikudur et al., 2012; Lachmayer, 2008; Lai et al., 2016). These graphs typically visualize abstract relationships rather than specific data (McKenzie & Padilla, 1986). While domain knowledge might not always be necessary, it is still essential to answer some of the questions, especially in tests that aim to connect graph features with domain concepts. One example item, where a general relationship must be interpreted, can be seen in Figure 9. It depicts different graphs and asks students to indicate which of them is described by two statements (McKenzie & Padilla, 1986).

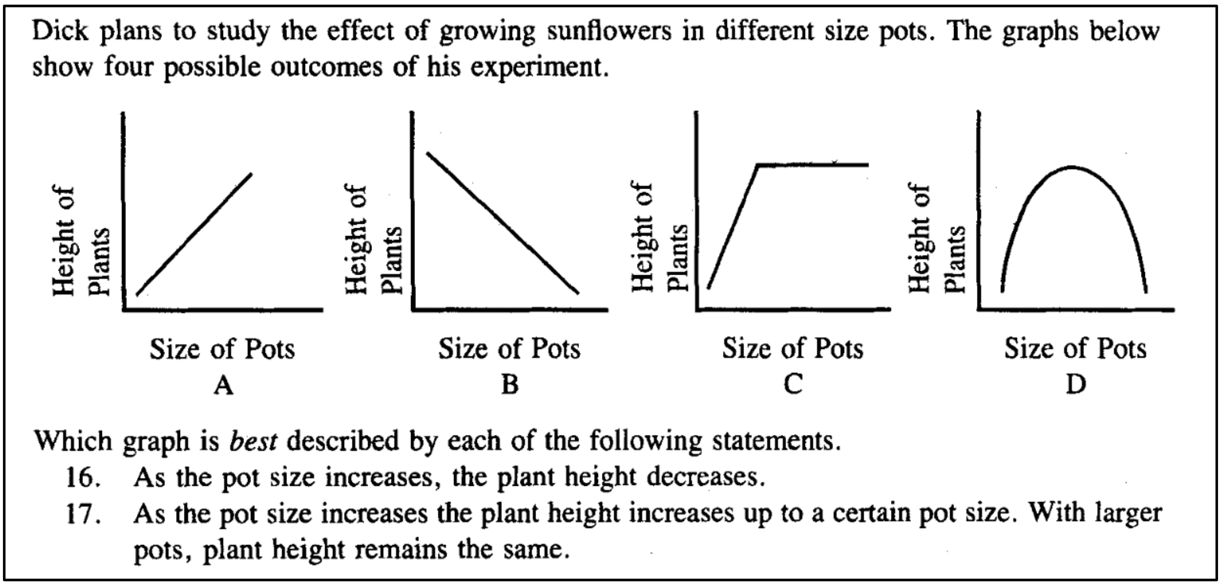


Figure 9. Item example of graph comprehension with a focus on interpreting general relationships with unspecific graphs (McKenzie & Padilla, 1986, p. 575)

A second example item can be seen in Figure 10. Here, the tasks are increasingly complex: For the first question, simply reading a data point is necessary, whereas the last question requires the explanation of a domain concept (Lai et al., 2016).

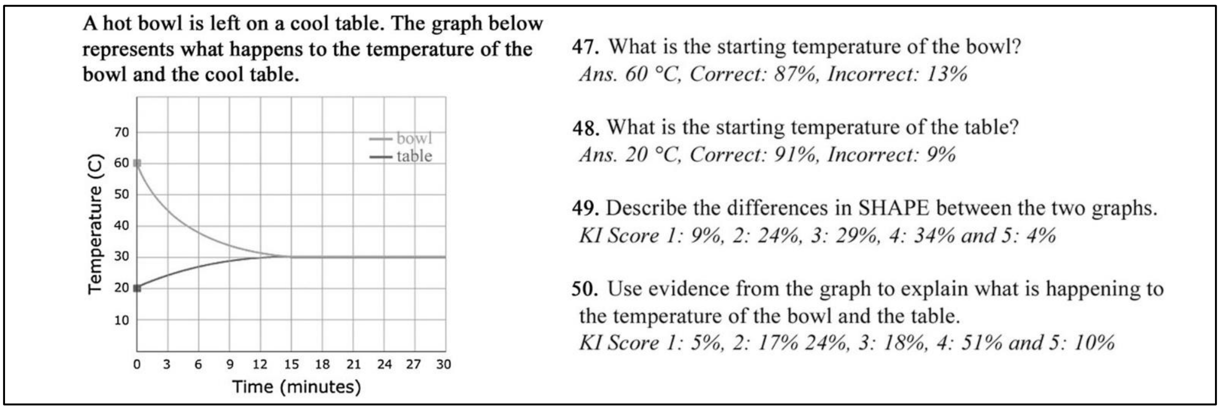


Figure 10. Item example of graph comprehension with a specific graph and domain concepts (Lai et al., 2016, p. 669)

Representational competence is used in this dissertation to describe the ability to work with multiple representations and the way in which learners use and connect (and possibly transform) information in different displays, such as graphs, tables and equations, to answer domain questions and to communicate about domain models (Kozma, 2003; Stieff et al., 2011). Researchers are interested in how the integration of multiple representations – and in turn learning outcomes – are influenced by learning material and instructional design (Bodemer et al., 2004; Scheiter & Eitel, 2015; Stieff et al., 2011; Ullrich et al., 2012). The practical implications of multiple representations and multimedia research are design principles for learning environments, for example the so-called “multimedia principle,” which states that learners tend to profit from studying text and pictures together compared to text alone (Mayer, 2014b). As different domains utilize various representations and might use the same representations differently, representational competence is seen as domain-specific (Ainsworth, 2006). Compared to novices, experts use different external representations fluently to explain domain problems (Kozma & Russell, 1997). Items that assess representational competence usually ask students to draw information from (multiple) representations or to either switch or translate between representations. An example is presented in Figure 11, where learners are asked to answer domain questions with the help of three chemical representations (Stieff et al., 2011).

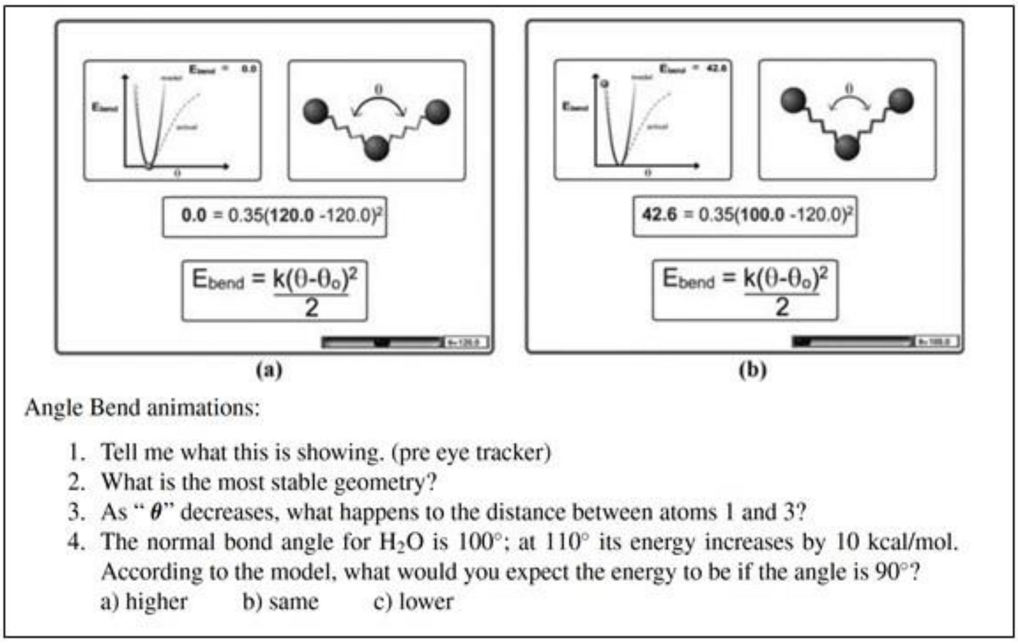


Figure 11. Item example of representational competence (Stieff et al., 2001, p. 125, 141)

To summarize, scholars have analyzed the ability of learners to work with visual representations in various research fields. The most relevant difference between models is what they expect of a graph-competent learner, ranging from specific tasks to more general interpretation (see the overview in Table 1). In the following subsections, the models are used to discuss the different tasks for a single visual representation (see the overview in Figure 12) as well as the typical challenges that learners face when performing the various operations. Lastly, numerous influencing factors for graph comprehension are discussed. As the integration of visual representations with other representations plays a larger role in multimedia and multiple representation research, frameworks for learning with multiple representations are presented in a separate section (see Section 1.4).

Table 1

Overview of exemplary measurement models for graph literacy, statistical literacy, graph comprehension and representational competence

Main Term / Main Question	Exemplary Authors	Task	Typical Visual Representation	Relevance of Domain	Extent of Abstraction
Graph literacy How well can learners read graphs?	Åberg-Bengtsson & Ottosson, 2006; Galesic & Garcia-Retamero, 2011	Reading	Graphs presenting specific data	Low	Low
Statistical literacy How well can learners draw conclusions from graphs?	Aoyama, 2007; Shaughnessy, 2007	Reading, Interpretation, Evaluation [focus on general relationships and conclusions from data]	Graphs presenting specific data		
Graph comprehension, Graph construction, Graph critique How well can learners use graphs to understand and represent a domain concept?	Lachmayer, 2008; Lai et al., 2016; McKenzie & Padilla, 1986	Reading, Interpretation [focus on general relationships and domain principles] Construction, Evaluation	Graphs presenting specific data, Graphs presenting general domain relationships, [sometimes text or data in a table to match to graphs or to construct graphs from]		
Representational competence How can learners use multiple representations to understand and represent a domain concept?	Kozma & Russell, 2005; Nitz et al., 2014; Stieff et al., 2016	Reading, Interpretation, Evaluation, Construction, Translation and Integration	Graphs, Equations, Models, Charts, … (depending on domain)	High	High

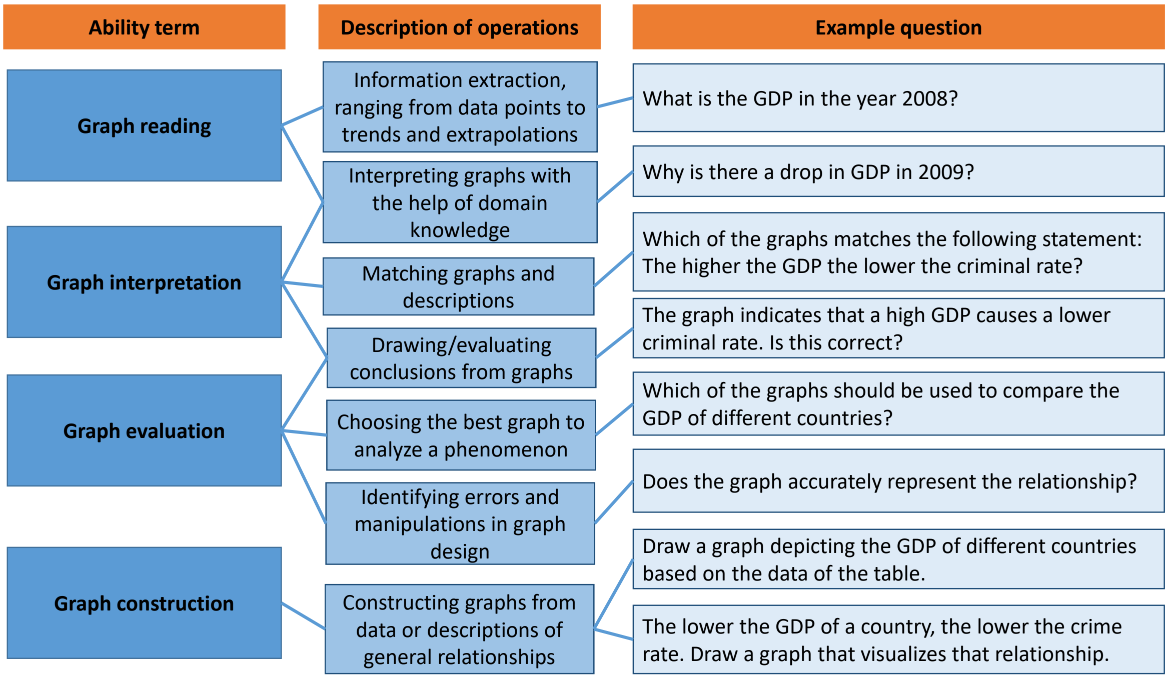


Figure 12. Overview of different graph tasks (own illustration)

1.3.2 Reading, evaluation and interpretation of graphs

A graph can efficiently display a variety of information, and a reader must make sense of that information. Therefore, models dealing with one's ability to work with visual representations encompass reading or interpreting graphically presented information as a relevant skill for a graph-competent learner (Bertin, 1983; Curcio, 1987; Lachmayer, 2008; Lai et al., 2016; Leinhardt et al., 1990).

Most models categorize the types of questions that a graph can be used to answer with a three-level approach (Curcio, 1987; Friel et al., 2001; Lai et al., 2016). The first level, "read the data" (Curcio, 1987), refers to the elementary question level, where the reader is asked to extract basic information about single data points (e.g., "What is the number of fish stocks in year X?"). To answer a question at the intermediate level, namely, "read between the data" (Curcio, 1987), learners must interpret and integrate information presented in a graph to find trends or relationships in the data (an example question might be, "How do fish stocks develop over time?"). For this second level, Lachmayer (2008) makes a more precise division and distinguishes between comparing two values (and identifying a trend) and comparing multiple values (or comparing multiple trends). The expert level, "read beyond the data" (Curcio, 1987), requires extrapolation; learners must draw general conclusions from the graph to form predictions or to answer domain questions. In some models, the "third" level questions can be answered without background knowledge, but readers are expected to extrapolate from the data (Curcio, 1987); an example question could be, "Based on the graph, how would you expect that fish stocks develop in the future?”. In contrast, other models expect learners at this level to be able to connect the graph’s information to domain concepts (Lai et al., 2016); for example, the following could be a question regarding fishery: “how would you expect the graph to develop if the principle of maximum sustainable yield would be used for fishing?”

Therefore, especially the third level is operationalized differently between studies: In some studies, prior knowledge is necessary to answer these questions (Lai et al., 2016), whereas in others, the tasks can be solved with a graph alone (Curcio, 1987). This methodological disparity might be the explanation for the contradictory findings where, in some studies, the three levels are consecutive – that is, Level-1 items are easier to solve than Level-2 items (Lai et al., 2016) – whereas in other studies, item difficulties and item levels are only partly related (Lachmayer, 2008). Researchers from statistical literacy in particular argue that it is more practical to distinguish between questions that can be answered with a graph alone (the first three examples) and questions where prior knowledge, critical thinking skills and evaluation are needed (the second Level-3 example; e.g., Gal, 1998). The line between reading, interpretation and evaluation is thus blurry. In some models, questions that require content or statistical knowledge are part of reading or interpretation, whereas other models view them as part of evaluation or critique. In general, the difficulty of items ranges from low difficulty for “read the data” questions to high difficulty for questions where learners need to extrapolate or more generally interpret the data (Curcio, 1987; Friel et al., 2001; Lachmayer, 2008). However, the questions of whether a precise level-wise distinction exists (e.g., different difficulties for single data points, trends and extrapolations) or whether it is more reasonable to distinguish between reading data questions and questions where data is interpreted (e.g., general conclusions) remain open.

In addition to information extraction (“quantitative interpretation of graphs”), some models focus on matching descriptions of relationships with graphs, sometimes termed “qualitative interpretation of graphs” (Leinhardt et al., 1990, p. 11). Instead of presenting learners a single data graph and asking them to identify a certain y-value at a certain x-value (which would be a Level-1 question in Curcio’s framework), items display multiple line graphs depicting different general relationships between variables and ask learners which of those graphs matches a certain description; for example, “Which graph is best described by the following statement: as the plot size increases, the plant height decreases” (McKenzie & Padilla, 1986, p. 575, see also Figure 9). To perform this qualitative interpretation of graphs, learners must consider the entire (line) graph and analyze the relationship between two variables focusing mainly on the pattern of covariation (Leinhardt et al., 1990).

Since evaluation could be understood as part of interpretation, a significant overlap exists between frameworks that describe interpretation and those that describe evaluation. Graph evaluation or critique is thus sometimes addressed in graph competence tests (e.g., Lai et al., 2016) and plays a key role in critical thinking and statistical literacy research (Aoyama, 2007; Gal, 2002; Nolan & Perrett, 2016; Watson & Callingham, 2003).

Although not focused on graphs per se, the interpretation and evaluation of graphical representation of statistical data is one aspect of statistical literacy (Aoyama, 2007; Watson & Callingham, 2003). In statistical literacy tests, students are asked to critique graphs according to their structure and design (e.g., identify missing axis titles) and according to their accurate representation of data (e.g., finding mistakes in graphs, use of accurate scaling; Watson & Callingham, 2003). Regarding the "understanding of meaning of the information," scholars have also analyzed whether students can draw the right conclusions from graphs and whether they can identify false conclusions, such as inferring causality from correlation (Aoyama, 2007). A high level of statistical literacy also includes the ability to critically analyze a graph's statistical design, for example regarding sampling or statistical power (Watson & Callingham, 2003).

In science, learners are often required to choose which graph might be the best fit to analyze a certain scientific phenomenon. For instance, in an instrument measuring graph comprehension, construction and critique, which was tested with approximately 460 middle school students, Lai et al. (2016) included a critique item where students were presented with two graphs depicting climate change data and asked to indicate which of the graphs should be used to analyze the phenomena.

1.3.3 Constructing graphs

When learning, students are expected not only to interpret and evaluate graphs but also to construct graphical representations. Although the abilities necessary to interpret and construct graphical representations are highly correlated (Lachmayer, 2008), they are seen as distinctive abilities: "*Whereas interpretation relies on and requires reaction to a given piece of data (e.g., a graph, an equation or a data set), construction requires generating new parts that are not given*." (Leinhardt et al., 1990, p. 12). The construction of graphs involves the generation of a quantitative visualization based on data, descriptions, tables or equations. For line graphs, this involves the construction of axes, scale as well as the identification of unit and data entries (Kotzebue et al., 2015). In one step prior, some authors include the choice of the appropriate graph or diagram type for different data sets or relationships as a relevant skill for constructing logical pictures (Baker, Corbett, & Koedinger, 2001; Lachmayer, 2008). In her competence model, Lachmayer (2008) describes the first part of construction as building the graph frame; prior to data entry, this process involves choosing the correct graph type, assigning the dependent and independent variables to the axes, labeling the axes as well as selecting and constructing scales. In her framework, the second part of construction includes three complexity levels: a single data point entry, construction of a trend line and construction of multiple trend lines. In comparison, Lai et al. (2016) use a digital graphing tool in their instrument, which provides students with a predefined set of axes and a point-plotting interface and allows for automatic evaluation of the resulting graph (Vitale, Lai, & Linn, 2015); therefore, their framework does not involve tasks before data entry (such as constructing and labeling axes).

1.3.4 Influencing factors and challenges in graph competence

Different factors influence the understanding and construction of graphs, and they are connected to typical challenges (a typical difficulty might be the effect of a certain format). In this subsection, the influencing factors and typical challenges are presented together. Although there are overlaps, the influence of task characteristics, design characteristics and learner characteristics are distinguished.

1.3.4.1 Task characteristics

For graph reading and interpretation, learners in secondary school face difficulties with general conclusions and open questions, whereas items that ask for specific information are mostly answered correctly for simple data graphs (Åberg-Bengtsson & Ottosson, 2006; Berg & Smith, 1994; Curcio, 1987; Friel et al., 2001; Gal, 1998). For example, Tairab and Khalaf Al-Naqbi (2004) analyzed the ability of biology students of in the 10^{th} grade (around 15–16 years) to read or interpret as well as construct biology graphs from given data tables. The students were mostly able to answer interpretation-based questions; however, they had more difficulties with questions that asked for a general interpretation of whole graphs (e.g., "*What general conclusion can be made from the graph above?*", p.132), compared to items that asked for specific data points (e.g., "*Which of the fruits shown on the graph contains more vitamin C?*", p.132). Similarly, the TIMS study 2011 demonstrated that internationally, around 60% of all 15-year-olds could read a single value of a graph that visualized the number of cars produced at a certain time, whereas learners struggled when they were asked to compute the average number of cars produced or identify the time interval in which the most cars were produced (29% and 34% solving rate respectively, TIMSS, 2013).

In their literature reviews, Glazer (2011) and Leinhardt et al. (1990) identify four major misconceptions that arise when learners interpret and construct general relationships depicted in line graphs. One challenging aspect of graph interpretation is the distinction between height and slope; for example, in a graph that depicted the position and the time (position on the x-axis and time on the y-axis) of two objects, learners falsely used the height (instead of the slope) to answer questions regarding speed (Hadjidemetriou & Williams, 2002; McDermott et al., 1987). Second, to answer comparison questions, learners tend to focus on points instead of intervals; for example, when they are shown a graph that depicts the height and age of girls and boys as two separate lines and are asked, "when are girls growing faster than boys?", they are more likely to answer with a specific age rather than an interval (Leinhardt et al., 1990; Preece, 1984). Furthermore, especially young learners are not always able to see a graph as an abstract relationship of variables and instead perceive it as a literal picture; for example, a graph that visualizes the relationship between time (x-axis) and distance (y-axis) is seen as a visualization of someone climbing a mountain instead of an abstract visualization of movement with changing direction (Kerslake, 1981; Lai et al., 2016; Leinhardt et al., 1990). Lastly, difficulties arise in understanding the nature of a graph as a general relationship, instead of separate points; this misconception is evident when learners prioritize a line connecting point-to-point rather than a best-fit trend line to describe a relationship between variables (Kerslake, 1981; McKenzie & Padilla, 1986).

Another challenge for learners is the interpretation of specific statistical graphs such as boxplots or histograms (e.g., Boels, Bakker, van Dooren, & Drijvers, 2019; Lem et al., 2015; Lem, Onghena, Verschaffel, & van Dooren, 2013). The roots of this challenge are misunderstandings about the statistical concepts that are visualized. For example, in a review by Boels et al. (2019), the authors identified misconceptions about histograms, which they linked back to key concepts of statistics: Regarding the key concept data, learners have difficulties with identifying the measured variable (e.g., they see histograms as a display of two variables) and understanding the measurement levels (e.g., they make reference to a normal distribution for nominal variables). Moreover, regarding the concept of distribution, the authors found misconceptions about variability (e.g., leaners interpret the difference in the height of the bars as a difference in variance), center (e.g., they use the mean of the frequency instead of the mean of the variable), information reduction through grouping (e.g., they have misinterpretations regarding the grouping in bins) and shape (e.g., students are unable to link histograms to corresponding boxplots).

Furthermore, learners have difficulties with graph selection (i.e., when they are asked to select the proper type of graph for the given data or relationship). In a study by Baker et al. (2001), under 25% of 52 ninth-grade students were able to select the proper graph type when they were shown four alternatives. Furthermore, when high school students are given a general relationship and are asked to choose the line graph setup (the axis combination) that is suitable to analyze the relationship, only around 50% of the students are able to do so (Padilla, McKenzie, & Shaw, 1986).

Learners in secondary education are often not able to critically analyze the statements that can be drawn from a data graph. For example, in a study by Aoyama (2007), students were asked to evaluate (false) conclusions from graphs (for item example, see Figure 8); that is, students were presented with a graph visualizing correlated data between two variables (hours spent playing video games and experiences of violence) and asked whether they agree with the conclusion that one variable (playing video games) causes the other (experiences of violence). Most students in secondary education (12–19 years) agreed with the statement, and only a small portion of students could separate correlation from causation, for example by reasoning that there might be another factor that is the cause for the two variables. Connected to difficulties regarding the evaluation of conclusions that can(not) be drawn from graphs, learners also struggle with choosing the best graph to analyze a certain phenomenon. In a study by Lai et al. (2016) sixth to eighth graders were tasked with choosing a graph that best visualizes climate change. The students regularly picked the graph with fewer data points and a smaller time frame, rationalizing their choice with readability instead of scientific arguments or prior knowledge. Lastly, learners (grades 3–9) have difficulties when they are asked to evaluate the graph features for a graph that is missing typical graph elements such as axis labels (Watson & Callingham, 2003).

Kotzebue et al. (2015) analyzed 437 first-year biology students' ability to construct graphical representations. They found that before actual data entry, learners had trouble with the construction of the data frame (e.g., selecting the correct graph type, assigning the variables to the axis, labelling the axes, drawing the scale and creating the legend). A typical mistake among the university students, for example, was the assignment of the dependent variable to the y-axis and the independent variable to the x-axis. Moreover, when entering data, the students were not always able to plot the data points and, when connecting points, often drew the line further than the data (Kotzebue et al., 2015). Similar findings, namely, that learners struggle with graph type selection and scaling and tend to extend their line graphs beyond the given data, were revealed in a study that was recently carried out with 46 eighth-grade students (Ozmen, Guven, & Kurak, 2020). Furthermore, for line graphs, drawing the right y-intersect as well as the slope of a line, when a general relationship is given, is difficult for sixth to eighth graders; for the y-intercept, the type of graph is relevant: sixth-grade students had more difficulties with the y-intercept for more generalized line graphs, where no specific numbers were displayed (termed qualitative graphs by the authors) compared to specific data graphs (termed quantitative graphs). The authors argued that graphs with quantitative features made it easier for the students to focus on local aspects and specific data points and thus helped them to construct the y-intercept (Hattikudur et al., 2012).

Lastly, mistakes in the construction of graphical representations are also the result of misconceptions of domain concepts. Examples of such misconceptions are as follows: when learners depict the process of cooling (and thus the relationship between temperature and time) as linear, even though it is exponential in physics (Lai et al., 2016; Vitale et al., 2015), or when learners are not able to construct accurate (equilibrium) graphs in economics (P. Davies & Mangan, 2007; Strober & Cook, 1992; see Subsection 1.2.3).

1.3.4.2 Representation design characteristics

The graph design influences both fact retrieval speed and conceptual understanding on different levels. Surface design characteristics, such as color, legend vs. labels and layout (detailed vs. simple) are mostly relevant regarding low-level perceptual aspects of graph comprehension; thus, they influence fact retrieval speed, but not necessarily conceptual understanding (Shah et al., 2005). The format of graphs, however, can make a meaningful difference regarding interpretation: Discrete comparison between values is easier when analyzing bar graphs, whereas line graphs are better for visualizing trends (Shah & Hoeffner, 2002); furthermore, learners are more likely to identify interactions in line graphs instead of bar graphs (Kosslyn, 2006; Shah & Freedman, 2011). Therefore, “*the effectiveness of a particular graphical presentation depends on the degree that the display contains the relevant information for the task demands*” (Shah et al., 2005, p. 451), and the presence of task-irrelevant information can impair performance (e.g., Canham & Hegarty, 2010). Moreover, the same data in different graph formats might lead to different conclusions; for example, in a study by Elting, Martin, Cantor and Rubenstein (1999), the authors demonstrated that decisions about a hypothetical clinical trial are influenced by the format of visualization. In studies that analyze graphical literacy or graph comprehension in the educational context, however, the format of the graphical display does not always influence graph comprehension performance (Åberg-Bengtsson & Ottosson, 2006; Lachmayer, 2008). A possible reason for this is that those studies usually “control” for graph–task fit by asking “fitting” questions for the different graphs.

In addition, in math education, some scholars examined graph specificity as an influencing factor: Hattikudur et al. (2012) analyzed graph construction for graphs with quantitative features (i.e., those with specific values and axis scaling) and graphs with qualitative features (i.e., those with axis labels but without specific values for the variables). The authors argue that graphs with only qualitative features might support students in focusing on the “bigger picture,” whereas specific graphs with quantitative features might be better for identifying and constructing specific quantities (Hattikudur et al., 2012; Lai et al., 2016).

1.3.4.3 Learner characteristics

Different learner characteristics have been shown to influence graph comprehension. General cognitive abilities, such as logical thinking and abstract reasoning, are positively associated with graph reading and interpretation (Berg & Smith, 1994; McKenzie & Padilla, 1986). Related to general cognitive abilities, academic language achievement (Åberg-Bengtsson & Ottosson, 2006) as well as mathematics knowledge and experience are positively correlated with graph competence (Åberg-Bengtsson, 1999; Curcio, 1987; Ludewig, Lambert, Dackermann, Scheiter, & Möller, 2019). Ludewig et al. (2019) demonstrated, for example, that beyond general cognitive ability, mathematical abilities such as performance in number line estimation and subtraction, as well as conceptual knowledge about arithmetic, were significant predictors of graph-reading performance. Although the finding is not surprising – as graph-reading tasks partly involve numerical operators (e.g., when learners compute the difference between bars in a bar graph, they usually need subtraction) – it is still relevant because it indicates that the foundation for graph reading is laid in primary school. Furthermore, it supports another argument made by Roth and McGinn (1997). They argue that instead of explaining a lack of graph competence in terms of cognitive ability, it is more useful to view it as a lack of meaningful practice and experience. In other words, although graph comprehension is related to cognitive abilities, it can (and should) be learned, which can be done through experience and practice. This effect, namely, that experience is relevant for graph reading, can be seen in different studies (overview in Shah et al., 2005), for example in a study done by Shah and Shellhammer (1999). They demonstrated that highly skilled graph viewers are more likely to identify general trends compared to low-skilled graph viewers. In addition, in a study by Peebles and Ali (2015), bar and line graphs that visualized a 2x2 interaction between variables from different (non-psychology-related) sources were shown to expert graph readers (N = 42 researchers of cognitive psychology or cognitive science). Whereas previous research has found that learners are more likely to identify interactions in line graphs compared to bar graphs (e.g., Kosslyn, 2006), their study revealed that this is not the case for expert readers, as their performance was the same for both graph formats. Graph-reading expertise thus reduces the effect of “design” effects such as format (Peebles & Ali, 2015).

Lastly, the content knowledge of a graph viewer is relevant in different ways. Even if a specific graph is unknown, the relationships and the relevant terminology used in graphs visualizing domain-specific data should be familiar to a learner with more content knowledge. It is reasonable to assume, for example, that experts in the economic domain know that the GDP is one of the central measures for economic growth and that they further know how the measure has developed over the last years.

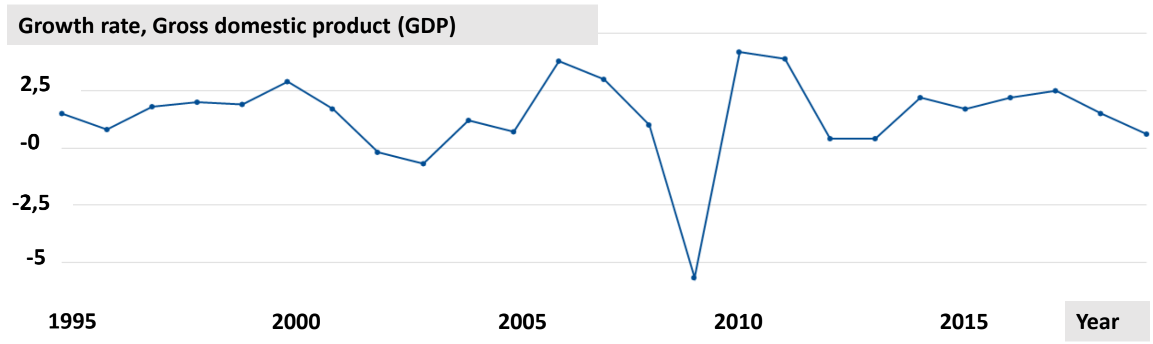


Figure 13. Growth rate of gross domestic product (annual in %) in Germany in the last 25 years

Own illustration based on open data from: The World Bank (2020) [Data Bank, World Development Indicators], https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.KD.ZG&country=DEU#advancedDownloadOptions, parameters used: GDP Growth annual in % (Indicator), Germany (Country), 1995 – 2020 (Time)

When encountering a data graph that depicts GDP growth rate over time (see Figure 13), this knowledge might be helpful for the interpretation. A novice learner, who has not yet encountered the term and does not know that it is a measure of economic growth (and has thus likely never thought about the growth rate of GDP over time), might have more trouble when interpreting the graph. An expert could connect the data to economic events and, for example, interpret the drop in GDP growth rate in 2009 as a result of an economic recession (the financial crisis in 2007–2008), whereas a novice learner might identify a sudden drop but not necessarily be able to connect the graph feature to an economic event. This effect, namely, that familiarity with the data and the underlying concepts plays a role in graph comprehension, was demonstrated in a study by Shah and Freedman (2011). The authors analyzed the graph comprehension of 55 psychology undergraduate students for graphs that depicted either familiar or unfamiliar data. Students achieved higher scores in graph comprehension (i.e., they generated more interferences and identified more interactions) when they viewed familiar graphs.

When the terms and the relationship between variables are not known, the domain principle and the graphical representation of the principle might still be similar enough that expertise might be relevant. For instance, a student in economics who has already encountered the Marshall graph should already know that the intersection of the two lines (demand and supply) is relevant for analysis and is further aware that the factors influencing demand or supply can be visualized as a shift in the respective line (see Figure 14, left panel). The student could then use the knowledge of the operators to work with a similar graph, for example the IS–LM model, which visualizes the relationship between the goods market (IS curve) and the money market (LM curve). In the same manner as in the Marshall diagram, influencing factors such as an increase in money supply are modeled as a shift of the LM curve, which results in a new interest rate i_1 and economic output Y_1 (Feenstra & Taylor, 2014, see Figure 14, right panel).

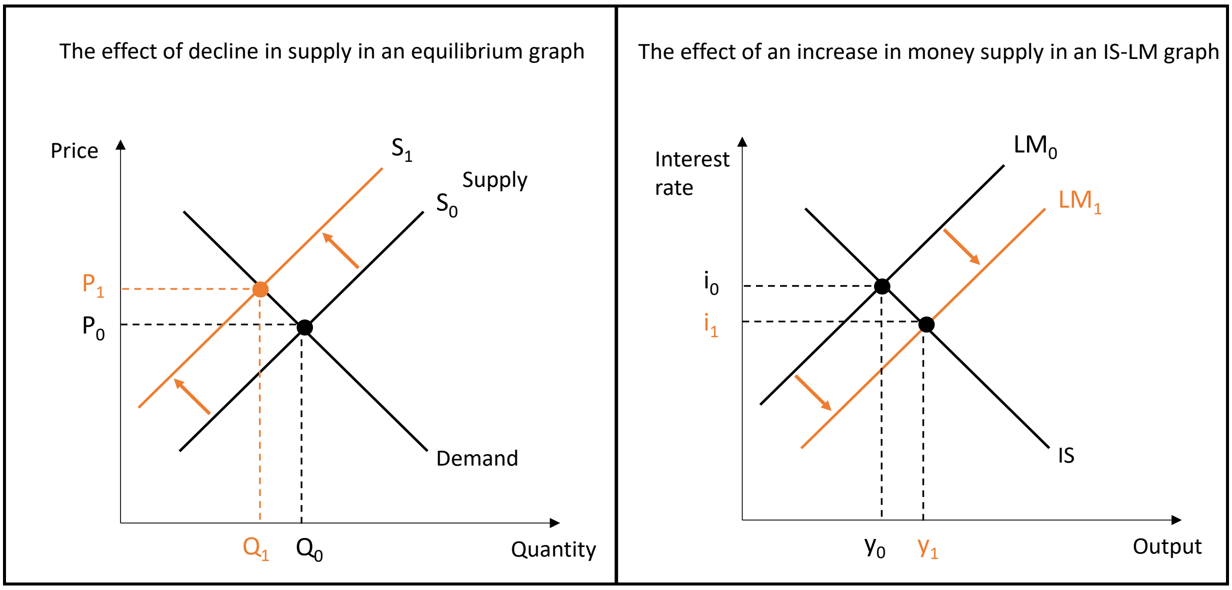


Figure 14. Marshall graph and IS–LM graph (own illustration, based on Feenstra and Taylor, 2014)

Content knowledge might thus be relevant not only due to a familiarity with the content or terminology but also due to knowledge about how certain relationships are visualized and what operators are used when working with the representation in the domain. This link between the ability to work with different representation and domain knowledge was demonstrated empirically in a study by Nitz et al. (2014). They measured the ability of biology students (931 students in the 12th grade, secondary school) to translate between different representations that are commonly used when teaching photosynthesis (e.g., photorealistic representations, graphs, chemical equations and formulae). To solve the items in their measure, the students needed to be able to use representations to describe scientific concepts; generate and/or select a representation; and lastly identify, describe and analyze features of different representations. To assess content knowledge, they used a measure with different multiple-choice items that covered the central concepts of photosynthesis. With confirmatory factor analyses, the authors revealed that the two concepts are two distinct variables and that both variables are positively related to each other. They concluded that the two abilities, namely, representational competence and content knowledge, were "*interactively related, but empirically distinguishable*" (Nitz et al., 2014, p. 18).

Although the relationship between content knowledge and graph or representational competence is mostly beneficial (i.e., in the sense that learners benefit from their previous knowledge), their expectations can also lead to an overestimation of graphically depicted effects (Freedman & Smith, 1996) or to arguing for an effect that is not visible in the graph (Shah & Hoeffner, 2002). All in all, it becomes evident that learner characteristics – especially prior knowledge – play a major role when the ability to work with graphs is analyzed. Knowledge of the terms and concepts used and knowledge of the reading rules of a domain representation influence how well learners can work with graphs.

1.3.5 Section summary

Different terms and models are used to analyze the ability to work with graphs and other visual representations. One central difference between the models pertains to the abilities they expect of a graph-competent person. For example, graph literacy researchers might be more interested in whether learners can extract data points (e.g., Åberg-Bengtsson & Ottosson, 2006), whereas representational competence researchers focus on how learners use (multiple) representations to answer domain questions (e.g., Stieff et al., 2011). Due to these discrepancies, they use a variety of different task when they consider graph or representational competence (for an overview, see Figure 12). Apart from the task characteristics, learner characteristics (e.g., prior knowledge), as well as representational design characteristics (e.g., different types of graphs), are relevant influencing factors for the effectiveness of graphical representations in the learning process (Shah et al., 2005; see Figure 15).

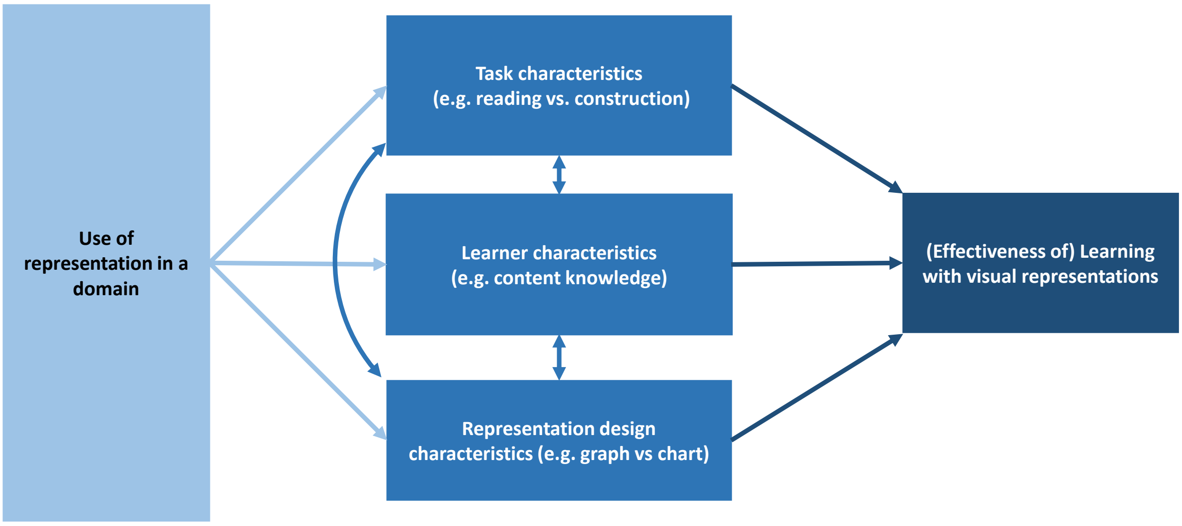


Figure 15. Influencing factors for the effectiveness of learning with visual representations (own illustration, partially based on influencing factors in Shah et al., 2005)

These factors might depend on the domain as different domains use different graphical representations (differently). Equilibrium graphs, such as Marshall's supply-and-demand graph, for example, are typical visual representations of economics (P. Davies & Mangan, 2007). Furthermore, these three influencing factors interact with one another; for example, learner characteristics such as content knowledge might be less relevant for a simple graph-reading task (e.g., how high is x at y?), whereas it is necessary to answer more complex interpretation questions (e.g., why is x high at point y?). Learners thus face a multitude of challenges when working with visual representation (e.g., difficulties in reading, interpreting or constructing a graph) that differ between domain, representation and learner (e.g., Boels et al., 2019; Glazer, 2011; Kotzebue et al., 2015; Shah et al., 2005). When analyzing the typical mistakes by learners in secondary education, it becomes evident that the most challenging aspects of various tasks involving graphical representation are the drawing of general conclusions and the connection of the underlying (domain) principle to the visual representation. This is illustrated when learners cannot answer open interpretation questions (Tairab & Khalaf Al-Naqbi, 2004), fail to separate causality from correlation in evaluation (Aoyama, 2007) or are unable to accurately construct equilibrium graphs in economics (Strober & Cook, 1992). Although the exact distinctions between different "levels" is not yet clear, most of the evidence indicates a relationship between the complexity of a reading or interpretation task and its difficulties for learners (Curcio, 1987; Friel et al., 2001; Lachmayer, 2008).

So far, mostly models of the ability to deal with single representations and the resulting empirical outcomes have been discussed. When learning within a domain, however, visual representations are usually one of many representational forms that are used to discuss models, ideas and systems. Assuming that the connection to the domain (and thus to other representations) is one of the central challenges when working with visual representations, it is necessary to focus on how learners use and connect multiple representations to gain an understanding of domain concepts. Therefore, the next section focuses on the ability to work with multiple representations and consequently on instructional support strategies that help learners to connect graphical representations with other representations to gain a more comprehensive understanding of domain concepts.

1.4 Instructional support for learning with (multiple) representations

1.4.1 Multimedia and multiple representations

The term "multimedia" is used in different ways: Mayer (2002) has promoted the understanding that multimedia is the presentation of learning material as a combination of words (e.g., by spoken or printed text) and graphical representations (e.g., not only graphs and diagrams but also dynamic graphics such as video and animations). This definition includes, for example, classical chalk-and-talk lectures, learning material in textbooks (as long as it does not contain only text or only pictures) and a screencast where PowerPoint slides are combined with explanations. The combination of text and graphical representation can be analyzed from different viewpoints: investigating the relevant devices and technologies (e.g., computer screen vs. blackboard); (re)presentation modes (text and graph vs. text and video); and sensory modalities such as pictures and printed text vs. pictures and spoken text (Mayer, 2014c). Due to these different perspectives, the term is not used consistently across the literature, and some authors refer primarily to digital media or different modalities when they write about multimedia (e.g., Korucu & Alkan, 2011). The term "multiple representations" is used to describe the part of multimedia research that focuses on multiple representational formats that learners are confronted with in learning material (Ainsworth, 1999, 2014). Therefore, the term "multiple representation" is mainly used in this dissertation as it underlines the primary interest in connecting multiple formats.

The overarching research goals of multimedia and multiple representation research are to understand how learners use and process multiple formats and, from a top-down perspective, to find rules or principles that can be used to design learning material and learning tasks in a way that optimizes understanding (Mayer, 2014a). Since actual understanding cannot (yet) be observed, researchers measure the effect of material (which is created based on theoretical assumptions) in experimental designs indirectly (i.e., as the learning outcome in performance tests). Newer technology, such as eye tracking, has recently been used to shed light on the process (rather than the outcome) of learning with multiple representations, and the resulting data can be used to gain a more nuanced understanding of learning with multiple representations (e.g., van Gog & Scheiter, 2010). As the main focus of this dissertation is on graphical representations and their connection to the (economic) domain, this section does not cover the entirety of multimedia or multiple representation research but is instead focused on the theoretical models and instructional support that are relevant for the combination of written text and static graphical representations.

1.4.2 Theoretical foundations of learning from text and graphical representation

Apart from the explicit cognitive process that involves the processing of text and graphical representations (i.e., mental model-building), the general cognitive architecture is relevant for the effectiveness of different instructional designs regarding the support of text–picture integration (Paas & Sweller, 2014). The human cognitive capacity is restricted due to the limited capacity of working memory load (Baddeley & Hitch, 1974). Therefore, the way in which these memory resources are used (or more precisely, how they are loaded) while learning is relevant for learning outcomes.

The cognitive load theory (Sweller, 1994, 2020) provides a framework to analyze, explain and predict learning outcomes resulting from learning environments consisting of graphical representation and text, and it is therefore relevant for this dissertation. The theory is based on the idea that relevant working memory resources are influenced by the complexity of the learning content (intrinsic cognitive load) and the design of the learning material (extrinsic cognitive load). Therefore, the aim of learning material and instructional design is to create a learning environment in which the limited working memory resources deal with intrinsic rather than extrinsic cognitive load (Paas & Sweller, 2014). For example, when a graphical representation does not add information and is instead redundant to the text, it might not influence intrinsic cognitive load (as it is irrelevant concerning the complexity of the relevant learning content) but depletes cognitive resources as learners still process the pictorial representations, thereby producing extrinsic cognitive load. This effect is termed the redundancy principle (Kalyuga & Sweller, 2014). Furthermore, the theory makes it necessary to identify factors that influence the processing of graphical representation and text, as those factors could be manipulated to reduce extrinsic cognitive load. For this, however, a framework for text–picture integration is necessary.

One of the theoretical starting points of contemporary frameworks for text–picture integration is the dual-code theory (Paivio, 1990). Paivio argued that contrary to prior assumptions, textual and pictorial information[3] is processed in functionally independent, but interconnected systems as mental representations. The interaction between the textual and pictorial channel might enable a positive effect of text–picture learning material regarding recall effects because the successful retrieval of the information from one representation might be sufficient to obtain the

[3] Paivio originally separated not according to representational form alone but also regarding modality and used the terminology “verbal“ and “non-verbal“ to describe the two channels. His assumptions were later used for other theories mainly based on the representational form rather than the modality (Mayer, 2014b), and accordingly, the later terminology is used.

information of the other representation as well. In other words: when learners have connected pictures and text in their mental model, the recall of a picture is sufficient to obtain the text information as well (Paivio, 1990).

Based on the dual-code assumption, two different contemporary models describe learning from text and pictures. In Mayer's cognitive theory of multimedia learning (CTML; Mayer, 2014b), the process begins by selecting information from text and pictures. In a second step, the information is organized within two mode-specific subsystems, resulting in separate mental representations for verbal and pictorial information, respectively. In the final step, the information in the two different subsystems is connected in one coherent mental model. In comparison, the integrated model of text and picture comprehension (ITPC) by Schnotz and Bannert (2003) also assumes that text and pictures are – in a first step – processed without interaction between the two representation modes. However, for text processing, this model states that learners must first process text on a descriptive level. In a second step, with the help of their mental lexicon and background knowledge, learners identify the text meaning based on semantic analysis, thereby constructing a propositional representation of it. Only then, a mental model is formed. The depictive representation (i.e., the picture) is used as a template for creating an internal visual representation. In the last step, text and pictures are interlinked through the identification and mapping of surface-level structures (verbal and pictorial elements) and deep-level semantic structures (Schnotz & Wagner, 2018). Therefore, in the second model, the process is initially more text-driven, although pictures may have a scaffolding function in the construction of the mental model and later, when recalling, might be used as an easily accessible visual tool (Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013; Schnotz et al., 2014).

Two central conclusions can be drawn from these models. First, the integration of text and pictures is a prerequisite for successful learning in text–picture environments; that is, learners must understand the particular representations, find the relevant correspondences and use them to connect the two representations in a mental model. Second, in both models, prior knowledge plays a relevant role in the construction of the mental model and its storage in long-term memory.

1.4.3 Integrating text and pictures

In accordance with the theoretical assumptions, different studies have demonstrated that, generally speaking, learning outcomes can be higher when learners study material that is composed of text and pictures compared to text alone (Carney & Levin, 2002; Guo, McTigue, Matthews, & Zimmer, 2020; Guo, Zhang, Wright, & McTigue, 2020; Levie & Lentz, 1982;

Levin et al., 1987), which is termed the multimedia effect (Mayer, 2014b). This effect, however, is not always visible, as adding graphs to a lecture can also lead to lower performance (Cohn et al., 2001; for more detail, see Subsection 1.2.3.1). The reason for this is that simply presenting a learner with two representations is not a sufficient requirement to increase the learning outcome. Similarly to graphs as single representations, other relevant factors determine how well learners can use an additional representation. One of the most prominent frameworks to describe these influencing factors is the **de**sign, **f**unctions, **t**asks (DEFT) framework (Ainsworth, 1999, 2006, 2014). In addition to parameters such as design and task (which is seen as the interaction between task and learner characteristics), Ainsworth argues that the function or role of a representation in a multiple representation learning environment is an important influencing factor. The framework thus extends the previously described frameworks for graphical representations by an important component and is used in this dissertation as the theoretical assumption to analyze the use of multiple representations. As the task, learner characteristics and representational design characteristics have already been discussed for graphical representations as single representations (see Subsection 1.3.4), only considerations on functions and interactions between the influencing factors that focus on text–picture material are further elaborated at this point.

According to Ainsworth (2006), additional representations can only be effective if they serve one of three functions, which she terms (1) complementary, (2) constraining and (3) constructing. With regard to the first function, representations complement one another either by offering unique information or by supporting different inferences. A classic example of the former is when a text about the human body is accompanied by an illustration such as an anatomical sketch. The illustration presents unique information as it informs the reader about the spatial position of different parts explained in the text, and text and visual representation thus complement each other. In terms of supporting different inferences, the author means that even if the information in two representations is the same, both representations might still be useful as each representation might be particularly useful for a certain task or a certain learner (Ainsworth, 2006). If learning material consists of a graph and a table that contain the same information (i.e., that display the same data), then the question of whether they complement each other depends on the task: When the learner needs to identify a trend, the graph might be more beneficial, whereas data points would be more easily found in the table (Meyer, Shinar, & Leiser, 1997). Therefore, when both tasks are given, the two representations complement each other. In contrast, when one representation is irrelevant for the task (e.g., when only data points are asked for), it is accordingly more sensible to only use one representation (e.g., only the table). Furthermore, learner characteristics – especially prior knowledge – are relevant. In a series of studies, Kalyuga, Chandler and Sweller (1998) demonstrated that inexperienced learners profit from a redundant text when learning from the visual representation, whereas experts are hindered by the text and perform better when only the visual representation is shown. This phenomenon is called the expertise reversal effect (Kalyuga, 2014) and generally states that design principles that help novice learners might be ineffective or even counterproductive for experts. This can be explained with the cognitive load theory (see above): For novice learners, an additional representation or a certain instruction might reduce the complexity of the material (as it helps them, for example, to select the correct representation for the right task) and thus reduce intrinsic load. An expert learner, however, sees additional representations or instructional guidance as an increase in extrinsic load because he or she already knows how to use (which) representation for the learning task.

With respect to the second function, using word constraining with a positive connotation, Ainsworth (2006) indicates that a representation might function as a scaffold to help a learner understand or more clearly interpret another, more complex representation. For example, in a study by Madden, Jones and Rahm (2011), learners discussed different chemical representations (equations, molecular sketches, tables and graphs), and experienced students in particular used a single representation that they were familiar with in a heuristic manner to interpret other representations with which they were not familiar (e.g., they understood how an equation could be manipulated and, as a result, how graphs or tables might be obtained from it).

The third function in Ainsworth's (2006) framework is the construction of a deeper understanding of the domain. Ainsworth (2014) argues that multiple representations might be aiding learners to achieve insights if they can separate the shared invariant features (the domain principle) from the properties of the individual representations. For teaching, this means that the ability to make connections and translate between representations might be an end in itself because when learners understand these connections and can see the invariant features, they identify the principle that is represented and thus become experts in the domain. This function of multiple representations is evident when experts (compared to novices) are able to explain domain principles with the help of multiple representations (Kozma, 2003).

The central conclusion that can be drawn from the framework is that no combination of representations is universally the best; rather, they must be chosen carefully in accordance with the design, task and learner so that they can function as unique learning opportunities.

Moreover, to unravel their potential, integration of the different representations is a relevant prerequisite as the representations cannot fulfill their function(s) if a learner is not able to identify connections. This, however, is not an easy task, and learners fail to integrate representations for different reasons (Renkl & Scheiter, 2017).

1.4.4 Challenges in learning from text and pictures

The first and “simplest” explanation for why learners fail to integrate text and pictures is that they sometimes ignore visual representations and focus solely or mainly on the text (e.g., Rayner, Rotello, Stewart, Keir, & Duffy, 2001; Schmidt-Weigand, Kohnert, & Glowalla, 2010; Schwonke, Berthold, & Renkl, 2009). In a study by Schmidt-Weigand et al. (2010), the authors used eye tracking to analyze how learners shift their attention in a learning environment consisting of text and dynamic visualization. The material was presented in one of three conditions: The exploratory text was either spoken, written close to (inside of) the animated visualizations or written below the animations. A general finding was that learners in the latter two conditions mainly focused on the text and spent less attention on the visual representation. As learning success was related to the time learners’ spent looking at the animated visualization, the first condition produced the highest results, although no difference was found between the other two conditions. Furthermore, learners might not identify the relevant parts of each representation and instead focus on information that is perceptually salient but not relevant for the task at hand (Renkl & Scheiter, 2017). This is particularly relevant for more complex visual representations, and the omission of task-unnecessary information within a representation consequently leads to higher learning outcomes (Canham & Hegarty, 2010; Hegarty, Canham, & Fabrikant, 2010).

Even when learners attempt to use the visual representations (and they are reduced to the most important information), they might lack the necessary graph comprehension or representational competence to understand or use them properly (e.g.., Boels et al., 2019; Glazer, 2011 / see Subsection 1.3.4). As a result, the learners are unable to connect the visual representation to other representations such as texts. Closely related to missing representational competence is missing domain knowledge (e.g., Nitz et al., 2014; Stieff et al., 2011 / see Subsection 1.3.4.3).

The relevance of graph-reading and -construction ability as well as content knowledge was demonstrated in two studies by Stern et al. (2003). In the first study, the authors tested participants with varying expertise regarding content knowledge and graph comprehension: business university students (presumed to have high content knowledge and high graph comprehension), math and computer science majors (presumed to have low content knowledge and high graph comprehension), business students from vocation school (presumed to have medium content knowledge and low graph comprehension) and humanities students (presumed to have low content knowledge and low graph comprehension). Their research aim was to determine whether “cross-content transfer” in text processing would be possible by means of a graphical representation and how content knowledge and graph comprehension would interact with the transfer. Therefore, every student from every group was randomly assigned to one of three learning text conditions: an active graph condition, where learners read a text, received a table and were asked to use the table to construct a graph; a passive graph condition, where learners received the text with an already completed graph; and a control condition, where learners received a text (with similar content to the second text) but no graph. As a dependent variable, all groups were required to read a second text with different content without a graph and answer transfer questions regarding the new text. The information that was necessary to answer the questions could mainly be achieved by using the same type of graphical representation as in the learning text; the participants were informed that drawing a graph might help them to answer the questions.

The following pattern emerged from their results: For the domain and graph experts (economic university students) as well as the graph experts (math/computer science students), learners profited more from the active than the passive and more from the passive than from the no-graph (but similar content) condition. Moreover, learners with less expertise – that is, learners with medium content but no graph knowledge (economic vocational students) and those with no content and no graph knowledge (humanities students) – were mostly unable to transfer the knowledge from the learning text independent of the conditions, and they thus did not profit from actively creating or passively encountering the graph. When the authors analyzed the graphical representations constructed by the two “novice” groups, they concluded that a possible explanation for the lack of effects of active graphing was that students without graph comprehension were unable to construct the graphs accurately. In a second study, Stern et al. (2003) then demonstrated that students with medium content knowledge but no graph knowledge (vocational school students) profited from creating a graphical representation as a learning task when they were supported in the construction (e.g., when they were provided with a printout of a coordinate system with labeled axes).

In summary, to optimally learn from text and pictures, learners must identify the most relevant information in both representations and then combine them in a mental model by identifying the structural or functional connections. This integration process is difficult as learners tend to focus on the textual representation, and even when considering the visual representation, they are not always able to identify the most relevant information. Potential explanations are the lack of relevant learner abilities (strategy, domain knowledge or representational competence) and an unfavorable design. To support learners with the integration of different representations, numerous approaches have been developed that focus either on promoting understanding for single (primarily graphical) representations or on the combination of text and graphical representations.

1.4.5 Instructional support for single graphical representations

Since the understanding of a single representation is a prerequisite for the integration of multiple representations, some interventions focus mainly on supporting students in that regard (e.g., Cromley, Perez et al., 2013; Mautone & Mayer, 2007; Miller, Cromley, & Newcombe, 2016; Shah, Mayer, & Hegarty, 1999). Based on the influencing factors for graph and representational competence, the interventions usually either focus on learner characteristics or optimize the material design for a certain task.

Regarding the design of graphical representations. Shah et al. (1999), for example, redesigned graphs in a way that the relevant information was presented in visual chunks (e.g., they changed the graphical format bar to line graphs when the question involved trends or transformed the relevant variable to a percentage) so that graph viewers did not need to compute the values in their head while working with the graph. Compared to the original graphs, where the relevant information was not perceptually grouped, the use of the redesigned graphs led to better identification of central trends by the students.

In comparison, Cromley, Perez et al. (2013) focused on learners – more precisely, on their knowledge about the relevant reading rules for logical and realistic representations in biology, which they termed "convention of diagrams." The authors implemented a training wherein the reading rules of graphical representations were explained and practiced (e.g., that arrows can have different meanings in different charts, for example movement or evolutionary change, or that captions are important and should be considered as part of a graph analysis). In classrooms where the training was implemented via a workbook, 10^{th}-grade students achieved higher comprehension scores on visual representations (that were not included in the training) compared to a business-as-usual control group (Cromley, Perez et al., 2013; Miller et al., 2016).

These interventions (for an overview, see Figure 16) support learners in understanding the visual representations and thus ensure that the relevant prerequisites for integration of different representations are met. As integration itself is difficult, further instructional approaches aim to support students in making connections between text and visual representations.

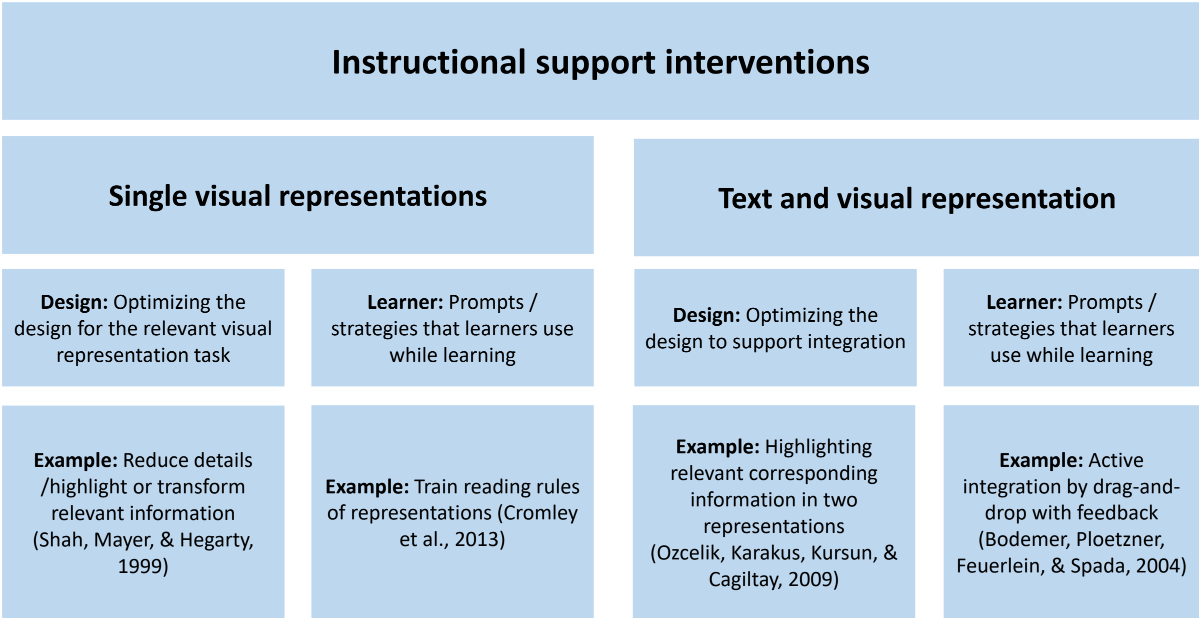


Figure 16. Overview of different instructional support interventions for single and multiple representations (own illustration, partially based on Renkl and Scheiter, 2017)

Renkl and Scheiter (2017) distinguish between interventions due to their focus on either the design of learning material or learner-centered instructions (Renkl & Scheiter, 2017; for an overview, see Figure 16). The former includes all instructional support, such as the spatial integration of text and pictures, that focuses on the design and is thus under the control of the supplier of teaching material (Ayres & Sweller, 2014; see below). The latter refers to strategies and prompts, such as reading strategies, that learners can use when they engage with text–graph material (Seufert, 2019; see below). Due to their specificity to the learning material and task, design strategies are often easily implemented, but they do not allow learners to develop a strategy that can be used with “original” learning material, such as textbook pages. Both are thus relevant from an educational perspective and serve different functions: When teachers construct or choose teaching materials, they can construct or select those materials in a way that promotes understanding. In addition, during teaching, to develop domain and representational competence that is transferable, strategies can be taught on material that is not specifically designed to support text–picture integration.

1.4.5.1 Design-oriented instructional support

Multiple approaches support learning with text and pictures via learning material design (e.g., Ayres & Sweller, 2014; Low & Sweller, 2014; van Gog, 2014). As integration is dependent on different factors, the same strategy might be more or less advantageous for different learners and learning material (Renkl & Scheiter, 2017).

Designing material in a way that text and pictures are in spatial and temporal proximity, or even physically integrating text into pictures, promotes learning from the representations as it optimizes the distribution of attention and reduces extrinsic cognitive load (Ginns, 2006). This finding is termed “spatial or temporal contiguity” or the “split-attention effect” (Ayres & Sweller, 2014; Ginns, 2006; Mayer & Fiorella, 2014). In his meta-analysis, Ginns (2006) summarized the results of 50 studies from various domains and concluded that this effect is larger for novice learners and for material with high element interactivity, which is defined as material that generates “*the need, during learning, to simultaneously hold multiple associated elements and their relations in working memory, understand those elements and their relations in concert, and hence construct a schema*.” (p. 512).

A second approach to promoting integration via design is a change in modality from written to spoken text (Low & Sweller, 2014). Although this so-called modality effect might help students to focus more on the graphical representation (Crooks, Cheon, Inan, Ari, & Flores, 2012; Schmidt-Weigand et al., 2010), only hearing the text might demand higher cognitive resources, especially for longer text, as learners need to remember the relevant text information and cannot use text strategies such as rereading (Renkl & Scheiter, 2017). The modality effect is thus more relevant for material that is not mainly based on text information (Leahy & Sweller, 2011).

Lastly, the term “signaling” or “cuing” is used to describe design elements that aim to direct one’s attention to the relevant correspondences between different representations (van Gog, 2014). This might be done, for example, by color coding (Ozcelik, Karakus, Kursun, & Cagiltay, 2009) or by using labels (Mayer & Johnson, 2008). Eye-tracking studies have demonstrated that signaling is effective in guiding a learner’s visual attention to signaled elements more frequently and earlier (Ozcelik, Arslan-Ari, & Cagiltay, 2010; Richter & Scheiter, 2019; Scheiter & Eitel, 2015). In different meta-analyses, signaled material is superior to not-signaled material regarding learning outcomes with small to medium effect sizes (Alpizar et al., 2020; Richter et al., 2016; Schneider et al., 2018). In some (but not all) studies, prior knowledge has been identified as a possible moderator: Signaling is primarily effective for low prior knowledge learners, whereas experts either do not profit or are even hindered in learning in an expertise-reversal effect (Richter & Scheiter, 2019; Richter, Scheiter, & Eitel, 2018).

To conclude, from a design perspective, text and graphical representations should be presented in a spatially integrated format, and especially for low-knowledge learners, highlights of the most relevant corresponding information are beneficial. A change in text modality to spoken text is only useful when the text is not the main information source.

1.4.5.2 Learner-centered instructional support

Researchers have tested various approaches that aim to support learners to use text and visual representation in a way that promotes learning. Renkl and Scheiter (2017) generally distinguish training studies from prompting.

In training interventions, learners practice text–picture integration strategies over a certain duration that can last from only a few minutes (e.g., Mason, Pluchino, & Tornatora, 2016) up to several weeks (e.g., Cromley, Bergey et al., 2013; Seufert, 2019). For instance, in a study by Seufert (2019), learners received a three-week training where they were taught how to work with text and visual representations. In the training, the students used a workbook that contained step-by-step explanations of how to apply reading strategies, first for single representations and then for integrating text and pictures (e.g., by identifying corresponding elements and structures in text and pictures). The training, however, was not effective for all learners when compared to a control group; only learners with high prior knowledge were able to benefit from practicing the reading strategies. This is in line with similar findings from earlier studies (Seufert, 2003; Seufert, Jänen, & Brünken, 2007). In contrast, Mason et al. (2016) demonstrated that when learners observed an expert's eye movement as an example of how to learn in a text–picture environment in a short video, they used the same strategies to connect text and pictures and profited in terms of learning outcome. Especially students with low prior knowledge benefited from the intervention.

Prompts generally aim at triggering a certain learner activity – in this case, a strategy to integrate text and pictures – and can be presented as part of the learning material. The approaches range from unspecific hints, such as a prompt to pay special attention to correspondences (Mayer, Dow, & Mayer, 2003), to more complex prompts where learners are required to connect representations by either performing a task, such as a drag-and-drop task (Bodemer et al., 2004), or writing important concepts from the text close to the visual representations (Leopold, Doerner, Leutner, & Dutke, 2015) while learning. In the study by Bodemer et al. (2004), 81 higher education students were shown text and visual representations in one of three conditions: The material was either physically integrated or physically separated, or learners were asked to connect the representations by drag-and-drop. In the drag-and-drop condition (termed active integration), learners were only able to connect the relevant parts correctly (i.e., they received system feedback if their connections were inadequate). Students achieved higher learning outcomes regarding comprehension in this active integration condition, which the authors explained with a more optimized use of cognitive resources. One problem with this intervention, however, is that learners could not transfer the strategy for material that was not designed in a manner that allowed for dragging and dropping, such as a classical textbook page with text and diagrams.

Furthermore, Leopold et al. (2015) analyzed how writing important concepts from the text next to the corresponding parts of the associated visual representation influenced learning outcomes for secondary students. They found that students did not achieve higher learning outcomes from this intervention compared to a group of learners who only received text and pictures without explicit strategy instructions. Nevertheless, within the group, they could demonstrate that the accuracy of referential connections was associated with learning performance. A possible explanation for this finding is a moderating effect of prior knowledge or experience with material that involves text–graph environments, as learners with higher prior knowledge are better at generating higher quality connections.

All in all, findings for learner-centered instructional support are inconclusive. The effect of instruction is sometimes moderated by prior knowledge: So far, it seems that if learners are required to use more general strategies to integrate text and graphical representations, then only learners with high prior knowledge profit (Seufert, 2019). However, when specific instruction is given or relevant connections are already highlighted, primarily novices are assisted in their learning process (Richter et al., 2018).

1.4.6 Section summary

The central assumption of the dual coding theory is that for learning material that consists of text and visual representations, both representations are processed in functionally independent, but interconnected systems (Paivio, 1990). Learning from text and pictures is not effective per se, and whether an additional representation can fulfill a relevant function is dependent on the combination of design, learner characteristics and task characteristics (Ainsworth, 2006). One central prerequisite for their effectiveness is the integration of text and visual representation (Mayer, 2014b; Schnotz & Bannert, 2003). Integration, however, is difficult, as learners might ignore the visual representations, might lack the relevant abilities to use them or cannot connect the visual representations to the text (Seufert, 2003).

Different strategies to support learners in either understanding the visual representation or in making the relevant connections have emerged from instructional design and multiple representation research (Renkl & Scheiter, 2017). The approaches that are used to aid integration can be distinguished as they either modify the learning material design (e.g., by signaling; van Gog, 2014) or support a learner's strategy when working with text and graphical representation (e.g., by asking the learner to integrate representations via drag-and-drop;

Bodemer et al., 2004). These strategies have a number of advantages and disadvantages: Design-oriented interventions can be controlled by the creator or supplier of the learning material and are easy to implement, and especially novice learners who do not have a strategy for integration might be supported in their attempts to integrate the material. From an educational perspective, however, only using design-oriented interventions seems problematic as they are task- and material-specific; thus, when learners are confronted with “authentic” material that is not optimized to support them in the integration of multiple representations, comprehension might be challenging. In contrast, strategies for learners (e.g., reading strategies for integrating text and picture; e.g., Seufert, 2019) can be used for different text–picture materials, but they are not easy to learn, as they take time and mental effort.

Regarding effectiveness, there is more evidence for the positive effect of design support (for signaling, c.f. Richter et al., 2016), whereas the findings for learner-centered interventions are less conclusive. This discrepancy might partly be explained by the interplay of learner characteristics such as prior knowledge, the complexity of instructional support (strategy) and the complexity of the learning material. There are, for example, some indications that more complex strategies might work especially well for learners with high prior knowledge (Seufert, 2019), whereas specific learning strategies as well as design interventions such as signaling are more effective for low prior knowledge learners (Richter et al., 2018). The exact nature of the influence of prior knowledge, learning material complexity and strategy implementation, however, is still an open question for design-centered and even more for learner-centered instructional support.

1.5 Research questions and methodology

So far, graphical representations have been discussed from three (sometimes overlapping) perspectives. These perspectives consider the following: the use of graphical representations as visualizations of domain models with a focus on economic education, the empirical models of a learner’s ability with a focus on graph literacy and graph comprehension and lastly, the effectiveness of visual representations as part of a multiple learning environment with a focus on text-picture integration and instructional support.

Based on the identified need for comprehension of graphical displays in general (e.g., Shah et al., 2005) as well as for economic education specifically (e.g., P. Davies & Mangan, 2007), the three perspectives lay the groundwork for the overarching goal of this dissertation: to understand and promote learning from visual representations. To date, most studies have focused on a learner’s ability to work with graphical representation either as a more general graph literacy skill (e.g., Åberg-Bengtsson & Ottosson, 2006; Galesic & Garcia-Retamero, 2011) or – based on the argument that the ability to work with typical graphical representations is specific to a certain domain (Ainsworth, 2006) – as domain competence mainly in science education (e.g., Lai et al., 2016; McKenzie & Padilla, 1986; Nitz et al., 2014). Due to the specific use of different representations between the domains, the results from the second stream of research in particular might not be transferable from science to social science education (Westelinck et al., 2005).

Although studies in higher education demonstrate that learners have difficulties in understanding graphical representations (Strober & Cook, 1992), little is yet known about learners in secondary education. Moreover, studies that have focused on secondary education mainly evaluate the design of the representation (Aprea & Bayer, 2010) or analyze how the design influences learning outcomes (e.g., Jägerskog, 2020). Up to now, far too little attention has been paid to learner abilities (and challenges) or the use of representations in classroom settings. Therefore, the research aim of this dissertation is to analyze how well learners can work with graphs, how visual representations are used in the classroom and how learners can be supported in utilizing multiple representations to their full potential.

The present dissertation is a starting point to extend research regarding graphical representations further to the economic domain. To this end, it has three main objectives:

(1) to improve the understanding of the construct graph literacy and to establish an empirical baseline for graph literacy skills at the beginning of secondary economic education;

(2) to understand how graphical representations are used in learning material as well as in teaching in secondary economic education;

(3) to investigate how text–graph integration can be fostered with the help of material and learner strategies and how this instructional support is influenced by prior domain knowledge.

Three empirical studies are consequently carried out:

In the first study (*Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs*), based on frameworks mainly from math and science educational research, a multiple choice test was developed in which the ability of eighth graders (highest school track) to read line graphs, bar graphs and pie graphs was examined. We used graphs that learners might encounter in newspapers and on the internet that visualize data related to sustainable development. This topic was chosen not only because it is an important issue for both economic and science education but also because graphs are relevant in communicating central problems (such as CO_2 and greenhouse gas emissions, e.g., Ritchie & Roser, 2017). To add to the understanding of graph literacy, we investigated the relationship between task characteristics and item difficulty with item response theory. Furthermore, the relationship between graph literacy and other learner abilities, such as motivation, interest and content knowledge, were examined.

Study 2 (*Logical pictures in secondary economic education: Textbook analysis and teacher perception*) addressed the question of how graphical representations are used in learning material as well as teaching in secondary economic education. To this end, we used an exploratory approach and combined two empirical approaches: First, we analyzed graphical representation in textbooks and used anchor examples to describe typical graphs and charts, and we discuss their relationship to domain principles. Second, we interviewed economic teachers in secondary schools (highest school track) to understand how they use graphs and charts in their everyday teaching. We discuss not only quantitative usage or learning goals but also relevant abilities and challenges that learners face when working with visual representations from the teacher's point of view, and we compare them to findings from other domains.

In Study 3 (*How to support text–graph integration: Comparing the effects of passive and active signaling on learning outcomes*), we investigated how text–graph integration can be supported with the help of material design and learner strategies in two domains. For this purpose, we used a quasi-experimental design, where higher education learners studied economics and biology material with text and typical domain graphs in one of three conditions: (1) no instructional support, (2) design support by passive signals in the learning material and (3) learner support by prompting learners to highlight central equivalent information in text and graphs via highlighter (active signaling). We hypothesized that an active signaling task might be beneficial, as learners take a more active role in connecting graphical representation and task. Furthermore, the role of prior knowledge and strategy implementation as boundary conditions for the effective use of instructional support are discussed.

In the following chapters, the three studies are described in greater detail. In the last chapter, their findings are summarized and evaluated before potential implications for future research concerning graphical representations in (secondary) economic education are discussed.

1.6 References

Åberg-Bengtsson, L. (1999). Dimensions of performance in the interpretation of diagrams, tables, and maps: Some gender differences in the Swedish scholastic aptitude test. *Journal of Research in Science Teaching, 36*(5), 565–582. https://doi.org/10.1002/(SICI)1098-2736(199905)36:5<565::AID-TEA4>3.0.CO;2-L

Åberg-Bengtsson, L., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*(1), 43–62. https://doi.org/10.1002/tea.20087

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2-3), 131–152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2014). The Multiple Representation Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 464–486). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.024

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development*, *104*(1), 250. https://doi.org/10.1007/s11423-020-09748-7

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, *2*(3), 298–318.

Aprea, C., & Bayer, D. (2010). Instruktionale Qualität von grafischen Darstellungen in Lehrmitteln: Kriterien zu deren Evaluation: [Instructional quality of graphical representations in teaching materials: Criteria for their evaluation]. *Beiträge Zur Lehrerinnen- und Lehrerbildung*, *28*(1), 73–83.

Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Stanat, P., Tillmann, K.-J., & Weiß, M. (2001). *PISA 2000: Zusammenfassung zentraler Befunde [PISA 2000: Summary of central findings]*. Retrieved from https://pure.mpg.de/rest/items/item_2102433/component/file_3189150/content

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 206–226). Cambridge: Cambridge University Press.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (pp. 47–89). Amsterdam: Elsevier. https://doi.org/10.1016/S0079-7421(08)60452-1

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2001). Toward a model of learning data representations. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 45–50). Mahwah, NJ: Erlbaum.

Bauer, M., Hamm-Reinöhl, A., Podes, S., & Riedel, H. (2012). *Unterrichtswerk für die Oberstufe Wirtschaft: Märkte, Akteure und Institutionen [Textbook for higher grade level: Economics - Markets, Actors and Institutions]* (1. Auflage). *Kolleg Politik und Wirtschaft*. Bamberg: Buchner, C C.

Baumert, J., Bos, W., & Watermann, R. (1998). *TIMSS/III: Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich: Zusammenfassung deskriptiver Ergebnisse [TIMSS/III: International comparison of student performance in mathematics and the natural sciences at the end of upper secondary school: Summary of descriptive results]*. Retrieved from http://hdl.handle.net/11858/00-001M-0000-0025-A35F-4

Beck, K., Krumm, V., & Dubs, R. (1998). *Wirtschaftskundlicher Bildungs-Test (WBT); Handanweisung [Test of economic literacy; Handbook]*. Hogrefe, Verlag für Psychologie.

Berg, C. A., & Smith, P. (1994). Assessing students’ abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments. *Science Education*, *78*(6), 527–554. https://doi.org/10.1002/sce.3730780602

Bertin, J. (1983). *Semiology of graphics: Diagrams networks maps*. Madison: University of Wisconsin Press.

Boatman, K., Courtney, R., & Lee, W. (2008). See How They Learn: The Impact of Faculty and Student Learning Styles on Student Performance in Introductory Economics. *The American Economist*, *52*(1), 39–48. https://doi.org/10.1177/056943450805200103

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*(3), 325–341. https://doi.org/10.1016/j.learninstruc.2004.06.006

Boels, L., Bakker, A., van Dooren, W., & Drijvers, P. (2019). Conceptual difficulties when interpreting histograms: A review. *Educational Research Review*, 100291. https://doi.org/10.1016/j.edurev.2019.100291

Brueck, H. (2018). *The author of one of Bill Gates' favorite books says the world isn't as apocalyptic as you might think — here are five of his top reasons why*. Retrieved from https://www.businessinsider.com/bill-gates-factfulness-by-hans-rosling-favorite-book-review-2018-4?r=DE&IR=T

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, *20*(2), 155–166. https://doi.org/10.1016/j.learninstruc.2009.02.014

Carney, R. N., & Levin, J. R. (2002). Pictorial Illustrations Still Improve Students' Learning from Text. *Educational Psychology Review*, *14*(1), 5–26. https://doi.org/10.1023/A:1013176309260

Chiou, C.-C. (2009). Effects of concept mapping strategy on learning performance in business and economics statistics. *Teaching in Higher Education*, *14*(1), 55–69. https://doi.org/10.1080/13562510802602582

Cohn, E., & Cohn, S. (1994). Graphs and learning in principles of economics. *The American Economic Review*, *84*(2), 197–200.

Cohn, E., Cohn, S., Balch, D. C., & Bradley, J. (2001). Do Graphs Promote Learning in Principles of Economics? *The Journal of Economic Education*, *32*(4), 299–310. https://doi.org/10.1080/00220480109596110

Cook, M. (2011). Teachers' Use of Visual Representations in the Science Classroom. *Science Education International*, *22*(3), 175–184.

Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction*, *26*, 45–58. https://doi.org/10.1016/j.learninstruc.2013.01.003

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving Students' Diagram Comprehension with Classroom Instruction. *The Journal of Experimental Education*, *81*(4), 511–537. https://doi.org/10.1080/00220973.2012.745465

Crooks, S. M., Cheon, J., Inan, F., Ari, F., & Flores, R. (2012). Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, *28*(3), 1063–1071. https://doi.org/10.1016/j.chb.2012.01.010

Curcio, F. R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, *18*(5), 382–393. https://doi.org/10.2307/749086

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? *Higher Education*, *62*(3), 279–301. https://doi.org/10.1007/s10734-010-9387-6

Davies, P., & Mangan, J. (2007). Threshold concepts and the integration of understanding in economics. *Studies in Higher Education*, *32*(6), 711–726. https://doi.org/10.1080/03075070701685148

Demir, I., & Tollison, R. D. (2015). Graphs in Economics. *Economics Bulletin*, *35*(3), 1834–1847.

Dreher, A., & Kuntze, S. (2015). Teachers Facing the Dilemma of Multiple Representations Being Aid and Obstacle for Learning: Evaluations of Tasks and Theme-Specific Noticing. *Journal Für Mathematik-Didaktik*, *36*(1), 23–44. https://doi.org/10.1007/s13138-014-0068-3

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. https://doi.org/10.1016/j.learninstruc.2013.05.002

Elting, L. S., Martin, C. G., Cantor, S. B., & Rubenstein, E. B. (1999). Influence of data display formats on physician investigators' decisions to stop clinical trials: prospective trial with repeated measures. *Bmj*, *318*(7197), 1527–1531.

Feenstra, R. C., & Taylor, A. M. (2014). *International economics* (third edition). New York: Worth Publishers a Macmillan Education Company.

Findeisen, S. (2017). *Fachdidaktische Kompetenzen angehender Lehrpersonen: Eine Untersuchung zum Erklären im Rechnungswesen (Pedagogical content knowledge of preservice teachers: an empirical study of explanations in accounting*) [Doctoral Dissertation, Universität Mannheim]. Wiesbaden: Springer Fachmedien Wiesbaden.

Freedman, E. G., & Smith, L. D. (1996). The role of data and theory in covariation assessment: Implications for the theory-ladenness of observation. *The Journal of Mind and Behavior*, 321–343.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education*, *32*(2), 124–158. https://doi.org/10.2307/749671

Gal, I. (1998). Assessing statistical knowledge as it relates to students' interpretation of data. *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12*, 275–295.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–25.

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *31*(3), 444–457. https://doi.org/10.1177/0272989X10373805

Gates, B. (2018). *Why I want to stop talking about the "developing" world: Factfulness is one of the most educational books I've ever read.* Retrieved from https://www.gatesnotes.com/Books/Factfulness#

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, *47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Goerwitz, R. (2019). Review of Factfulness: Ten Reasons We're Wrong About the World—and Why Things Are Better Than You Think, by Hans Rosling, Anna Rosling Rönnlund, and Ola Rosling. *Numeracy: Advancing Education in Quantitative Literacy*, *12*(2). https://doi.org/10.5038/1936-4660.12.2.19

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25.

Gültepe, N. (2016). Reflections on High School Students' Graphing Skills and Their Conceptual Understanding of Drawing Chemistry Graphs. *Educational Sciences: Theory & Practice*, *16*(1), 53–81. https://doi.org/10.12738/estp.2016.1.2837

Guo, D., McTigue, E. M., Matthews, S. D., & Zimmer, W. (2020). The Impact of Visual Displays on Learning Across the Disciplines: A Systematic Review. *Educational Psychology Review*, *32*(3), 627–656. https://doi.org/10.1007/s10648-020-09523-3

Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do You Get the Picture? A Meta-Analysis of the Effect of Graphics on Reading Comprehension. *AERA Open*, *6*(1), 233285842090169. https://doi.org/10.1177/2332858420901696

Hadjidemetriou, C., & Williams, J. (2002). Children's graphical conceptions. *Research in Mathematics Education*, *4*(1), 69–87.

Harris, R. L. (2000). *Information graphics: A comprehensive illustrated reference*. Oxford: Oxford University Press.

Hattikudur, S., Prather, R. W., Asquith, P., Alibali, M. W., Knuth, E. J., & Nathan, M. (2012). Constructing Graphical Representations: Middle Schoolers' Intuitions and Developing Knowledge About Slope and Y-intercept. *School Science and Mathematics*, *112*(4), 230–240. https://doi.org/10.1111/j.1949-8594.2012.00138.x

Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 37–53. https://doi.org/10.1037/a0017683

Humphrey, T. M. (1992). Marshallian cross diagrams and their uses before Alfred Marshall: the origins of supply and demand geometry. *Economic Review*, *78*(2), 3–23.

Jägerskog, A.-S. (2020). *Making Possible by Making Visible: Learning through Visual Representations in Social Science* [Doctoral Dissertation, Stockholm University]. Stockholm: Department of Humanities and Social Sciences Education, Stockholm University.

Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 576–597). Cambridge: Cambridge University Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, *40*(1), 1–17.

Kalyuga, S., & Sweller, J. (2014). The Redundancy Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 247–262). Cambridge: Cambridge University Press.

Kerslake, D. (1981). Graphs. *Children's Understanding of Mathematics*, *11*(16), 120–136.

Korucu, A. T., & Alkan, A. (2011). Differences between m-learning (mobile learning) and e-learning, basic terminology and usage of m-learning in education. *Procedia - Social and Behavioral Sciences*, *15*, 1925–1930. https://doi.org/10.1016/j.sbspro.2011.04.029

Kosslyn, S. M. (1999). *Image and brain: The resolution of the imagery debate* (4. print). *A Bradford book*. Cambridge, Mass.: MIT Press.

Kosslyn, S. M. (2006). *Graph design for the eye and mind*. New York: Oxford University Press. Retrieved from http://www.myilibrary.com/?id=294674

Kotzebue, L. von, Gerstl, M., & Nerdel, C. (2015). Common Mistakes in the Construction of Diagrams in Biological Contexts. *Research in Science Education*, *45*(2), 193–213. https://doi.org/10.1007/s11165-014-9419-9

Kourilsky, M., & Wittrock, M. C. (1987). Verbal and graphical strategies in the teaching of economics. *Teaching and Teacher Education*, *3*(1), 1–12. https://doi.org/10.1016/0742-051X(87)90030-8

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*(2), 205–226. https://doi.org/10.1016/S0959-4752(02)00021-X

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*(9), 949–968.

Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representational Competence. In J. K. Gilbert (Ed.), *Models and Modeling in Science Education: Vol. 1*.

Visualization in Science Education (pp. 121–145). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3613-2_8

Lachmayer, S. (2008). *Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht [Development and validation of a graph competence model for biology education]* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. Retrieved from https://d-nb.info/1019667389/34

Lachmayer, S., Nerdel, C., & Prechtl, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht [Modelling of cognitive abilities regarding the handling of graphs in science education]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *13*, 161–180.

Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R., & Linn, M. C. (2016). Measuring Graph Comprehension, Critique, and Construction in Science. *Journal of Science Education and Technology*, *25*(4), 665–681. https://doi.org/10.1007/s10956-016-9621-9

Law, K. (2019). *Factfulness: Ten Reasons We're Wrong About The World…: How a book offering a bright worldview became a surprise bestseller.* Retrieved from https://www.standard.co.uk/lifestyle/books/factfulness-hans-rosling-review-a4180336.html

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*(6), 943–951. https://doi.org/10.1002/acp.1787

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1–64.

Lem, S., Kempen, G., Ceulemans, E., Onghena, P., Verschaffel, L., & van Dooren, W. (2015). Combining Multiple External Representations and Refutational Text: An Intervention on Learning to Interpret Box Plots. *International Journal of Science and Mathematics Education*, *13*(4), 909–926. https://doi.org/10.1007/s10763-014-9604-3

Lem, S., Onghena, P., Verschaffel, L., & van Dooren, W. (2013). On the misinterpretation of histograms and box plots. *Educational Psychology*, *33*(2), 155–174. https://doi.org/10.1080/01443410.2012.674006

Leopold, C., Doerner, M., Leutner, D., & Dutke, S. (2015). Effects of strategy instructions on learning from text and pictures. *Instructional Science*, *43*(3), 345–364. https://doi.org/10.1007/s11251-014-9336-3

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology Journal*, *30*(4), 195–232.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of picture in prose. In D. M. Willows & H. A. Houghton (Eds.), *The Psychology of Illustrations (Vol. 1)* (51–114). Harrisonburg, VA: R. R. Donnelley & Sons.

Lohse, G. L., Biolsi, K., Walker, N., & Rueter, H. H. (1994). A classification of visual representations. *Communications of the ACM*, *37*(12), 36–49. https://doi.org/10.1145/198366.198376

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (227-246). Cambridge: Cambridge University Press.

Ludewig, U., Lambert, K., Dackermann, T., Scheiter, K., & Möller, K. (2019). Influences of basic numerical abilities on graph reading performance. *Psychological Research*, 1–13. https://doi.org/10.1007/s00426-019-01144-y

MacArthur, R. H., & Wilson, E. O. (2001). *The theory of island biogeography*. *Princeton landmarks in biology*. Princeton: Princeton University Press. Retrieved from http://www.jstor.org/stable/10.2307/j.ctt19cc1t2

Madden, S. P., Jones, L. L., & Rahm, J. (2011). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research Practice*, *12*(3), 283–293. https://doi.org/10.1039/C1RP90035H

Marangos, J., & Alley, S. (2007). Effectiveness of concept maps in economics: Evidence from Australia and USA. *Learning and Individual Differences*, *17*(2), 193–199. https://doi.org/10.1016/j.lindif.2007.03.003

Marks, M., & Kotula, G. (2009). Using the Circular Flow of Income Model to Teach Economics in the Middle School Classroom. *The Social Studies*, *100*(5), 233–242. https://doi.org/10.1080/00377990903221939

Marshall, A. (2009). *Principles of economics: unabridged eighth edition*: Cosimo, Inc.

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology*, *47*(6), 1083–1095. https://doi.org/10.1111/bjet.12271

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, *93*(2), 377–389. https://doi.org/10.1037/0022-0663.93.2.377

Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, *99*(3), 640–652. https://doi.org/10.1037/0022-0663.99.3.640

Mayer, R. E. (2002). Multimedia learning. In B. H. Ross (Ed.), *Psychology of Learning and Motivation: v.41. Psychology of Learning and Motivation: Advances in Research and Theory* (1st ed., Vol. 41, pp. 85–139). Amsterdam: Elsevier textbooks. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. (Ed.) (2014a). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Mayer, R. E. (2014b). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Mayer, R. E. (2014c). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia Learning in an Interactive Self-Explaining Environment: What Works in the Design of Agent-Based Microworlds? *Journal of Educational Psychology*, *95*(4), 806–812. https://doi.org/10.1037/0022-0663.95.4.806

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 279–315). Cambridge: Cambridge University Press.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380–386. https://doi.org/10.1037/0022-0663.100.2.380

McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503–513. https://doi.org/10.1119/1.15104

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (togs). *Journal of Research in Science Teaching*, *23*(7), 571–579. https://doi.org/10.1002/tea.3660230702

Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, *39*(2), 268–286.

Miller, B. W., Cromley, J. G., & Newcombe, N. S. (2016). Improving diagrammatic reasoning in middle school science using conventions of diagrams instruction. *Journal of Computer Assisted Learning*, *32*(4), 374–390. https://doi.org/10.1111/jcal.12143

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2016a). *Bildungsplan des Gymnasiums Biologie [Curriculum Biology]*. Retrieved from http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_BIO.pdf

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2016b). *Bildungsplan des Gymnasiums Wirtschaft/ Berufs- und Studienorientierung (WBS) [Curriculum Economics, Vocational and Study orientation]*. Retrieved from http://www.bildungsplaene-bw.de/,Lde/LS/BP2016BW/ALLG/SEK1/WBS

Niederhaus, C. (2011). Informierende Bilder in Fachkundebüchern: berufsfeldabhängige Variation und didaktische Konsequenzen [Informative pictures in textbooks: variation and didactic consequences depending on the occupational field]. *Bwp@ Berufs-und Wirtschaftspädagogik–online, Ausgabe Bwp@ Spezial – Hochschultage Berufliche Bildung 2011*, 1–14. Retrieved from http://www.bwpat.de/ht2011/ft18/niederhaus_ft18-ht2011.pdf

Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction*, *31*, 13–22. https://doi.org/10.1016/j.learninstruc.2013.12.003

Nolan, D., & Perrett, J. (2016). Teaching and Learning Data Visualization: Ideas and Assignments. *The American Statistician*, *70*(3), 260–269. https://doi.org/10.1080/00031305.2015.1123651

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, *26*(1), 110–117. https://doi.org/10.1016/j.chb.2009.09.001

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education*, *53*(2), 445–453. https://doi.org/10.1016/j.compedu.2009.03.002

Ozmen, Z. M., Guven, B., & Kurak, Y. (2020). Determining the Graphical Literacy Levels of the 8th Grade Students. *Eurasian Journal of Educational Research*, *20*(86), 1–24. https://doi.org/10.14689/ejer.2020.86.13

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 27–42). Cambridge: Cambridge University Press.

Paivio, A. (1990). *Mental representations: A dual coding approach.* Oxford: Oxford University Press.

Peebles, D., & Ali, N. (2015). Expert interpretation of bar and line graphs: The role of graphicacy in reducing the effect of graph format. *Frontiers in Psychology*, *6*, 1673. https://doi.org/10.3389/fpsyg.2015.01673

Pindyck, R. S., & Rubinfeld, D. L. (2016). *Mikroökonomie [Microeconomics]* (8., aktualisierte und erweiterte Auflage). Hallbergmoos: Pearson. Retrieved from http://lib.myilibrary.com?id=921250

Preece, J. (1984). *Some theoretical issues related to interpreting graphs*: Milton Keynes, England: Institute of Educational Technology, The Open University, Walton Hall.

Raso, A. (2018). Visualisierungen zur Förderung des kognitiven Wissensaufbaus [Visualizations for the promotion of knowledge construction]. *Bwp@ Spezial at-1: Wirtschaftspädagogische Forschung Und Impulse Für Die Wirtschaftsdidaktik-Beiträge Zum 12. Österreichischen Wirtschaftspädagogikkongress*, 1–16.

Rau, M. A., Aleven, V., & Rummel, N. (2017). Making connections among multiple graphical representations of fractions: sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science*, *45*(3), 331–357. https://doi.org/10.1007/s11251-017-9403-7

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, *7*(3), 219–226. https://doi.org/10.1037/1076-898X.7.3.219

Reingewertz, Y. (2013). Teaching macroeconomics through flowcharts. *International Review of Economics Education*, *14*, 86–93. https://doi.org/10.1016/j.iree.2013.10.004

Renkl, A., & Scheiter, K. (2017). Studying Visual Displays: How to Instructionally Support Learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4

Richter, J., & Scheiter, K. (2019). Studying the expertise reversal of the multimedia signaling effect at a process level: evidence from eye tracking. *Instructional Science*, *47*(6), 627–658. https://doi.org/10.1007/s11251-019-09492-3

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Richter, J., Scheiter, K., & Eitel, A. (2018). Signaling text–picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology*, *110*(4), 544–560. https://doi.org/10.1037/edu0000220

Ritchie, H., & Roser, M. (2017). *CO_2 and Greenhouse Gas Emissions*. Retrieved from https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions

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. *The Reading Teacher*, *67*(1), 12–24. https://doi.org/10.1002/TRTR.1174

Rosling, H., Rosling, O., & Rosling-Rönnlund, A., (2018). *Factfulness: Ten reasons we're wrong about the world and why things are better than you think*. London: Sceptre.

Roth, W.-M., & McGinn, M. K. (1997). Graphing: Cognitive ability or practice? *Science Education*, *81*(1), 91–106. https://doi.org/10.1002/(SICI)1098-237X(199701)81:1<91::AID-SCE5>3.0.CO;2-X

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). Explaining the modality and contiguity effects: New insights from investigating students' viewing behaviour. *Applied Cognitive Psychology*, *24*(2), 226–237. https://doi.org/10.1002/acp.1554

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, *23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

Schnotz, W. (2001). Wissenserwerb mit Multimedia [Knowledge acquisition with multimedia]. *Unterrichtswissenschaft*, *29*(4), 292–318.

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

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instructional Science*, *36*(3), 175–190. https://doi.org/10.1007/s11251-007-9029-2

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

Schopf, C., Raso, A., & Kahr, M. (2019). How to give effective explanations: Guidelines for business education, discussion of their scope and their application to teaching operations research. *RISTAL-Research in Subject-Matter Teaching and Learning*, *2*, 32–50. https://doi.org/10.23770/rt1823

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology*, *23*(9), 1227–1243. https://doi.org/10.1002/acp.1526

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*(2), 227–237. https://doi.org/10.1016/S0959-4752(02)00022-1

Seufert, T. (2019). Training for Coherence Formation When Learning From Text and Picture and the Interplay With Learners' Prior Knowledge. *Frontiers in Psychology*, *10*, 193. https://doi.org/10.3389/fpsyg.2019.00193

Seufert, T., Jänen, I., & Brünken, R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior*, *23*(3), 1055–1071. https://doi.org/10.1016/j.chb.2006.10.002

Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in Cognitive Science*, *3*(3), 560–578. https://doi.org/10.1111/j.1756-8765.2009.01066.x

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The Comprehension of Quantitative Information in Graphical Displays. In P. Shah & A. Miyake (Eds.), *Cambridge Handbooks in Psychology*. *The Cambridge Handbook of Visuospatial Thinking* (pp. 426–476). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.012

Shah, P., & Hoeffner, J. (2002). Review of Graph Comprehension Research: Implications for Instructions. *Educational Psychology Review*, *14*(1), 47–69. https://doi.org/10.1023/A:1013180410169

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, *91*(4), 690.

Shah, P., & Shellhammer, D. (1999). *The Role of Domain Knowledge and Graph Reading Skills in Graph Comprehension.* Paper Presented at the 1999 Meeting of the Society for Applied Research in Memory and Cognition, Boulder, CO.

Shaughnessy, J. M. (2007). Research on statistics' reasoning and learning. *Second Handbook of Research on Mathematics Teaching and Learning*, 957–1009.

Slough, S. W., & McTigue, E. (2013). Development of the Graphical Analysis Protocol (GAP) for Eliciting the Graphical Demands of Science Textbooks. In M. S. Khine (Ed.), *Critical analysis of science textbooks: Evaluating instructional effectiveness* (pp. 17–30). Dordrecht, Heidelberg, New York, London: Springer. https://doi.org/10.1007/978-94-007-4168-3_2

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction*, *13*(2), 191–203. https://doi.org/10.1016/S0959-4752(02)00020-8

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction*, *29*(1), 123–145. https://doi.org/10.1080/07370008.2010.507318

Stieff, M., Scopelitis, S., Lira, M. E., & Desutter, D. (2016). Improving Representational Competence with Concrete Models. *Science Education*, *100*(2), 344–363. https://doi.org/10.1002/sce.21203

Strober, M. H., & Cook, A. (1992). Economics, lies, and videotapes. *The Journal of Economic Education*, *23*(2), 125–151.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312.

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, *68*(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Tairab, H. H., & Khalaf Al-Naqbi, A. K. (2004). How do secondary school science students interpret and construct scientific graphs? *Journal of Biological Education*, *38*(3), 127–132. https://doi.org/10.1080/00219266.2004.9655920

TIMSS (2013). *Trends in International Mathematics and Science Study 2011: Released mathematical items.* Retrieved from https://nces.ed.gov/timss/pdf/TIMSS2011_G8_Math.pdf

Ullrich, M., Schnotz, W., Horz, H., McElvany, N., Schroeder, S., & Baumert, J. (2012). Kognitionspsychologische Aspekte eines Kompetenzmodells zur Bild-Text-Integration [Cognitive psychological aspects of a competence model of picture-text-integration]. *Psychologische Rundschau*, *63*(1), 11–17. https://doi.org/10.1026/0033-3042/a000105

Van Gog, T. (2014). The Signaling (or Cueing) Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 263–278). Cambridge: Cambridge University Press.

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*(2), 95–99. https://doi.org/10.1016/j.learninstruc.2009.02.009

Vitale, J. M., Lai, K., & Linn, M. C. (2015). Taking advantage of automated assessment of student-constructed graphs in science. *Journal of Research in Science Teaching*, *52*(10), 1426–1450. https://doi.org/10.1002/tea.21241

Volterra, V. (1928). Variations and fluctuations of the number of individuals in animal species living together. *ICES Journal of Marine Science*, *3*(1), 3–51.

Vries, E. de (2012). Learning with External Representations. In N. M. Seel (Ed.), *Springer reference. Encyclopedia of the sciences of learning* (pp. 2016–2019). New York: Springer. https://doi.org/10.1007/978-1-4419-1428-6_675

Wallman, K. K. (1993). Enhancing Statistical Literacy: Enriching Our Society. *Journal of the American Statistical Association*, *88*(421), 1–8. https://doi.org/10.1080/01621459.1993.10594283

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3–46.

Westelinck, K. D., Valcke, M., Craene, B. de, & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior*, *21*(4), 555–573. https://doi.org/10.1016/j.chb.2004.10.030

Wheat, D. (2007). The feedback method of teaching macroeconomics: is it effective? *System Dynamics Review*, *23*(4), 391–413.

Winn, W. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustrations (Vol. 1)* (pp. 152–198). Harrisonburg, VA: R. R. Donnelley & Sons.

Wu, H.-K., & Puntambekar, S. (2012). Pedagogical Affordances of Multiple External Representations in Scientific Processes. *Journal of Science Education and Technology*, *21*(6), 754–767. https://doi.org/10.1007/s10956-011-9363-7

2 Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs

Ring, M., Brahm, T., & Randler, C. (2019). Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs. *International Journal of Science Education*, 41(13), 1787-1804. https://doi.org/10.1080/09500693.2019.1640915

Abstract

In science, graphs are used to visualize data, relationships and scientific principles. Basic graph reading operations, i.e. reading single data points, recognizing trends as well as conducting small extrapolations from the data are important skills for pupils and lay the groundwork for a comprehensive understanding of data visualizations across all disciplines. To explore the relationship between question level and item difficulty, we analyzed the graph reading skills of eight-grade pupils in German Grammar Schools (N = 198) with Item Response Theory. In this study, pupils were asked to identify data points, trends and make small extrapolation from realistic descriptive graphs, used as teaching material. Furthermore, we examined the relationship between graph reading skills and motivational/emotional factors as well as academic achievement. Results show that the eight-graders mastered the basic abilities well and – contrary to our expectation – are as well able to extrapolate as they are to read trends or data points. We also find some relationship between graph reading skill and academic achievement and motivational/emotional factors as well as content knowledge.

Keywords: Graphical literacy; graphicacy; graph comprehension; question levels; information extraction; assessment; item response theory

2.1 Introduction

Understanding graphs is essential for participating in a 21st century´s information society. To understand modern data-driven science, it is necessary that pupils are trained in reading, analyzing, constructing and evaluating graphs or representations (Nitz, Ainsworth, Nerdel, & Prechtl, 2014; Yore & Hand, 2010).

In order to capture the processes around pupils' capturing of graphs, different terms are used in scientific literature. Representational competence is used more generally, when the ability to work with different representations (graphs, diagrams, formulas, concept maps etc.) within a specific domain is analyzed (Bergey, Cromley, & Newcombe, 2015; Nitz et al., 2014; Stieff, Scopelitis, Lira, & Desutter, 2016). The focus is here not only on the ability to interpret or extract information but also on integrating information from different representations or translate them from one representation into another. The term graph reading or graphical literacy (also termed graphicacy) is used, when individuals' ability to extract information from graphs 'independent' of a specific content domain is analyzed (Åberg-Bengtsson & Ottosson, 2006). Graph comprehension or graphing ability is mostly used when the focus lies more on graphs that represent domain-specific principles (Friel, Curcio, & Bright, 2001; Lai et al., 2016; Peterman, Cranston, Pryor, & Kermish-Allen, 2015). In this paper, we use the term of graph reading, graph literacy (graphicacy) as well as information extraction interchangeably in order to highlight the focus of the presented study on the ability to extract information from graphs. Other terms are used in order to distinguish between the different concepts. In the broader scope of representational competence, models from multimedia research focus less on single representations and more on the ability to connect information from different representations into a bigger picture and on design principles to improve learning outcomes (Ainsworth, 2014; Scheiter & Eitel, 2015; Seufert, 2003).

Even if the distinction between these research areas is not entirely delineated, most studies that analyses representational competence or examine graphing ability in science work with representations or graphs that represent domain/scientific principles and ask domain-relevant questions (Lai et al., 2016). In contrast, studies on graphical literacy mostly use artificial graphs and ask questions that can be answered with the help of the graph alone (e.g. Åberg-Bengtsson & Ottosson, 2006).

To our knowledge, there are no studies that focus on the ability to extract information of 'authentic' graphical displays that do not visualize domain-specific principles (e.g. prey-predator-relationship in a Lotka–Volterra model), but show descriptive data (e.g. the decline of fish stocks in the sea). However, such data-driven graphs are used by a variety of (science) disciplines in the educational context and are prevalent in media coverage and science communication (Glazer, 2011). For instance, in the context of sustainable development, pupils are demanded to understand whether resources are exploited. This is usually depicted in a data-based graph which can often be found in pupils' textbooks, in newspapers and on the internet.

To analyze the relationship between task characteristics and the ability to read graphs, most studies in science education work with a three-level approach (Friel et al., 2001). A graph task on the first level might ask for a single data point, a graph task on the second level might ask for a trend or an average whereas the third level includes tasks that ask for a small extrapolation or prediction. Whether these question levels are related to pupils' graphicacy is debated: Although there are several studies that show the graduated nature of the three question levels (for an overview see (Friel et al., 2001) – meaning that that level 3 questions are more difficult than level 1 questions, there are also studies questioning the consecutive approach to graph understanding (Lachmayer, 2008). In order to shed light on this debated issue, our first research question for this study asks whether differences in graph reading performance can be explained with the help of different graph reading levels (RQ 1).

Furthermore, pupils' performance is also dependent on their motivation and interest into a topic (Guay et al., 2010). However, this relationship has not been explored for pupils' graphicacy. Consequently, the following study also takes into account several other variables (which are explained in further detail in subsection 2.4). Our second research question (RQ2) aims at providing insights into how graph reading relates to motivational and interest variables.

2.2 Theoretical background and prior research

2.2.1 Graph comprehension models

According to Shah, Freedman, and Vekiri (2005), graphs are '*a unique form of visuospatial depictions that represent quantitative information via an analogy between quantitative scales and visual or spatial dimensions, such as length, colour, or area*' (Shah et al., 2005, p. 428). Lachmayer, Nerdel and Prechtl (2007) make a distinction between diagrams that represent qualitative information, e.g. processes or principles which do not have agreed-upon general reading instructions and graphs, which are quantitative representations such as line plots, bar graphs and pie charts, which have clear labels and reading rules.

Regarding the graph comprehension ability, there are different models that can be divided into groups in relation to their main focus: 'perceptual' models focus more on the perceptual process and the ability to retrieve simple facts from or to make distinctions between different parts of graphs. General process models analyze the 'conceptual' processes: how people draw more general conclusions from graphs, how they perform open-ended tasks and what roles different factors (e.g. content knowledge) play in that process (Shah et al., 2005). Most models within the general process models describe graph comprehension as a three-step process: encoding visual information and identifying relevant elements, relating the visual features to the represented concepts and lastly connecting the information from the graph to the disciplinary context (Shah & Hoeffner, 2002).

For instance, regarding a bar graph that depicts fish stocks in the open sea, the model would assume that a pupil would first identify a bar as a relevant element, then would link this element to the depicted concept: i.e. the share of overfished stocks in a certain year and in the last step might link the information to a relevant question (e.g. based on the data in the graph, how would you expect the overfished stocks to develop?).

Shah et al. (2005) describe an interactive model of graph comprehension and outline five general factors that play a major role in predicting how viewers interpret graphs and how long it takes them to accomplish certain tasks: the display characteristics (type of graph, colour, legend vs. labels...), the complexity of the data (complexity of relationships, number of data points…), the task of the viewer, the viewer’s prior knowledge about the content and the viewer’s graph comprehension skills (Shah et al., 2005).

2.2.2 Task characteristics: Question levels

Regarding the task demands of graph reading, Friel et al. (2001) summarize how different authors structure the different kinds of questions that graphs can be used to answer. They report three levels of graph comprehension which were briefly mentioned above. These three levels are connected to the three processes of the general process models (for an overview of studies using the three-level approach see (Friel et al., 2001). The authors name these three levels after Curcio (1987) (1) Reading the data, (2) Reading between the data, (3) Reading beyond the data. (1) The elementary level focuses on extracting single data from a graph, (2) the intermediate level of understanding is characterized by finding trends and relationships in the data as shown in the graph. (3) Finally, the expert level requires extrapolation, or – in other words – extending, predicting, or inferring from the representation of data to answer questions and/or draw general conclusions (Friel et al., 2001). According to different definitions of graph comprehension, these three levels are operationalized differently in the respective studies. While the first level is mostly seen as 'simple' data extractions such as single data point questions, the second level ranges from questions reading simple data trends (Curcio, 1987) to tasks extrapolating relationships out of complex graphical displays of data (Lai et al., 2016). The third level then involves predictions (Curcio, 1987) as well as answering domain specific/opinion questions with the help of the graph and content knowledge (Lai et al., 2016). Whereas most studies find a positive relationship between different item levels and item difficulty (e.g. Curcio, 1987, Lai et al., 2016) – the higher the level of graph understanding, the more difficult the respective assessment item – there are also studies that cannot find any meaningful differences for the first two levels but can find differences in regards to 'beyond the data' -questions (Lachmayer, 2008). It has been suggested to focus on two types of questions - those that can be answered with 'reading data off a graph' (or a table) – which would be equivalent to level 1 and 2 – and opinion questions which require more evaluation and critical thinking skills as well as background knowledge (Gal, 1998).

This conceptual discussion around task characteristics is related to the question of item types and their influence on graphicacy. One of the more prominent tests regarding graph comprehension – the Test of Graphing in Science (TOGS) developed by McKenzie and Padilla for science pupils from grades seven through twelve uses only multiple-choice items (McKenzie & Padilla, 1986). Berg and Smith´s (1994) research showed that multiple choice items could not reveal the complex thinking that pupils showed when interpreting science graphs and the authors, therefore, argued that free response items are a necessary assessment tool when it comes to interpreting graphs in science (Berg & Smith, 1994). In contrast to Berg and Smith, Åberg-Bengtsson & Ottoson (2006) did not find that the question format has an influence on test performance for reading the data and reading between the data questions (Åberg-Bengtsson & Ottosson, 2006).

2.2.3 Learner characteristics

A person’s ability to derive meaning from graphs as graphical literacy is associated with different learner characteristics. Among these factors are cognitive abilities such as visual and spatial thinking, logical reasoning, or basic numeric abilities (Ludewig, Lambert, Dackermann, Scheiter, & Möller, 2019; Vekiri, 2002). Åberg-Bengtsson & Ottoson (2006) identified different factors that are related to pupil´s graphical literacy skills in secondary education and found a general 'graphicacy-test factor' that was strongly correlated with a mathematic/science factor, an overall school achievement factor. The authors also found a page-specific factor, which they interpreted as content/graph specific difficulties and a weak connection between graphical literacy and a language factor (Åberg-Bengtsson & Ottosson, 2006, pp. 56–57).

2.3 The present study

To contribute to graph literacy research, we aim to better understand graphical literacy as a generic ability for descriptive graphs. In the first research question, we analyze the relationship between task characteristics and pupils' performance:

RQ 1: can differences in graph reading performance be explained with the help of different graph reading levels?

Second, we aim to explore the relationship between information extraction ability with learner characteristics. Some of these relationships – such as the relationship between academic achievement as well as the pupils' language and content knowledge (Åberg-Bengtsson & Ottosson, 2006), - are already known and thus can be used to establish validity of the test. The relationship between graph reading and motivational or interest variables, however, are not yet fully understood and deserve further attention. In the second research question we explore the relationship between graph reading performance and motivational as well as interest variables:

RQ 2: how does graph reading relate to motivational and interest variables?

2.4 Methods and Materials

2.4.1 Participants and data collection

This study was conducted at four different German grammar schools (Gymnasium) in Baden-Württemberg, south-western Germany in January and February 2018 about halfway through the academic year. The pupils were tested during normal school hours at different times and not in the context of a certain subject. Participation was voluntary and a coupling with formal academic achievement was anonymously established via pupils-IDs. In this paper, data from 198 pupils (49% female, 47% male, 4% NA) aged 13-16 years (M=13.3, SD =0.65) in eleven grade 8-classes is reported. Around 90% of the pupils indicated German (or German and another language) as their first language. The same test was used in all classes although the item-sequence did permute systematically in three different versions.

2.4.2 Pupils' prior educational experiences with graphs

Graph reading is part of the German School Curriculum. Within the Grammar School (Gymnasium) in Baden-Württemberg, it is part of the math curriculum in 6th grade, part of the physics curriculum in different grades as well as part of geography (beginning 6th grade) (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2004). Within the disciplines the use of graphs is mostly connected to certain content: e.g. in math graphs are connected to fractions and statistics (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2004, p. 96), whereas in geography graph reading might be taught when pupils learn about climate (graphs) (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2004, p. 240). There is, however, no information available on how intensely or how profoundly graph reading is taught. Although there is an assumed linkage of domain and graphs, we do not know if teachers work on a more "general" graph reading ability (i.e. reading different graphs) or on certain graph types (e.g. climate charts) only. In this study, we were not able to control for pupils' pre-study involvement with graphs but rather aim to assess pupils' graphicacy after certain exposure with graphs.

2.4.3 Test instrument

2.4.3.1 Graphs

In order to ensure authenticity in the material, we used graphs that pupils could encounter in school as well as in newspapers or on the internet. The graphs were chosen from different sources of (teaching) material. The graph-sources are the "Bundeszentrale für politische Bildung" (Federal Agency for Civic Education), "dpa-Globus-Grafiken" (German press agency infographics), "Handelsblatt macht Schule: Infografik" (teaching material supplemented by a German-language Business Newspaper, particularly aimed at schools), "Statista Infografiken" (German online portal and data-base for statistics). Most of the graphs are commonly licensed and were explicitly created to serve as teaching material. Comparable graphs in terms of format and theme by the same sources can be found in textbooks of all domains. For example, in the economic domain, one textbook depicts fishing catches of different countries as well as overfished stock percentages in the context of sustainable development (Biehahn, Jüngling, Machoczek, Michael, & Ottmar, 2018, p. 23).

The chosen graphs do not highlight certain domain-specific relationships but were chosen due to their visualization of descriptive data. The overall topic of the graphs is sustainable development. It is an increasingly relevant topic in science but also a field where data graphs are commonly used as a communication tool. Since we aimed to test for a generic information extraction ability, we decided to use different graph formats (line graphs, bar graphs, pie charts or a mix of these types). Although graphs in learning material might sometimes be accompanied by some explanatory or descriptive narrative in the text, the visualizations were used without any corresponding text (besides the items and the source) in the test instrument. The rationale for this decision is our focus on the graph reading ability.

2.4.3.2 Information extraction items

To address our first research question – is item difficulty is associated with different question levels? – we created items based on the three level approach and compared their difficulty. In the test to assess the pupil´s ability to extract information, 52 dichotomous items were developed for 7 graphs (overview in table 1).

Table 1

Overview of graphs and items

Content	Type of Graph	Number of Items		
		Level 1	Level 2	Level 3
Fish stocks	Mix type (line/bar graph)	3	3	3
Oil reserves	Mix type (bar graph/pie chart)	3	3	3
CO_2 Emissions	Multiple line graph (2 lines)	2	3	3
Malnutrition	Multiple line graph (8 lines)	3	4	3
Water consumption	Bar graph (horizontal)	2	0	3
Population development	Mix type (line/bar graph)	3	3	0
Renewable energy	Mix type (bar graph/pie chart)	3	0	2

For line graphs (or mixed types with line graph elements), we developed items that asked for single data points as level 1 ('read the data'), items that asked for trends as level 2 ('read between the data') and items regarding predictions and evaluations as level 3 ('read beyond the data'). For pie charts, the items regarding single data points (comparisons) are level 1 ('read the data'), and evaluations are level 3 ('read beyond the data'). In order to make the levels comparable and to achieve a test to be easy to use in classrooms, we used a dichotomous format. Even if test takers can guess with a higher chance, the format still proves to be advantageous to portray varying degrees of comprehension (Brassil & Couch, 2019). Two example graphs with respective example items of every level can be seen in Figure 1 and 2.

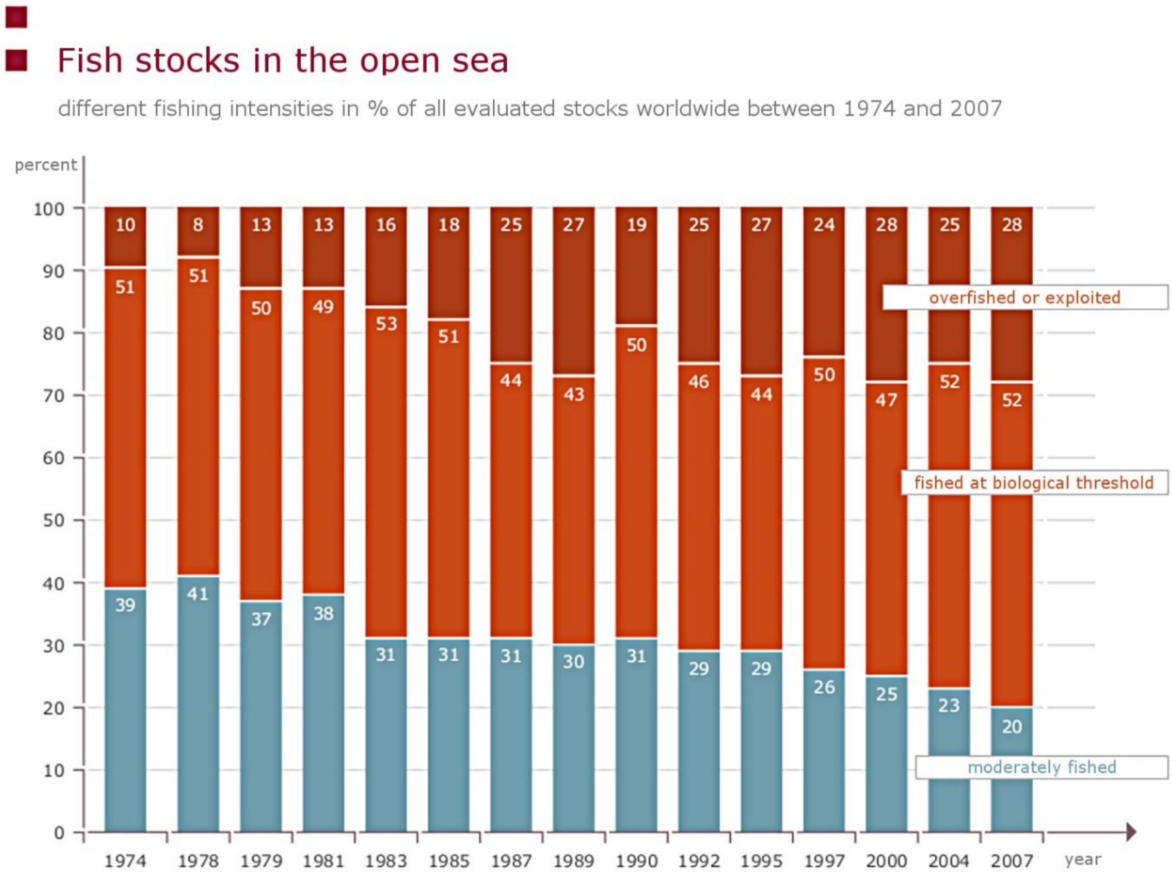


Figure 1. Graph example: Fish stocks in the open sea (Bundeszentrale für politische Bildung, 2017 / Food and Agriculture Organization of the United Nations (FAO): The State of World Fisheries and Aquaculture, license: cc by-nc-nd/3.0/de/ - translated and used with permission by the authors)

Items of different levels	True	False
(1) The share of fish stocks that was fished at biological threshold in 1978 is 51%.	☐	☐
(2) The share of fish stocks that was overfished or exploited has mostly decreased over the whole period.	☐	☐
(3) Based on the information in the graph, it is realistic to assume that the share of fish stocks that will be fished at biological threshold in 2010 will be close to the same share in 1974.	☐	☐

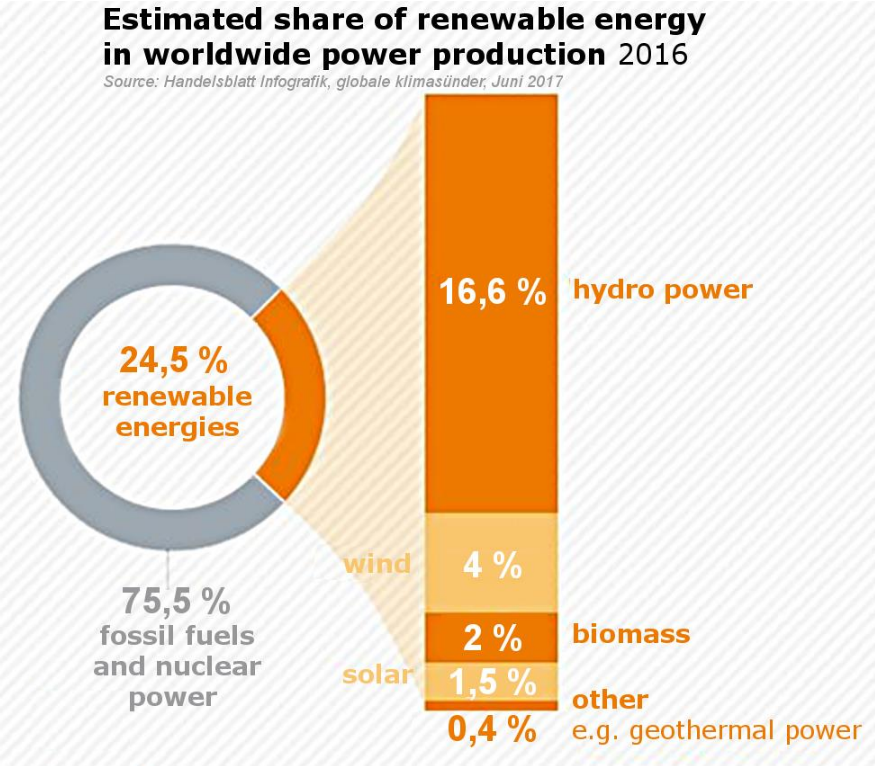


Figure 2. Graph example: Renewable energy (Handelsblatt, 2017, Grafik des Tages "Globale Klimasünder" [Graph of the day: 'Global climate sinners']; Handelsblatt Nr.111; 12.06.2017, p. 24- translated and used with permission by the authors)

Item examples for renewable energy graph	True	False
(1) The share of energy produced by wind energy sources in worldwide power production is 4%.	☐	☐
(3) Imagine it is no longer possible to produce energy from biomass. Based on the graph it is realistic to assume, that the consequences are similar to a loss of possibility to produce energy from fossil and nuclear power.	☐	☐

2.4.3.3 Motivational and interest factors

To establish reliability and convergent validity of the graph reading instrument, we also surveyed gender, math grade, German grade and first language. Different motivational and interest factors were assessed to analyze their relationship to graphical literacy (second research question).

Based on the assumption that content knowledge as well as previous exposure to graphs might have an effect on the graphing ability, we asked the pupils whether they had already encountered the theme of the graphs (sustainable development) whether they are interested in the topic and used a knowledge test about sustainable development. For all three scales (previous engagement, interest and knowledge about sustainable development), we used established items from another test (Michaelis, 2017) and adapted them for our purpose. Due to minimal changes, we did not revalidate the items prior this study. The scale for previous engagement with sustainable development contains 7 items and is Likert-scaled from 1-4. An example item is: 'How often in the last year did you engage yourself with the following topic: renewable energy' (Cronbach’s alpha: 0.66, Mean score: 2.31 Min: 1.0, Max: 3.43). The scale for interest in sustainable development contains four items is Likert-scaled from 1-4. Example item: 'How much are you interested in the following topic: environment and nature?' (Cronbach’s alpha: 0.64, Mean score: 2.66 Min: 1,0, Max: 4,0). The knowledge test consists of six multiple choice items. An example item is: Which of the following objectives could be assigned to sustainable development? The number of all correct items divided by the number of all items was used as a measure (percent score, mean score: 0.62, Min: 0.00, Max: 1.00).

We also assessed subject motivation adapted from Prenzel and Drechsel (1996) in German. Motivation is highly relevant for academic achievement and different pupils might have different levels of motivation for different subjects (Guay et al., 2010). We assessed motivation in subjects, where pupils might encounter similar graphs (geography and biology). The used test contains seven items from two dimensions and is Likert-scaled from 1-4. A high score represents a high motivation. An example item for intrinsic motivation is: In school (think of classes in biology or geography), I really had fun learning. (Cronbach’s alpha: 0.57, Mean score 2.74, Min:1.0, Max: 4.0). An example item for extrinsic motivation is: In school (think of classes in biology or geography), I only participated in class when the teacher asked me to. To achieve higher reliability, one of the extrinsic motivation items was removed (Cronbach’s alpha: 0.69, Mean score: 1.78, Min: 1.0, Max: 3.67).

Due to the voluntary participation and the fact that there are no direct personal consequences associated with their scores, test-motivation/effort might vary and play a significant role in performance (Barry, Horst, Finney, Brown, & Kopp, 2010). The test motivation item was adapted from PISA in German (Ramm et al., 2006). The item asks the participants to rate the effort with which they took the test on a scale from 1 to 10. (Mean Score: 7.55, Min: 1.00, Max: 10.00). Together with the items for graphicacy, the test consisted of 77 items.

2.4.4 Data analysis procedure

To analyze the ability to extract information, Item Response Theory (IRT) was used. IRT is based on the assumption that the probability to solve an item is a function of the test-taker ability as well as the item difficulty. With the help of IRT it is possible to estimate abilities and item difficulties on the same logit scale, ranging from -4.0 (lower abilities/easy items) to +4.0 (higher abilities/difficult items) (Embretson & Reise, 2013).

In order to link the data and the assessment framework (Kuo, Wu, Jen, & Hsu, 2015), our goal in using IRT was twofold: for Research Question 1, we grouped items according to their characteristics (or to stay in the terminology of graph comprehension: according to their question levels) and compared the mean difficulties of these groups. If items that ask for predictions or extrapolations (level 3) are harder for pupils than items that ask for single data points (level 1) than this would show in the respective item difficulties. If question levels (as operationalized in this study) are not relevant for descriptive authentic graphs than we might expect a random distribution of the difficulty for the respective items. For Research Question 2, we used the IRT scaled ability estimates of the test takers and related these to other measures, namely to motivational and interest factors in order to find out whether pupils' motivation or interest are associated with their test performance.

2.4.4.1 Model fit, item exclusion and Wright-Map

A one-dimensional Rasch model was used to estimate the latent abilities and item difficulties, it established the best fit based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) values in comparison with other models. For abilities, we used weighted likelihood estimations (WLE), whereas marginal maximum likelihood estimates (MMLE) were used for the item difficulties. All estimates were done with the TAM-package in the R Studio Software (R Core Team, 2020; Robitzsch, Kiefer, & Wu, 2018). Regarding the whole test instrument, we also computed standard error and test information in relationship to skill levels.

When analyzing items with IRT in science education, researches may exclude items based on low discrimination, model fit or redundancy (Neumann, Neumann, & Nehm, 2010). We did not exclude items based on their discrimination or redundancy since we focus on the question of differences between the three levels, and thus do not want to exclude items that fit the assessment framework and are meant to check whether pupils mastered the associated skill. Premised on the model, however, items were analyzed based on outfit (unexpected answers for items far from a person's ability) and infit (unexpected answers for items close to a person's ability). We excluded 10 items that showed weighted fit mean squares lower than 0.7 or higher than 1.3. Based on the items with satisfactory properties a Wright-map was created to show the distribution of pupil ability and item difficulty. We then compared the estimated difficulties of the different item groups that asked for data points (now: 16 items), data trends (now: 12 items) or extrapolations/evaluations (now: 14 items). The means of all groups were compared with an analysis of variance (ANOVA).

2.4.4.2 Reliabilities, subgroup performance and learner characteristics

The reliabilities of the other measures were calculated using Cronbach's Alpha. Criterion validity was analyzed by inspecting correlations and subgroup differences between the information extraction ability estimated by the IRT model and the gender, math grade, German grade and first language. Due to different variances and sample sizes, the subgroup differences were calculated with a Welch`s t-test. Furthermore, we analyzed the relationship by inspection correlations between the results of the graphical literacy test (estimated abilities), with interest in sustainable development, prior engagement with sustainable development as well as knowledge about sustainable development one the one hand, and motivational and test motivation /effort on the other hand.

2.5 Results

2.5.1 Graph extraction ability

The Wright-Map (Figure 3) shows the estimated ability distribution on the left-hand and the estimated item difficulty on the right-hand side. Whereas the pupil's estimated ability to extract information ranges from -3 to +3 on the logit scale, most of the item difficulties are between -4 and -1 on the same scale. To provide accurate measures, it would be ideal, that the distribution of the item difficulty matches the distribution of ability estimates. This, however, is not the case – the Wright-Map clearly shows a lack of difficult items in comparison to the estimated ability. In the Wright-Map the items are ordered according to the information extracted /their question levels (Data points: 1-16, Data trends: 17-28, Extrapolations: 29-42).

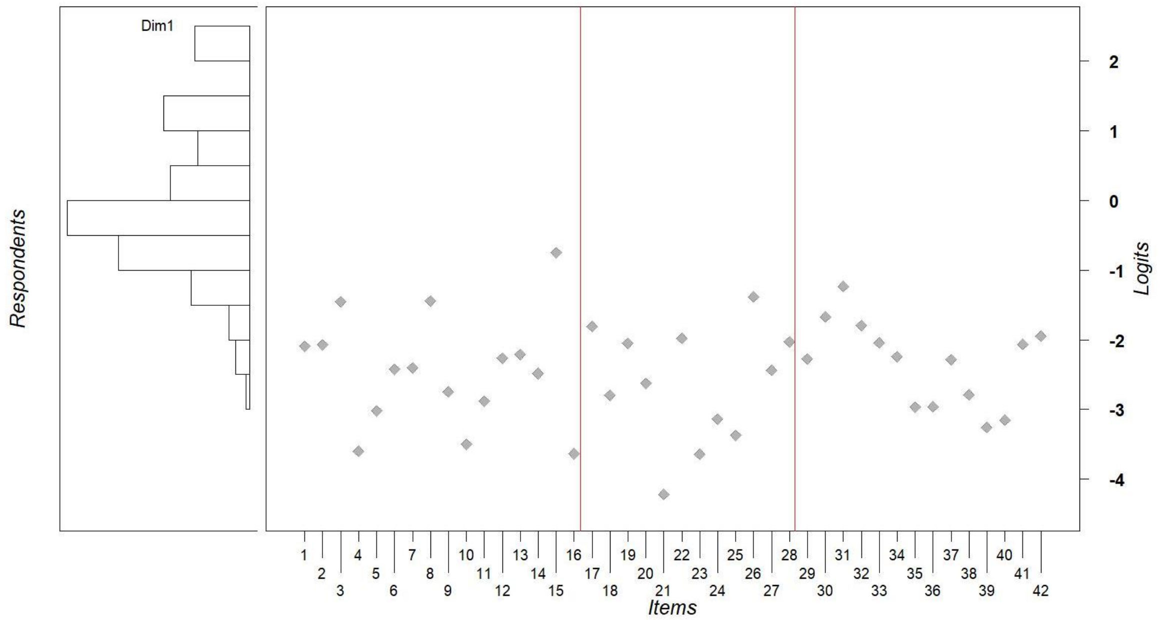


Figure 3. Wright-Map: Item difficulties and ability estimates (Items 1-16 'Read the data', 17-28 'Read between the data', 29-42 'Read beyond the data', own illustration)

Figure 4 shows the test information and standard error. Whereas the test information for lower ability pupils might be accurate, the test is not able to differentiate between medium to higher levels of information extraction ability. Due to the lack of difficult items in comparison to the estimated difficulty, we have no viable test information for higher skill levels.

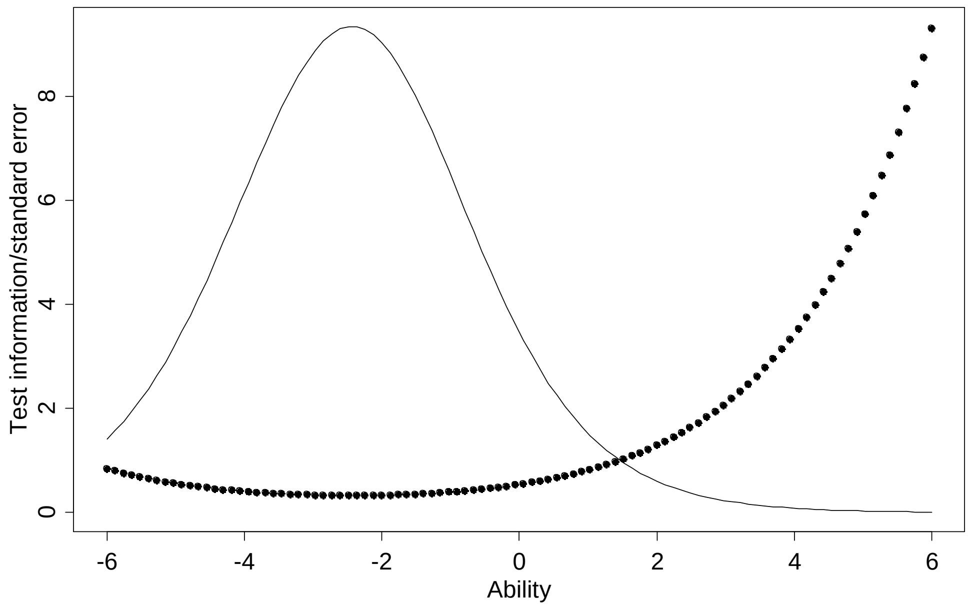


Figure 4. Test information and standard error in relation to ability (own illustration)

2.5.2 First Research Question: relationship between item difficulty and item characteristics

The estimates for item difficulty range from -4.22 to -0.75 with a mean of -2.46. Figure 5 (below) shows the item difficulty means of all level 1 (read the data), level 2 (read between the data) and level 3 (read beyond the data) items in comparison. An ANOVA shows that there are no systematic differences between the three item groups (F = 0.481 < P = 0.622).

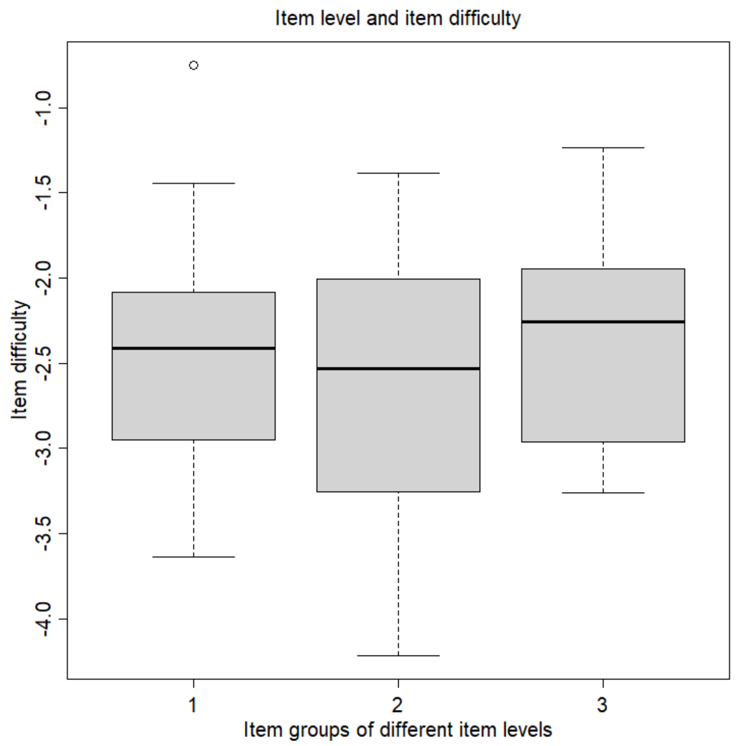


Figure 5. Item difficulty of item groups of different question levels (own illustration)

Since different kinds of graphs were used in the study, we examined whether there are differences regarding the information extraction ability in order to find out whether graphs need to be excluded from our further analysis. Indeed, some differences could be detected between different graphs. It appears that the graphs on water consumption and malnutrition were slightly easier than the other five graphs (ANOVA results: F=2.448 > p =0.0441). Overall, however, the graphs are comparable and thus, were all used for further analyses.

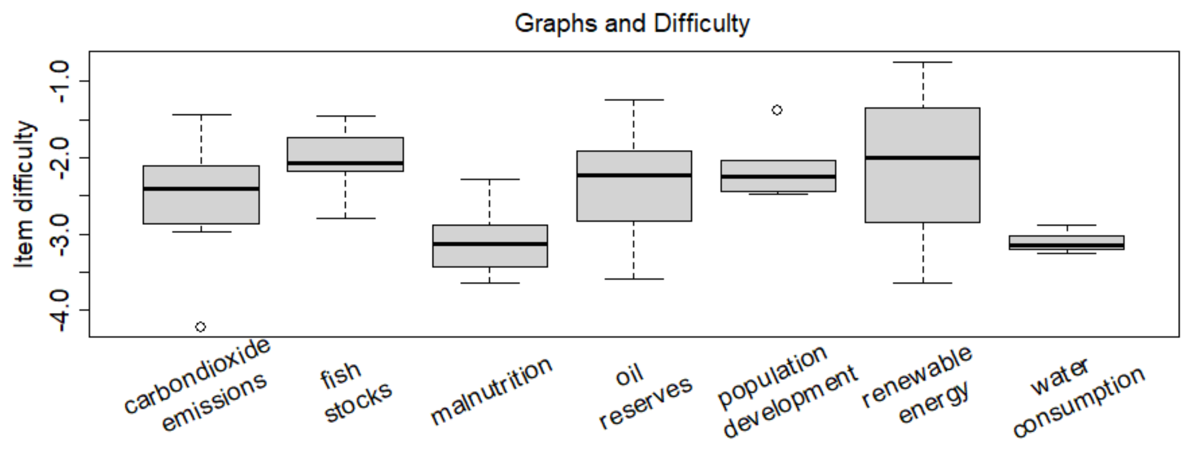


Figure 6. Item difficulty of different graphs in comparison (own illustration)

2.5.3 Second Research Question: Motivation, interest and subgroup performances

To assess criterion validity, correlations of the estimated abilities and convergent validity measures were analyzed. Results (see Table 2) show that there are significant correlations with pupils' performance in German, math as well as prior engagement with sustainable development and knowledge about sustainable development. We find a negative correlation between extrinsic motivation and graphicacy but no meaningful correlation between the extraction ability and age, intrinsic motivation, test motivation or interest in sustainable development.

Table 2

Correlations of control variables and graph ability (Kendall's Tau)

Construct	Correlation with Ability estimate	p-value
Performance in German class	0.23	$p<.001$
Performance in math class	0.22	$p<.001$
Extrinsic motivation	-0.18	$p<.001$
Knowledge about sustainable development	0.19	$p<.001$
Prior engagement with sustainable development	0.13	$p<.05$

As can be seen from table 3 (below), the only meaningful subgroup differences were between pupils who reported German as first language and those who did not. We cannot report gender differences in regard to graphical literacy.

Table 3

Subgroup performance parameters and differences (Welch's T-Test)

Subgroup	Parameters		Differences	
	n	mean	p-value	t-value
German as first language	176	0.06	p<.01	3.08
Not-German as first language	19	-0.64		

2.6 Discussion

2.6.1 Research question 1

The items were constructed to find out whether task characteristics (question levels) matter for item difficulty. Although both the test and an IRT analysis have their limitations (see below), the paper includes some contributions to graphicacy research: The pupils in this sample are rather competent in solving graph items and are quite capable of reading data points, identifying trends and are also able to extrapolate, i.e. make predictions or evaluate new information with the help of the graph. There seems to be no systematic differences between item difficulties for item groups of different question levels. In other words, what exactly an item asked for – a point, a trend or an extrapolation – is not systematically associated with its difficulty in this sample.

Regarding the question level of extrapolations and evaluations, the results clearly differ from most of the previous research (Curcio, 1987; Lai et al., 2016). There could be a number of reasons for these discrepancies based on methodology and research design. The most obvious reason could be the operationalization of the three levels: All our items are dichotomous and solvable with the help of the graph alone and might be closer to reading questions than to comprehension questions (and thus, could be argued, all belong to level 1). However, if level three items are not solvable without content knowledge, it is difficult to isolate a generic information extraction ability. The second explanation could be the focus on performance as item-solving without considering time on task, where hypothetically different cognitive steps behind the three levels might have shown a greater variance in the dependent variable. To further analyze the relationship between task characteristics and (efficient) performance, it might be fruitful to either measure time on task or to conduct think-aloud studies, in which pupils are asked to explicate the difference as well as the difficulty of the three level questions.

Another influence could be the graph complexity. Whereas the graphs in this test only showed descriptive data in an authentic data-oriented format, other graph researchers (Lachmayer et al., 2007; Lai et al., 2016; Nitz et al., 2014) work with graphs/representations which visualize domain-specific (science) principles with potentially more complex relationships between data. Even without using representations that visualize important domain principles, it would be possible to use a wider array of graphs. Within the scientific and mainstream discussion about sustainable development, there are graphs that are more complex visualizations of data (Ritchie & Roser, 2019). On a theoretical level, the results of this research support the idea that an operationalization on the basis of task characteristics alone might not be meaningful. If there are no differences with respect to different kinds of 'reading the graph' questions, it might not be the valid research framework. Or in other words: A too narrow focus. From our point of view, these results could be interpreted as an empirical argument to Gal's idea that is necessary to focus on two types of questions - those that can be answered with 'reading data off a graph' (or a table) and 'opinion questions' which require more evaluation and critical thinking skills as well as background knowledge (Gal, 1998).

2.6.2 Research Question 2

A major restriction of our study is that the test is – due to a lack of difficult items – not able to distinguish between higher levels of information extraction skills. All medium to higher ability estimates have no empirical basis and a high standard error. The relationships between the estimates of the IRT model and the different factors should therefore be interpreted with caution.

However, we find most of the connections that are already established in graphicacy-research (Åberg-Bengtsson & Ottosson, 2006; Lai et al., 2016) or German school assessments (OECD, 2016) which may be a cautionary proof of some validity.

There are significant correlations between pupils' estimated ability and performance in German/math class, prior engagement with sustainable development as well as knowledge about sustainable development.

A relationship with performance in German class was expected due to the high portion of text in the items and the literacy character of the test instrument. Since graphs are prominent in the math curriculum and math grade tends to correlate with similar cognitive operations (OECD, 2016), a correlation with math performance was also expected.

The relationship between prior engagement with sustainable development and performance in the graph reading test is an interesting finding. We cannot be sure, however, whether pupils profited from exposure to the content (or theme) of the graphs or the use of descriptive graphs (and thus a training of graphicacy) within the lessons involving sustainable development or both. The relationship between content knowledge and the graphicacy test could be a further hint, that exposure to the theme of the graphs might have a beneficial effect for graph reading ability. However, correlation does not imply causality and a second possible explanation for these correlations could be a common denominator, e.g. an overall school achievement factor as described by Åberg-Bengtsson and Ottosson (2006). Further research could expand on this question, with an experimental design and analyzing graphicacy for a content knowledge treatment.

We do not find the expected correlation between age and graphicacy. This might be due to the fact that the testing was in the same grade level and the few pupils that were older than the mean, either were enrolled in school later or had to repeat a year. Overall, the variance in the age variable is very small.

While there are no significant differences between male and female pupils, the test shows differences between pupils who indicated German (or German and another language) as their first language and those who did not. This effect is also reported in most other German education assessment studies (OECD, 2016) as well as in graph research (Lai et al., 2016).

With respect to motivational and interest factors, our findings are inconclusive: we cannot report a correlation between the graphicacy and test motivation, intrinsic motivation or interest in the broader theme. There is, however, a negative correlation with extrinsic motivation. Further research with more reliable scales for motivational constructs and a more comprehensive test instrument for graphicacy could address this question and determine whether the reported findings are systematic or simply data noise.

Although the sample is large enough for the purpose of this study, it is not a representative sample of all eighth-graders due to the focus on the German Grammar School. The implications of this study are also limited by the selection of certain graphs with rather descriptive relationships that can be found in authentic teaching material. All in all, the generalizability of the results is subject to certain limitations.

2.6.3 Conclusion

In this paper, items were developed to assess pupils’ ability to read data-oriented graphs in the context of sustainable development. The aim of this paper was to operationalize graphicacy very clearly and to check whether there is a three-level structure that can be explained through certain item characteristics (and independently of graph type or content). An in-depth analysis with item response theory suggests that the grade-eight-pupils in this sample are quite capable of reading data points, identifying trends and are also able to extrapolate to a certain extent. The study was not able to find a three-level structure based on the relationships between the kind of information an item asked for and the item difficulty.

With this study, we have established a baseline – results show that pupils can use the basic graph operators (equally) well. However, teachers often report challenges when working with graphs or representations in authentic classroom settings. If these challenges do not arise from a lack of graph reading ability, future research could further look into the reasons for such challenges. Our study shows that pupils can use basic graph reading operators but it might be worthwhile to analyze their skills regarding more complex (statistical) concepts which are relevant for data literacy (e.g. variance, standard deviation, correlations…). This could be done with more complex graphs and in the same testing format.

A second relevant open question could be pupils' ability to make sense of a graph in a meaningful way within a discipline. Of course, a science teacher is not solely interested in whether pupils might mark pre-given statements as right or wrong but rather in the pupils' ability to analyze a graph, identify the relevant features and to critically discuss the value and the implication of the graph for the question at hand.

Accordingly, the following recommendations can be made for future research:

Focus on how graphs (or external representations in general) are used in authentic learning settings. In science, graphs often are embedded in the scientific process (in order to visually show data and to prove or disprove certain hypothesis). If identifying points, trends or extrapolations is not difficult for 8^{th} graders (as this study suggest) it might be interesting to focus on the usage of graphs as reasoning and explanation tools.

Identify strategies that pupils might have when they encounter graphs in learning material – our study suggests that pupils can read graphs well. However, the closed item format does not show how pupils proceed when reading graphs by themselves. More open ended questions as well as thinking-aloud studies might be beneficial in learning more what strategy pupils use, where pupils show difficulties and how graphical literacy can be further developed.

2.7 References

Åberg-Bengtsson, L., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching, 43*(1), 43–62. https://doi.org/10.1002/tea.20087

Ainsworth, S. (2014). The Multiple Representation Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 464–486). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.024

Barry, C. L., Horst, S. J., Finney, S. J., Brown, A. R., & Kopp, J. P. (2010). Do Examinees Have Similar Test-Taking Effort? A High-Stakes Question for Low-Stakes Testing. *International Journal of Testing, 10*(4), 342–363. https://doi.org/10.1080/15305058.2010.508569

Berg, C. A., & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments. *Science Education, 78*(6), 527–554. https://doi.org/10.1002/sce.3730780602

Bergey, B. W., Cromley, J. G., & Newcombe, N. S. (2015). Teaching High School Biology Students to Coordinate Text and Diagrams: Relations with Transfer, Effort, and Spatial Skill. *International Journal of Science Education, 37*(15), 2476–2502. https://doi.org/10.1080/09500693.2015.1082672

Biehahn, M., Jüngling, H., Machoczek, M., Michael, P., & Ottmar, C. (2018). *Auer Wirtschaft - Berufs- und Studienorientierung. Ausgabe Baden-Württemberg: Schülerbuch Klasse 8-10 [Auer Economics, Vocational and Study orientation, Textbook 8th-10th grade]* (1. Auflage). Stuttgart: Klett.

Brassil, C. E., & Couch, B. A. (2019). Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison. *International Journal of STEM Education, 6*(1), 1289. https://doi.org/10.1186/s40594-019-0169-0

Bundeszentrale für politische Bildung (2017). *Fischbestände auf offener See [Fish stocks in the open sea]*. Retrieved from http://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52733/fischbestaende

Curcio, F. R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, *18*(5), 382–393. https://doi.org/10.2307/749086

Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory for Psychologists*. *Multivariate applications series*. Hoboken: Taylor and Francis.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education*, *32*(2), 124–158. https://doi.org/10.2307/749671

Gal, I. (1998). Assessing statistical knowledge as it relates to students' interpretation of data. In S. P. Lajoie (Ed.), *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12* (275–295). Mahwah, NJ.: Lawrence Erlbaum Ass.

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, *47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H. W., Larose, S., & Boivin, M. (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *The British Journal of Educational Psychology*, *80*(4), 711–735. https://doi.org/10.1348/000709910X499084

Handelsblatt (2017). *Infografik des Tages: Globale Klimasünder [Graph of the day: "Global climate sinners"]*. Retrieved from https://www.handelsblattmachtschule.de/fileadmin/PDF/Sonstiges/Infografik-pdfs/Klimasuender_Seiten_24_25_Handelsblatt_2017-06-12.pdf

Kuo, C.-Y., Wu, H.-K., Jen, T.-H., & Hsu, Y.-S. (2015). Development and Validation of a Multimedia-based Assessment of Scientific Inquiry Abilities. *International Journal of Science Education*, *37*(14), 2326–2357. https://doi.org/10.1080/09500693.2015.1078521

Lachmayer, S. (2008). *Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht [Development and validation of a graph competence model for biology education]* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. Retrieved from https://d-nb.info/1019667389/34

Lachmayer, S., Nerdel, C., & Prechtl, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht [Modelling of cognitive abilities regarding the handling of graphs in science education]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *13*, 161–180.

Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R., & Linn, M. C. (2016). Measuring Graph Comprehension, Critique, and Construction in Science. *Journal of Science Education and Technology*, *25*(4), 665–681. https://doi.org/10.1007/s10956-016-9621-9

Ludewig, U., Lambert, K., Dackermann, T., Scheiter, K., & Möller, K. (2019). Influences of basic numerical abilities on graph reading performance. *Psychological Research*, 1–13. https://doi.org/10.1007/s00426-019-01144-y

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (togs). *Journal of Research in Science Teaching*, *23*(7), 571–579. https://doi.org/10.1002/tea.3660230702

Michaelis, C. (2017). *Kompetenzentwicklung zum nachhaltigen Wirtschaften: Eine Längsschnittstudie in der kaufmännischen Ausbildung. [Competence Development Toward Sustainable Operational Management: A Longitudinal Study in Vocational Educational Training]* [Doctoral dissertation, Georg-August-Universität Göttingen]. Frankfurt a.M, Germany: Peter Lang.

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2004). Bildungsplan 2004: Allgemeinbildendes Gymnasium. [Curriculum 2004] Retrieved from http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/Bildungsplaene/Bildungsplaene-2004/Bildungsstandards/Gymnasium_Bildungsplan_Gesamt.pdf

Neumann, I., Neumann, K., & Nehm, R. (2010). Evaluating Instrument Quality in Science Education: Rasch-based analyses of a Nature of Science test. *International Journal of Science Education*, *33*(10), 1373–1405. https://doi.org/10.1080/09500693.2010.511297

Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction*, *31*, 13–22. https://doi.org/10.1016/j.learninstruc.2013.12.003

OECD (2016). *PISA 2015 Ergebnisse (Band I): Exzellenz und Chancengerechtigkeit in der Bildung [PISA 2015 Results (Volume I) Excellence and Equity in Education]*. Bielefeld, Germany: Bertelsmann Verlag. Retrieved from http://dx.doi.org/10.1787/9789264267879-de

Peterman, K., Cranston, K. A., Pryor, M., & Kermish-Allen, R. (2015). Measuring Primary Students’ Graph Interpretation Skills Via a Performance Assessment: A case study in instrument development. *International Journal of Science Education*, *37*(17), 2787–2808. https://doi.org/10.1080/09500693.2015.1105399

Prenzel, M., & Drechsel, B. (1996). Ein Jahr kaufmännische Erstausbildung: Veränderungen in Lernmotivation und Interesse [One year of vocational educational training: changes in learning motivation and interest]. *Unterrichtswissenschaft*, *24*, 217–234.

R Core Team (2020). *A language and environment for statistical computing (Version 3.4.4) [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M.; Pekrum, R., Rolff, H.-G., Rost, J., & Schiefele, U. (2006). *PISA 2003: Dokumentation der Erhebungsinstrumente [PISA 2003: Scale documentary]*. Münster, Germany: Waxmann.

Ritchie, H., & Roser, M. (2019). *CO_2 and other Greenhouse Gas Emissions*. Retrieved from https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions

Robitzsch, A., Kiefer, T., & Wu, M. (2018). *TAM: Test analysis modules.: R package version 2.13-15*. Retrieved from https://CRAN.R-project.org/package=TAM

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*(2), 227–237. https://doi.org/10.1016/S0959-4752(02)00022-1

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The Comprehension of Quantitative Information in Graphical Displays. In P. Shah & A. Miyake (Eds.), *Cambridge Handbooks in Psychology*. *The Cambridge Handbook of Visuospatial Thinking* (pp. 426–476). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.012

Shah, P., & Hoeffner, J. (2002). Review of Graph Comprehension Research: Implications for Instructions. *Educational Psychology Review*, *14*(1), 47–69. https://doi.org/10.1023/A:1013180410169

Stieff, M., Scopelitis, S., Lira, M. E., & Desutter, D. (2016). Improving Representational Competence with Concrete Models. *Science Education*, *100*(2), 344–363. https://doi.org/10.1002/sce.21203

Vekiri, I. (2002). What Is the Value of Graphical Displays in Learning? *Educational Psychology Review*, *14*(3), 261–312. https://doi.org/10.1023/A:1016064429161

Yore, L. D., & Hand, B. (2010). Epilogue: Plotting a Research Agenda for Multiple Representations, Multiple Modality, and Multimodal Representational Competency. *Research in Science Education*, *40*(1), 93–101. https://doi.org/10.1007/s11165-009-9160-y

3 Logical pictures in secondary economic education: textbook analysis and teacher perception

Ring, M., & Brahm, T. (in press). Logical pictures in secondary economic education: textbook analysis and teacher perception. *RISTAL. Research in Subject-matter Teaching and Learning*.

Abstract

Logical pictures, such as graphs and charts are an important part of instruction, not only in economic education. Learning with these logical pictures might be beneficial under appropriate conditions, however, domain-specific and visualization-specific challenges might impede learning. In this paper, we study the use of logical pictures in secondary economic education learning material and in economics teaching. In a mixed-method approach, we first analyze 450 logical pictures and propose a category system which distinguishes between the form of a logical picture as well as its domain-specificity. In a second step, we conducted teacher interviews with economic teachers. Results show that logical pictures are used frequently in textbooks, with graphs occurring more often than charts. The interview findings support the relevance of graphs and charts for instruction and provide information about the necessary student abilities and their challenges when working with different logical pictures in economic education from the teacher’s perspective.

Keywords: visual representations; logical pictures; graphs; charts; diagrams; secondary economic education, Germany

3.1 Introduction

In social sciences, logical pictures such as graphs and charts are used to visualize and communicate about data, ideas and systems. Accordingly, they are a ubiquitous part of instruction and experts regard visualizations as an important explanation tool for economics (Schopf, Raso, & Kahr, 2019). In higher education, some authors even argue that graphs are more important than algebra or calculus for teaching and learning economics (Hey, 2005).

For the use of logical pictures in learning settings, the question *which* logical pictures are used and *how* they are used is highly relevant for multiple reasons. First, different formats of representations can influence mental model building (Schnotz & Kürschner, 2008) and thus affect how students understand economic concepts such as price building (Jägerskog, 2020). Furthermore, different formats come with their own set of necessary abilities and learner challenges, for example missing graphical literacy for graphs and maps (Åberg-Bengtsson & Ottosson, 2006) or difficulties with statistical concepts for histograms (Boels, Bakker, van Dooren, & Drijvers, 2019). When comparing between domains, different logical pictures are used, (e.g. maps in geography, equilibrium graphs in economics) and even the same logical picture might be use differently, as different information is relevant for the domain question (Cook, 2011). To sum up: the use in classrooms, the tasks as well as the challenges of logical pictures for learners are highly specific for a certain domain (Ainsworth, 2006) and for the respective representations (Schnotz & Kürschner, 2008).

For economic education, most research discusses the use of graphs and charts in higher education (e.g. Cohn, Cohn, Balch, & Bradley, 2001; M. Davies, 2011), but little research has been conducted at secondary level (e.g. Aprea & Bayer, 2010; Jägerskog, 2020).

Although the understanding of logical pictures depends on their form and their use in teaching, few studies have focused on what kind of logical pictures are used in learning material or how economic teachers use logical pictures in their teaching. Thus, the objective of this study is to investigate the use of logical pictures in secondary economic education in an exploratory approach.

3.2 Literature review

3.2.1 Logical pictures as graphical representations

Graphical representations in general are illustrations that are used in learning material and, thus, encompass pictures, drawings or caricatures as well as graphs, charts or diagrams. There are different taxonomies of graphical representations in learning material that differ in their scope (i.e. what kind of representations they focus on), their terminology (what different representations are called) and their criteria for classification (form vs. function). Although there are some similarities, no classification is universally accepted yet (Ainsworth, 2006).

Regarding *scope*, some classifications are focused only on specific types of representations (and, thus, differentiate, for example, between graphs, charts and diagrams, see Winn, 1987); others classify different representations on a higher level and therefore distinguish between text, realistic pictures and logical pictures (e.g. Schnotz, 2001). Furthermore, the various classifications use different *terminology*, a "chart" in one classification, might be a "graph", a "diagram", an "infographic" or a "visual representation" in another framework (e.g. Harris, 2000; Kosslyn, 1999). From a literacy perspective, researchers also use the term discontinuous text, for example, when describing the use of visualizations in geography (e.g. Huber & Stallhofer, 2010). Lastly, most scholars use structural-form characteristics to categorize (e.g. Winn, 1987) whereas others use the relationship between the representation and the represented object (e.g. Schnotz, 2001) or distinguish pictorial elements in learning material according to the function they serve when they are combined with text (e.g. Levin, Anglin, & Carney, 1987).

In this paper, we use the classification of Schnotz (2001) to categorize different visual representations based on the similarity of the represented object and the representation and to clarify our focus on logical pictures in economic education (see Figure 1).

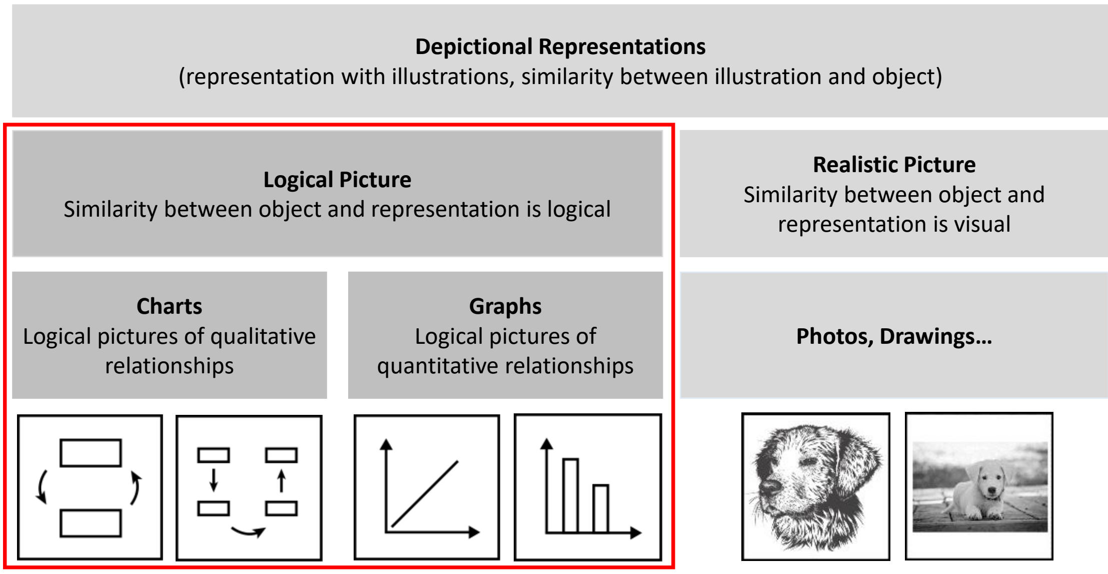


Figure 1. Classification of depictional representations

Own illustration based on the classification by Schnotz, 2001; pictures from pixabay were used as realistic representations: User Gorkhs (2020) [Vector drawing of a dog], pixabay, https://pixabay.com/de/vectors/hund-tier-haustier-h%C3%BCndchen-1728494/, User 3194556 [Picture of a dog], pixabay, https://pixabay.com/de/photos/hund-haustier-tier-niedlich-wei%C3%9F-1903313/

A “depictional representation” describes a real object with a visual-graphic representation. Schnotz (2001) further distinguishes between pictures, where the relation between the object and the representation is visible (e.g. a photograph) and charts and graphs (=logical pictures), where the relation between the object and the representation is logical (Schnotz, 2001). Based on this terminology, we use the term “logical pictures” to refer to representations where the relationship between the represented object and the representation is logical. Within the category of logical pictures, the term “charts” is used to describe visualizations of relationships between distinct objects (hierarchy, process flow, structure…). In charts, the relationships are often displayed with lines and arrows, which are interpreted differently depending on the context (e.g. as “part-of”, “consequence of”, “before-after” etc.). In comparison, we refer to “graphs” as logical pictures of quantitative relationships (bar graphs, line graphs…). For graphs, spatial distances can be meaningfully interpreted as differences between represented objects or relationships (e.g. if one slice of a pie graph is bigger than another, the number it represents is also bigger by the same percentage); as a result, they normally have clear reading rules, i.e. compared to charts, where the meaning of an arrow can change between charts, the height of a bar in a bar graph can be interpreted in the same way in different bar graphs.

3.2.2 Representational competence and domain knowledge

In science education, rather specific terms like graph comprehension (Lai et al., 2016; Peterman, Cranston, Pryor, & Kermish-Allen, 2015; Shah, Freedman, & Vekiri, 2005) or graph literacy/graphicacy (Åberg-Bengtsson & Ottosson, 2006; Galesic & Garcia-Retamero, 2011; Ring, Brahm, & Randler, 2019; Roberts et al., 2013) are used to describe the ability to work with certain kinds of representations. In contrast, the broader set of knowledge and skills to make sense of different representations is referred to as representational competence (Gebre & Polman, 2016; Kozma & Russell, 1997; Nitz, Ainsworth, Nerdel, & Prechtl, 2014; Stieff, Scopelitis, Lira, & Desutter, 2016). The ability to deal with typical representations and to fluently use them to communicate about domain principles is seen as the mark of an expert in a domain (Kozma, 2003; Kozma & Russell, 2005). In economic education, for example, graphs are used to discuss the relationship between price, supply and demand. The ability to work with these graphs is an important part of expertise and learners who fail to understand these graphs reveal “that they have not developed an economist’s way of thinking and practising" (P. Davies & Mangan, 2007, p. 721). In this regard, one of the ongoing debates in this research field is the relationship between the content or domain of a representation and the representational competence (Ainsworth, 2006). The author argues that in order to learn properly with representations within a domain, learners need to understand the form of representation, the relationship between the representation and the domain, as well as the rules how to select and construct an appropriate representation within the domain. Another debate refers to the importance of prior knowledge: Different studies identified the so-called *representational dilemma* (Dreher & Kuntze, 2015; Rau, Aleven, & Rummel, 2017): teachers have to enable learning *about* representations (e.g. how to read a certain graph) and learning *from* representations (e.g. what does the graph tell us for the domain question).

3.2.3 Challenges when learning with logical pictures

In graph research, scholars address students' strategies to work with logical pictures as well as their common mistakes, errors or misconceptions (Glazer, 2011). For instance, Åberg-Bengtsson and Ottosson (2006) show that tasks that go beyond the most obvious relationships and the reading of simple data points may be difficult for older students and even for college students when reading science graphs. Similarly, the connection between science concepts and graphical representations is a major challenge for students at the secondary level (Lai et al., 2016). Moreover, young adults are challenged when constructing a graph from data and given text and struggle with labelling axes or choosing the right graph type (Kotzebue, Gerstl, & Nerdel, 2015). Additional challenges regarding graph comprehension include confusing an interval and a point, difficulties with graph interpretation which result from design/format choices (e.g. features such as color, size, scale…) and (preservice-)teachers lacking expertise (Glazer, 2011). For histograms, Boels et al. (2019) summarized conceptual misconceptions possibly resulting from misunderstanding central statistical concepts - data (e.g. number of variables and measurement level) and distribution (shape, center and variability or spread). In addition to challenges with specific logical pictures, one of the major challenges is connecting and using multiple representations (text, equations, logical pictures) to answer domain questions (Kozma, 2003).

Most of these problems, however, are reported in the context of STEM-education and primarily for line graphs or histograms. We might expect some of these problems to be representation-specific and thus, potentially valid across different domains. Nevertheless, it remains unclear if these are "typical problems" for representation-related tasks which are also relevant for the economic domain and whether there are additional problems in the context of social sciences (as suggested by Westelinck, Valcke, Craene, & Kirschner, 2005).

3.2.4 Research aims

Since research on representational competence in economic education is scarce, we apply an exploratory approach and focus on which logical representations are used to what extent and how they are used in authentic classroom settings. Our aim is to shed light both on the most relevant representations and on the relevant abilities and typical challenges that arise in the context of economic education as a social science in secondary education. For this purpose, we conducted two studies: First, we analyzed logical pictures, i.e. the graphs and charts in the available economic textbooks in south-west Germany. As a result, we developed a category system that not only differs between the form of a representation (graph vs. chart) but also distinguishes between different levels of domain-specificity. With that in mind, in a second step, we aim to include the teaching perspective in the study by interviewing economic teachers how these different logical pictures are used in authentic economic classroom settings and what domain-specific and representation-specific challenges and competences students show when working with these representations.

3.3 Textbook analysis

3.3.1 Methods

3.3.1.1 Selection of textbooks

In order to identify and analyze relevant logical pictures, we used textbooks for the new school subject "economics, vocational and study orientation" [Wirtschaft, Berufs- und Studienorientierung] (grade level 8-10) in Baden-Württemberg/Germany for higher secondary school ("Gymnasium"). School books were chosen since they can be seen as a close representation of the potentially implemented curriculum (Valverde, Bianchi, Wolfe, Schmidt, & Houang, 2002) and are used as indicators of teaching practice in (social) science education (e.g. Strippel, Tomalla, & Sommer, 2018; Zimenkova, 2008). The analyzed corpus consists of all four available school books which were written in accordance with the official curriculum for the subject (Altmann, Boss, & Göser, 2018; Biehahn, Jüngling, Machoczek, Michael, & Ottmar, 2018; Burghardt et al., 2018; Kochendörfer, 2018) as well as the only textbook available (to our knowledge) for the upper secondary level (Bauer, Hamm-Reinöhl, Podes, & Riedel, 2012). Since the subject was newly introduced for grade levels 8-10 in the curriculum reform of 2016, it has only been taught in schools since the school year 2018/2019. Since these school books are the only ones available, they provide good insights into the amount and content of visualizations used for the subject economics. In total, we analyzed all 450 logical pictures available in the five school books in order to gain a comprehensive overview of the visualizations used.

3.3.1.2 Procedure

Due to the focus on logical pictures, in a first step, we excluded every visualization that was either an image, a drawing, a comic or a text-representation which showed minimal design elements (e.g. pro-contra lists) from the classification. Second, based on the two different groups of logical pictures described in subsection 3.2, a representation with a logical connection between object and representation, was labelled as chart if it represented qualitative relationships (e.g. flowcharts, hierarchical charts etc.). In contrast, a representation of a quantitative relationship was labelled as graph (e.g. line graphs, bar graphs, equilibrium graphs…). In a third step, we distinguished according to the domain-specificity of the logical pictures and distinguished between descriptive logical pictures which use every day terminology, logical pictures with domain terms and lastly typical visualizations of domain principles. With this category, we connect research focusing primarily on graph reading of rather descriptive – quantitative – representations (e.g. Åberg-Bengtsson & Ottosson, 2006), and studies that analyze the interplay between domain knowledge and visual representations (e.g. Lai et al., 2016). The latter stream of research works with visualizations of domain principles (e.g. Prey-Predator Relationship in biology).

Table 1

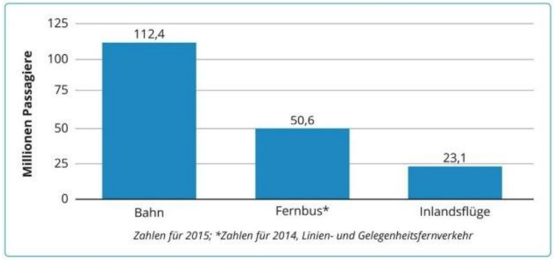
Category system for textbook analysis

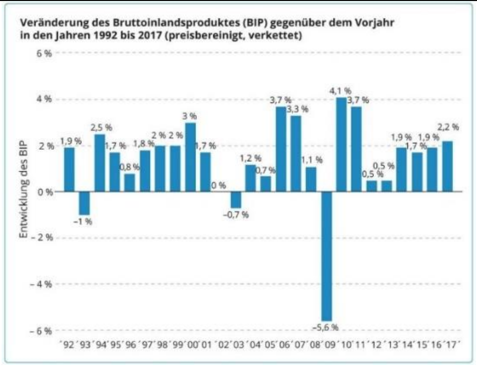
		Domain specificity		
		None	*Low-middle*	*High*
Connection between object and logical picture	Quantitative relationship “Graph”	Descriptive graph	Graph with domain terms	Graph with domain principle
	Qualitative relationship “Chart”	Descriptive Chart	Chart with domain terms	Chart with domain principle

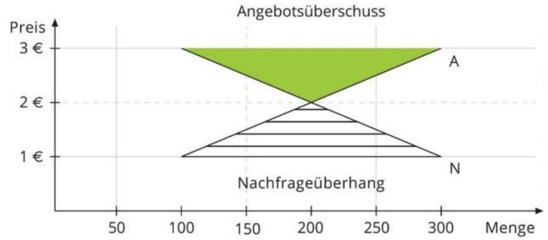
All 450 visualizations of the school books were then rated by two independent raters based on these categories (see Table 1). The inter-rater reliability was *Cohens Kappa* of 0.93 for the category “form” and 0.58 for “domain-specificity”. To achieve a clear count for further analysis, cases of dispute were solved by a third rater.

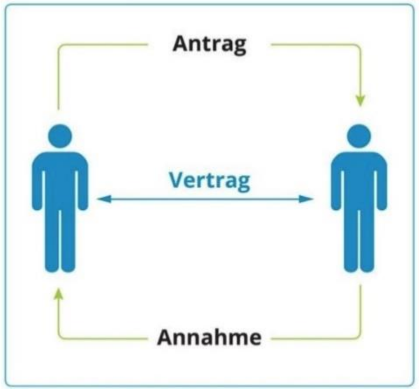
Table 2

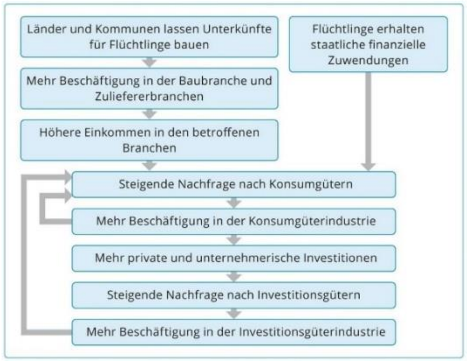
Category system for textbook analysis with anchor examples (sources of anchor examples in appendix, publication with permission by the right holders)

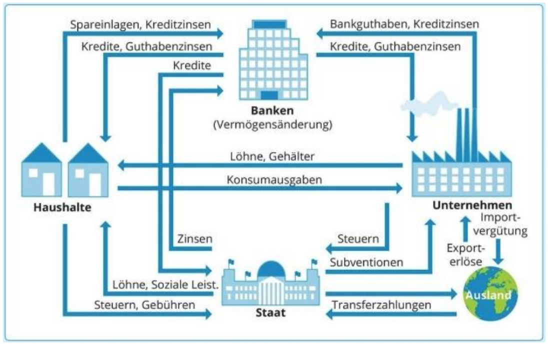
<table>
<tr><td colspan="2"></td><td colspan="3">Domain Specificity</td></tr>
<tr><td colspan="2"></td><td>No domain specificity</td><td>Low-middle domain specificity (terms)</td><td>Highly domain specific (principles)</td></tr>
<tr><td rowspan="2">Form</td><td>„Graph“</td><td>


</td><td>


</td><td>


</td></tr>
<tr><td>„Chart“</td><td>


</td><td>


</td><td>


</td></tr>
</table>

3.3.2 Results

3.3.2.1 Anchor Examples

Based on the category system, table 2 shows the anchor examples of the different categories in German. In the following, the examples will be briefly described to illustrate the different categories that we found in the school books.

Descriptive graphs without domain-specific terms do not require any domain knowledge. The anchor example (upper left panel in Table 2) visualizes the number of passengers in millions that use railway, long-distance buses or domestic flights as means of transportation. These graphs require knowledge of graph reading rules.

The anchor example for **graphs with domain-specific terms** (upper middle panel in Table 2) is a bar graph that depicts the change in the gross domestic product compared with the previous year between 1992 and 2016. In addition to graph rules, economic knowledge of the term gross domestic product is necessary. Furthermore, context knowledge might be necessary to interpret some parts of the graph, e.g. the drop of GDP in 2009 as a result of the global financial crisis.

The anchor example for **graphs visualizing domain-specific principles** (upper right panel in Table 3) shows the excess supply/demand in Marshall's supply and demand graph. Knowledge of the underlying terms (e.g. excess demand) and of the principles (price develops from an equilibrium of aggregated supply and demand) is necessary to work with this model. With respect to the graph reading rules, this graph additionally differs from graphs in other disciplines. While the independent variable is regularly represented by the x-axis (in math and STEM disciplines), in this graph, the independent variable (price) is depicted on the y-axis and the dependent variable (supply/demand) is depicted on the x-axis.

The anchor example for **descriptive charts without domain-specific terms** (lower left panel in Table 2) is a simple visualization visualizing a contract between two individuals. The design elements (here: arrows) show the relationship between the two contract partners. The terms used in the chart are rather everyday language than specific economic terminology. In consequence, no domain knowledge is needed to interpret the terms or the design elements.

For **charts with domain-specific terms**, the interpretation of the design elements does not necessarily require domain-specific knowledge, while understanding the terms used is necessary to grasp the whole chart. The anchor example (lower middle panel in Table 2) is a visualization of the possible economic effects of government spending on refugees. The design elements (arrows) can be interpreted as “consequence of”, the terms used, e.g.

“Investitiongüter” [capital goods], “Konsumgüter” [consumer goods], are economic terminology.

The anchor example for **charts representing domain-specific principles** (lower right panel in Table 2) is a visualization of the model of the circular flow of income, which visualizes economic interrelations between different actors (state, bank, households, companies). The design elements (arrows) are labelled with economic terms and visualize the interplay between the actors. The interpretation of this chart is difficult without comprehensive knowledge of the terms used and the underlying economic principles.

3.3.2.2 Graphs and charts in textbooks

In the five textbooks, a total of 450 visualizations were identified as logical pictures (graph or chart). On average over all books, there is one logical picture every 3-4 pages. Of these visualizations, 276 were rated as graphs, and 174 were rated as chart. An overview of the textbooks can be found in Table 3.

Table 3

Overview of logical pictures in different economic textbooks

	Grade level	Pages	Logical pictures Total	Graphs	Charts	Page/ Logical Picture Ratio	Correlation between page number and domain specificity (Kendall’s Tau)
Book 1	8-10	312	87	50	37	3.6	0.24 (p = .004)
Book 2	8-10	280	49	26	23	5.7	0.28 (p = .016)
Book 3	8-10	213	87	32	55	2.4	0.26 (p = .002)
Book 4	8-10	312	85	67	18	3.6	0.17 (p = .040)
Book 5	11-12	455	142	101	41	3.2	-0.19 (p = .004)

In the following Figure 2, we provide an overview of the number of graphs and charts in relation to their domain specificity.

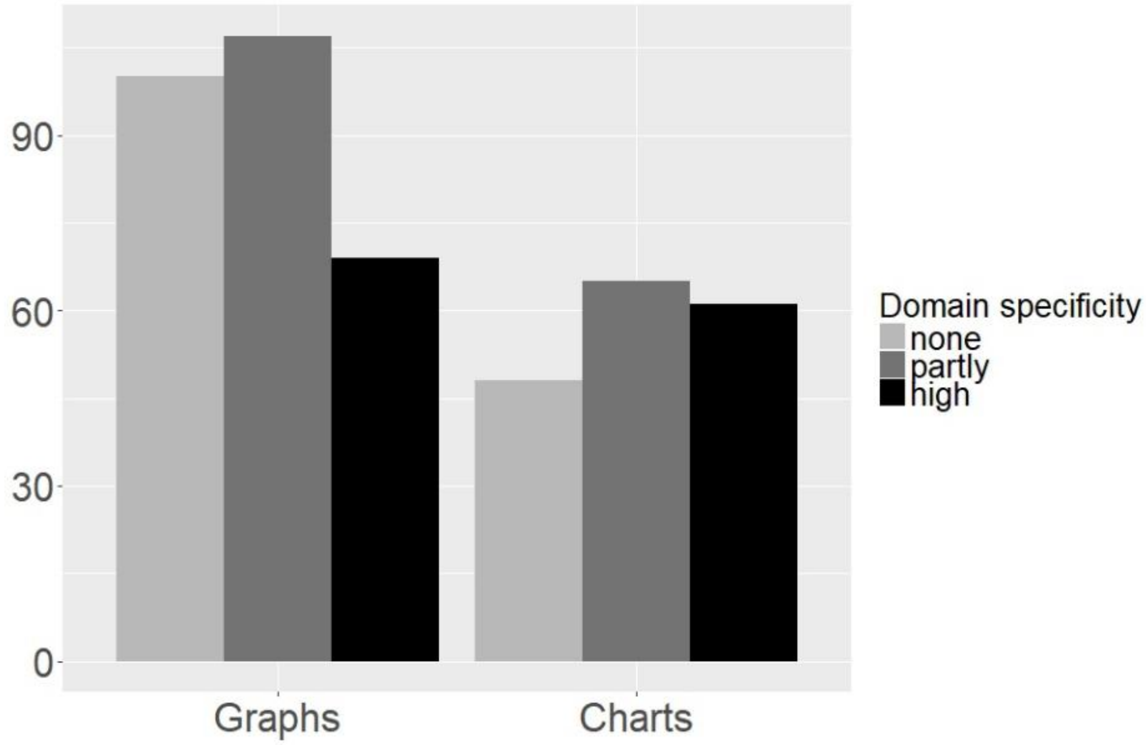


Figure 2. Number of graphs and charts in economic textbooks of the secondary level, subdivided according to the domain-specificity (own illustration)

In total, 100 graphs (in comparison 48 charts) were rated as descriptive with no domain-specific terms, which should, therefore, be understandable without content/prior knowledge. 107 graphs (65 charts) were rated as partly domain-specific, due to the use of domain-specific terms. 69 graphs (61 charts) were rated as highly domain-specific, due to their visualization of economic principles. To explore whether the domain-specificity raises over the course of a school-year, we analyzed the correlation between domain-specificity and page number (assuming that a book would be used in a rather linear fashion). For all grade level 8-10 textbooks, a correlation between 0.17 and 0.27 was found and therefore more domain-specific graphs are used later in the books. In the textbook for the upper secondary level, domain-specificity and page number correlated negatively ($r = -0.19$).

3.4 Teacher interviews

3.4.1 Methods

3.4.1.1 Participants and data collection

For the second study, we interviewed a sample of N=10 economics teachers in higher secondary schools (Gymnasium). Due to the new subject (see above), teachers had only taught economics at the upper secondary level (grades 11/12) at the time of this study. In consequence, we interviewed these already experienced teachers. Although the teachers represent a convenience sample, they differ in age (ranging from 34 to 63) and their second/third subject (languages, math, geography, history, physical education).

After a short introduction, where different logical pictures from the textbooks were shown as examples, the teachers were asked about the quantity of usage in teaching, their purpose/learning aims when using logical pictures, the importance of logical pictures for teaching, as well as relevant student competencies and difficulties. Each interview lasted between 20-45 minutes. All of the interviews were carried out by the first author, audiotaped and transcribed shortly afterwards. Although the questions varied depending on the course of conversation and the use of logical pictures of individual teachers, an interview guideline was used (see Appendix).

3.4.1.2 Content analysis procedure

The transcripts were analyzed applying qualitative content analysis (Kuckartz, 2016), using the software MAXQDA. Initially, we coded the answers according to three main categories that emerged from the guiding questions (use of logical pictures, competencies and challenges). For the use of logical pictures, the following subcategories were created inductively: use on average, importance and rationale for usage/learning goals. For competencies and challenges, we established subcategories deductively based on already established challenges and competencies in the literature, e.g. mathematical competencies (e.g. Ludewig, Lambert, Dackermann, Scheiter, & Möller, 2019), reading comprehension (e.g. Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014), integrating multiple representations (e.g. Kozma, 2003), prior knowledge/background knowledge (e.g. Nitz et al., 2014); other subcategories were inductively supplemented based on the answers of the teachers (e.g. standardized description). An overview of the categories (and results) can be found in Table 4. For reasons of readability, we will only show the English translations (by the first author) of relevant quotations (the original German versions are available upon request).

3.4.2 Results

Table 4

Overview of teacher interviews results

Category	**Rating / Subcategory**
Use on average	Every lesson (2/10) Every second lesson (5/10) Every third/fourth lesson (3/10)
Importance	Important or very important for teaching economics (10/10)
Rationale of usage and learning goals	Derive a problem, introduce a topic, activation of prior knowledge or fact-checking (4/10) Visualize economic principles and relationships, model thinking (7/10) Methods training, critical evaluation skills, graph reading skills (6/10) Place topic in larger context, structure domain knowledge, overview (5/10)
Competencies	Reading comprehension (5/10) Mathematical/logical abilities (5/10) Integrating multiple representations (3/10) Content and background knowledge (6/10) Critical thinking and evaluation (6/10) Standardized description/analysis (4/10) Construction of logical pictures (4/10)
Challenges	Mathematical/logical abilities (7/10) Precision and concentration (4/10) Connecting logical picture with tasks and identifying key points (7/10) Critical evaluation (4/10)

3.4.2.1 Quantity of usage and relevance of logical pictures

Teachers use graphs and charts regularly for their teaching, from every hour to once a week. For the teachers, the quantity depends on the content of the lesson, some topics, for example markets, are always taught with graphs.

> *"So if we examine market and prices, it's in every hour, even in the practice phases, but otherwise perhaps on average once a week." (Interview 9, #00:03:03-5#)*

Teachers consistently deem graphs and charts as important to very important. They justify this both based on the discipline/subject content (e.g. thinking in models and data in economics), the advantage of logical pictures for learners (e.g. different forms of access to content, multiple representations), but also the advantages from the teacher's point of view (e.g. using charts as a way of securing results/checking understanding).

> *"Yes, they [the logical pictures] are from my point of view one of the most central sources, even more so than others, I think the economy or economic education is characterized by the fact that statistical analysis plays a major role, and thus logically […] logical pictures are central" (Interview 3, #00:07:12-2#).*

3.4.2.2 Rationale for usage and learning goals

Teachers use different types of graphs and charts, depending on content and learning goals. Descriptive graphs are used to derive a problem, introduce a new topic or activate prior knowledge (Interview 10, #00:15:08-9#).

More complex graphs (e.g. supply and demand graphs, Lorenz curve etc.) and charts (e.g. circular flow of income), are used to develop economic knowledge, i.e. to learn about economic principles, to apply model thinking and to assess interrelations and processes.

> *"… this supply and demand graph to somehow see changes from one element to the whole, so […], the supply increases the demand decreases, what happens then […], or if there is a price change due to tariffs […], what effect does this have?" (Interview 5, #00:05:22-2#)*

In a similar vein, a teacher described the circular flow of income as a good visualization to place the topic in a larger context and to get an overview regarding the economic system:

> *"Well, let's take this circular flow of income as an example, if you introduce the system of economics, what instances there are, which actors are active, and this is gradually developed in a chart, then of course [...] it is good for clarity, and ultimately, like everything we do in class, it ensures learning success." (Interview 6, #00:03:59-3#, #00:04:16-1#)*

Teachers also use certain logical pictures to train methodological competence, to work with data, e.g. to check facts/hypotheses, to foster graph comprehension/statistical literacy and critical thinking competencies (Interview 5, #00:05:22-2#).

3.4.2.3 Relevant competencies

Reading and text comprehension is regarded as a basic (but not unimportant!) prerequisite by the teachers, to understand the relevant part of graphs as well as accompanying text and tasks:

"So what I basically notice in teaching economics, as in any other lesson, is that text, language ability, is the basic prerequisite. […], it first needs a lot of language skills to understand the axis label, the heading in context." (Interview 6, #00:07:40-8#)

To be able to work with logical images – especially with graphs in economics, students need certain *mathematical competencies* (e.g. absolute versus relative numbers). It is also important that they recognize in which unit of measurement the values are displayed and which concept is displayed on which graph axis (*graph reading competencies).*

"…so exactly with these supply and demand curves [...] they must be able to recognize what is represented there, so just they must look at the numerical values, what is represented on the x- y-axis, in which unit of measurement are these values represented…" (Interview 5, #00:09:17-0#)

According to the teachers, students also need the ability to *integrate information from multiple representations*. This includes to be able to "translate" between different external representations:

"If you introduce logical pictures you say you can display something as graph, as table, as flow chart and as text and they [the students] have to be able to jump back and forth between these display forms." (Interview 6, #00:05:35-0#)

Content and background knowledge were also deemed important by the teachers for analysis, interpretation and evaluation. As content knowledge, they explicitly mentioned the relevant keywords in the logical pictures (e.g. gross domestic product, aggregated income…) as well as "historical" context knowledge (e.g. to explain economic development over a longer period of time). Some teachers noted that the amount of expertise needed depends very much on the graph and task and cannot be generalized.

"Do I need economic knowledge? That depends, with the supply and demand curves yes, I need economic knowledge. If I have a statistic on GDP, first of all, to understand the visualization, I don't need quite as much economic knowledge, of course I apply it later [...] At the second step I need it, clearly" (Interview 4, #00:06:31-2#, #00:06:32-9#)

Some teachers point out that students need to be able to *construct logical pictures*. One teacher regularly asks the students to construct a flow chart from a given text, whereas another teacher instructs his students to draw graphs from data (e.g. the course of unemployment rate, price increase and economic growth since the 1950s).

According to the teachers, *critical evaluation* of the content and the visualization is a crucial ability. Therefore, they expect their students to examine logical pictures and their content regarding their possibility to influence the reader on different levels. As examples of what to analyze critically, they cited design choices (e.g. manipulation through a limited display of data or displaying axis in different units etc.), the publisher/origin/source of graph and data, their underlying interests as well as the deconstruction of common economic models (that are often depicted graphically).

> *"…on the other hand, it is important that you can also somehow evaluate diagrams from different sources, depending on who made them, that's always an important part of it. A critical approach to logical pictures." (Interview 1, #00:04:33-8#)*

The teachers assessed it as important that students use a *standardized description/analysis procedure* to deal with graphs. In addition to an introductory sentence, teachers expect three different requirement levels: (1) describing the graph, (2) analyzing (e.g. what is important, what is not important for a question at hand), (3) evaluation (where does the graph come from, who published it and what could be the interests behind it?):

> *"If you divide this into three parts, okay, I have the description level, I have the classic interpretation level, which then means that from this graph you can clearly see that this and that will happen or is predicted and the third part is than with what intention was the logical picture published by whom" (Interview 10, #00:13:12-6#)*

3.4.2.4 Difficulties and challenges

According to the teachers, students often view a graph or a chart as a single representation and fail to see *connections to other representations or to the domain task*. As a consequence, they struggle to identify key points of a logical picture, cannot separate important and unimportant information and lack understanding:

> *"[…] the analytical competence, so how do I read the thing and what does it really mean what I see there, and can I summarize it in one sentence and if I can't do that, it often shows that they simply haven't understood what the axes say, or what the percentages or the numbers as a whole are supposed to say, that's actually rather the problem" (Interview 1, #00:06:55-2#)*

The teachers also identified *critical evaluation* as one of the more challenging parts of working with logical pictures (in line with requiring it as a competence, see 4.2.3). Students are not always able to detect manipulation (e.g. through a limited display of data or displaying axis in different units) and they do not always include the source of the graph/chart in their analysis.

> *"Only a fraction of students achieves this third level […], namely to work out this interpretational sovereignty so that their opinion is not externally determined. That, of course, is what we want, the responsible citizen." (Interview 10, #00:13:12-6#)*

Among the problems that teachers face when working with quantitative visualizations in economics are *missing mathematical abilities*. According to the teachers, students are not always able to calculate percentage scores, distinguish absolute from relative numbers, calculate the median or average or work with index numbers although this is identified as a crucial competence.

> *"Then what I really notice is that some of them really lack the rudimentary abilities [...] that they highlight certain numbers by putting them in relation [...] and then I notice that sometimes it's hard for them to calculate the percentage scores, [...] not everyone is able to use the rule of three." (Interview 7, #00:08:17-5#)*

Finally, some teachers point out that students are not *precise* enough. When analyzing a graph, they might sometimes miss details, e.g. the unit in which a certain parameter is displayed. The teachers attribute this kind of challenges to *concentration* rather than to conceptual understanding or lacking mathematical abilities.

3.5 Discussion

The objective of these two studies was to show how logical pictures are used in secondary economic education by a) analyzing logical pictures in textbooks with regard to their form (chart or graph) and their domain-specificity and b) interviewing economic teachers about their experiences.

With our study, we distinguished between different forms of logical pictures (e.g. Harris, 2000; Kosslyn, 1999), and, between different levels of domain-specificity. The latter adds another layer to the already existing frameworks as it allows for a separation of descriptive representations that visualize (economic) content with usual graphs or charts from visualizations which show economic principles.

This separation is helpful in understanding the different abilities and challenges that learners face in the classroom: whereas certain reading rules are necessary for all visualizations, content knowledge and more specific reading rules seem to be necessary for the "more domain-specific" visualizations and thus should be addressed in secondary economic education in order to support learners to become domain experts. The fact that not all logical pictures in a textbook are domain-specific can be explained by the nature of the subject: it is the first time students are confronted with economic concepts. In consequence, some logical pictures (more in earlier sections of the textbooks as the correlations between page number and domain-specificity overall shows for the earlier text books) rather use everyday language – and, accordingly, are not rated as domain-specific.

As described by the representational dilemma (Dreher & Kuntze, 2015; Rau et al., 2017), the teachers in the *second study* explain how they use logical pictures not only for the development of domain knowledge (learning from representations) but also to develop strategies to learn *with* logical pictures. They expect their students to use a strategy (e.g. standardized description) for their analysis.

The teachers mentioned some of the abilities and challenges that are already established in graph comprehension and representational competence literature, e.g. reading competence (e.g. Åberg-Bengtsson & Ottosson, 2006) or (basic) math competencies (e.g. Ludewig et al., 2019). Comparable to findings in science education (e.g. Cook, 2011; Kozma, 2003), they mentioned the need to connect multiple representations, to switch between different representations and to construct their own logical pictures (based on data or text). Furthermore, this study also confirms the need for content and background knowledge (e.g. Stern, Aprea, & Ebner, 2003).

Teachers also argued that the evaluation of logical pictures is a very important step in the classroom. Accordingly, one of the central challenges for students (from the teachers' perspective) is to identify the relevant key features of a logical picture for the task at hand and to critically evaluate the visualization also in light of purposefully using statistics to influence the reader. This is, above all, important as graphs are used as plausibility tools in texts (Isberner et al., 2013) and might tempt readers to focus less on the text (Ögren, Nyström, & Jarodzka, 2017).

Overall, both studies showed that logical pictures are important for economic education. The high number of graphs and charts in all textbooks corresponds to their relatively frequent usage in authentic classroom settings. The fact that there are more graphs than charts in the textbooks might be connected to the demand for more "mathematical" competencies and the prevalence of math-related challenges the teachers described in study 2. Furthermore, the different levels of domain-specificity in the textbook analysis might correspond to different demands on students. For example, for descriptive visualizations, students might need graph reading and analytical competencies, while more conceptual understanding and content knowledge would be a prerequisite for more domain-specific logical pictures.

3.5.1 Limitations

When interpreting these results, however, it is important to keep the limitations of the two studies in mind: First of all, we only interviewed teachers who were teaching on the upper secondary level. However, we argue that it is reasonable to assume that most of the described competencies and challenges are accurate for different grade levels in a broad sense whereas details (e.g. which mathematical competencies do students need exactly?), might differ between grade levels (e.g. basic math competencies and graph reading rules might be more relevant in earlier grade levels whereas conceptual understanding and more complex math / statistical knowledge might be expected at the upper secondary and university level). Thus, the differences between the competencies needed could be an interesting array for future research.

Although we analyzed all textbooks available for grammar schools in south-western Germany, it is not a representative sample for all secondary economic education textbooks in Germany – the same is true for the teacher-interviews. Since we were only able to conduct interviews with a convenience sample of teachers, the relevance and quantity of usage of visualizations might be overestimated (teachers that use many visualizations and consider them important are much more likely to participate in a voluntary interview). However, we attempted to find different teachers with varying teaching experience and from multiple backgrounds (concerning their other school subjects) to allow for some variability among the teachers. Lastly, we analyzed the instructional material and interviewed teacher about their use, i.e. we focused on the teaching input. However, for now, we were not able to observe the actual use in the classroom, which could be the next step in future research.

3.5.2 Implications and further research

Despite the limitations, the two studies emphasize the importance to distinguish between different logical pictures such as graphs and charts and regarding their domain-specificity as they are powerful tools to explain and visualize economic concepts (Hey, 2005; VanderMolen & Spivey, 2017). Although they are used frequently in teaching, the teacher interviews indicate that students need support to understand and evaluate the logical pictures. At the same time, as different logical pictures are offered in textbooks and since experienced teachers report different challenges when using them in the classroom, it is advisable to include the topic in teacher education for social sciences as well.

In addition to these practical implications, there are different avenues for future research. First of all, it would be necessary to replicate the findings for different grade levels and other local contexts as well as to empirically confirm the necessary students' competencies and the challenges identified.

There is also a need to examine how to foster representational competence in the social sciences – e.g. to design material in a way that connects graph-text and task (as done for other domains in multimedia research, e.g. Scheiter & Eitel, 2015) or to design trainings where learners use certain graph and text-reading strategies (e.g. Seufert, 2019) or construct their own representations (e.g. Stern et al., 2003). As especially critical evaluation is difficult for students, it would be interesting to analyze how teachers (and/or technologies) can support students in that regard.

From our findings, it is clear that such interventions should be adapted to the domain-content that is taught with the representations, since – especially for more domain-specific graphs and charts – content and representation are to be seen as integrated teaching material.

3.6 References

Åberg-Bengtsson, L., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*(1), 43–62. https://doi.org/10.1002/tea.20087

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Altmann, G., Boss, G., & Göser, U. (2018). *Wirtschaft & Du: Wirtschaft/Berufs- und Studienorientierung [Economy & You - Economics, Vocational and Study orientation]* (Gymnasium, SI, Baden-Württemberg, Druck A, Prüfauflage). Braunschweig: Westermann.

Aprea, C., & Bayer, D. (2010). Instruktionale Qualität von grafischen Darstellungen in Lehrmitteln: Kriterien zu deren Evaluation: [Instructional quality of graphical representations in teaching materials: Criteria for their evaluation]. *Beiträge Zur Lehrerinnen- und Lehrerbildung*, *28*(1), 73–83.

Bauer, M., Hamm-Reinöhl, A., Podes, S., & Riedel, H. (2012). *Unterrichtswerk für die Oberstufe Wirtschaft: Märkte, Akteure und Institutionen [Textbook for higher grade level: Economics - Markets, Actors and Institutions]* (1. Auflage). *Kolleg Politik und Wirtschaft*. Bamberg: Buchner, C C.

Biehahn, M., Jüngling, H., Machoczek, M., Michael, P., & Ottmar, C. (2018). *Auer Wirtschaft - Berufs- und Studienorientierung. Ausgabe Baden-Württemberg: Schülerbuch Klasse 8-10 [Auer Economics, Vocational and Study orientation, Textbook 8th-10th grade]* (1. Auflage). Stuttgart: Klett.

Boels, L., Bakker, A., van Dooren, W., & Drijvers, P. (2019). Conceptual difficulties when interpreting histograms: A review. *Educational Research Review*, 100291. https://doi.org/10.1016/j.edurev.2019.100291

Burghardt, Y., Heuser, J., Hamm-Reinöhl, A., Podes, S., Riedel, H., & Straub, J. (2018). *Wirtschaft & Co: Wirtschaft für das Gymnasium [Economics & Co, Economics for Gymnasium]*. Bamberg: C.C. Buchner Verlag.

Cohn, E., Cohn, S., Balch, D. C., & Bradley, J. (2001). Do Graphs Promote Learning in Principles of Economics? *The Journal of Economic Education*, *32*(4), 299–310. https://doi.org/10.1080/00220480109596110

Cook, M. (2011). Teachers' Use of Visual Representations in the Science Classroom. *Science Education International*, *22*(3), 175–184.

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? *Higher Education*, *62*(3), 279–301. https://doi.org/10.1007/s10734-010-9387-6

Davies, P., & Mangan, J. (2007). Threshold concepts and the integration of understanding in economics. *Studies in Higher Education*, *32*(6), 711–726. https://doi.org/10.1080/03075070701685148

Dreher, A., & Kuntze, S. (2015). Teachers Facing the Dilemma of Multiple Representations Being Aid and Obstacle for Learning: Evaluations of Tasks and Theme-Specific Noticing. *Journal Für Mathematik-Didaktik*, *36*(1), 23–44. https://doi.org/10.1007/s13138-014-0068-3

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *31*(3), 444–457. https://doi.org/10.1177/0272989X10373805

Gebre, E. H., & Polman, J. L. (2016). Developing young adults' representational competence through infographic-based science news reporting. *International Journal of Science Education*, *38*(18), 2667–2687. https://doi.org/10.1080/09500693.2016.1258129

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, *47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Harris, R. L. (2000). *Information graphics: A comprehensive illustrated reference*. Oxford: Oxford University Press.

Hey, J. D. (2005). I Teach Economics, Not Algebra and Calculus. *The Journal of Economic Education*, *36*(3), 292–304. https://doi.org/10.3200/JECE.36.3.292-304

Huber, M., & Stallhofer, B. (2010). Diskontinuierliche Texte im Geografieunterricht [Discontinuous texts in geography teaching]. *ProLesen. Auf Dem Weg Zur Leseschule–*

Leseförderung in Den Gesellschaftswissenschaftlichen Fächern. Donauwörth: Auer, 223–240.

Isberner, M.-B., Richter, T., Maier, J., Knuth-Herzig, K., Horz, H., & Schnotz, W. (2013). Comprehending conflicting science-related texts: graphs as plausibility cues. *Instructional Science, 41*(5), 849–872. https://doi.org/10.1007/s11251-012-9261-2

Jägerskog, A.-S. (2020). *Making Possible by Making Visible: Learning through Visual Representations in Social Science* [Doctoral Dissertation, Stockholm University]. Stockholm: Department of Humanities and Social Sciences Education, Stockholm University.

Kochendörfer, J. (2018). *Startklar! (Oldenburg) - Wirtschaft/Berufs- und Studienorientierung: Schülerbuch Klasse 8 -10 [Economics, Vocational and Study orientation, Textbook 8th-10th grade]*. Berlin: Oldenburg Schulbuchverlag.

Kosslyn, S. M. (1999). *Image and brain: The resolution of the imagery debate* (4. print). *A Bradford book*. Cambridge, Mass.: MIT Press.

Kotzebue, L. von, Gerstl, M., & Nerdel, C. (2015). Common Mistakes in the Construction of Diagrams in Biological Contexts. *Research in Science Education*, *45*(2), 193–213. https://doi.org/10.1007/s11165-014-9419-9

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*(2), 205–226. https://doi.org/10.1016/S0959-4752(02)00021-X

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*(9), 949–968.

Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representational Competence. In J. K. Gilbert (Ed.), *Models and Modeling in Science Education: Vol. 1. Visualization in Science Education* (pp. 121–145). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3613-2_8

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung [Qualitative Content Analysis: methods, practice, computer support]* (3., überarbeitete Auflage). *Grundlagentexte Methoden*. Weinheim, Basel: Beltz Juventa.

Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R., & Linn, M. C. (2016). Measuring Graph Comprehension, Critique, and Construction in Science. *Journal of Science Education and Technology*, *25*(4), 665–681. https://doi.org/10.1007/s10956-016-9621-9

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of picture in prose. In D. M. Willows & H. A. Houghton (Eds.), *The Psychology of Illustrations (Vol. 1)* (pp. 51–114). Harrisonburg, VA: R. R. Donnelley & Sons.

Ludewig, U., Lambert, K., Dackermann, T., Scheiter, K., & Möller, K. (2019). Influences of basic numerical abilities on graph reading performance. *Psychological Research*, 1–13. https://doi.org/10.1007/s00426-019-01144-y

Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction*, *31*, 13–22. https://doi.org/10.1016/j.learninstruc.2013.12.003

Ögren, M., Nyström, M., & Jarodzka, H. (2017). There's more to the multimedia effect than meets the eye: is seeing pictures believing? *Instructional Science*, *45*(2), 263–287. https://doi.org/10.1007/s11251-016-9397-6

Peterman, K., Cranston, K. A., Pryor, M., & Kermish-Allen, R. (2015). Measuring Primary Students' Graph Interpretation Skills Via a Performance Assessment: A case study in instrument development. *International Journal of Science Education*, *37*(17), 2787–2808. https://doi.org/10.1080/09500693.2015.1105399

Rau, M. A., Aleven, V., & Rummel, N. (2017). Making connections among multiple graphical representations of fractions: sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science*, *45*(3), 331–357. https://doi.org/10.1007/s11251-017-9403-7

Ring, M., Brahm, T., & Randler, C. (2019). Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs. *International Journal of Science Education*, *41(13)*, 1787–1804. https://doi.org/10.1080/09500693.2019.1640915

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. *The Reading Teacher*, *67*(1), 12–24. https://doi.org/10.1002/TRTR.1174

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

Scheiter, K., Schüler, A., Gerjets, P., Huk, T., & Hesse, F. W. (2014). Extending multimedia research: How do prerequisite knowledge and reading comprehension affect learning from text and pictures. *Computers in Human Behavior*, *31*, 73–84. https://doi.org/10.1016/j.chb.2013.09.022

Schnotz, W. (2001). Wissenserwerb mit Multimedia [Knowledge acquisition with multimedia]. *Unterrichtswissenschaft*, *29*(4), 292–318.

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instructional Science*, *36*(3), 175–190. https://doi.org/10.1007/s11251-007-9029-2

Schopf, C., Raso, A., & Kahr, M. (2019). How to give effective explanations: Guidelines for business education, discussion of their scope and their application to teaching operations research. *RISTAL-Research in Subject-Matter Teaching and Learning*, *2*, 32–50. https://doi.org/10.23770/rt1823

Seufert, T. (2019). Training for Coherence Formation When Learning From Text and Picture and the Interplay With Learners' Prior Knowledge. *Frontiers in Psychology*, *10*, 193. https://doi.org/10.3389/fpsyg.2019.00193

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The Comprehension of Quantitative Information in Graphical Displays. In P. Shah & A. Miyake (Eds.), *Cambridge Handbooks in Psychology*. *The Cambridge Handbook of Visuospatial Thinking* (pp. 426–476). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.012

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction*, *13*(2), 191–203. https://doi.org/10.1016/S0959-4752(02)00020-8

Stieff, M., Scopelitis, S., Lira, M. E., & Desutter, D. (2016). Improving Representational Competence with Concrete Models. *Science Education*, *100*(2), 344–363. https://doi.org/10.1002/sce.21203

Strippel, C., Tomalla, L., & Sommer, K. (2018). A cross-subject content analysis of science textbooks using the understandings about scientific inquiry rubrics. *RISTAL. Research in Subject-Matter Teaching and Learning*, *1*, 66–81.

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the Book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Dordrecht: Springer Netherlands.

VanderMolen, J., & Spivey, C. (2017). Creating infographics to enhance student engagement and communication in health economics. *The Journal of Economic Education*, *48*(3), 198–205. https://doi.org/10.1080/00220485.2017.1320605

Westelinck, K. D., Valcke, M., Craene, B. de, & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior*, *21*(4), 555–573. https://doi.org/10.1016/j.chb.2004.10.030

Winn, W. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), *The Psychology of Illustrations (Vol. 1)* (pp. 152–198). Harrisonburg, VA: R. R. Donnelley & Sons.

Zimenkova, T. (2008). Citizenship Through Faith and Feelings: Defining Citizenship in Citizenship Education. An Exemplary Textbook Analysis. *JSSE - Journal of Social Science Education, Transformation and Citizenship Education II*, *9*(1), 81–111. https://doi.org/10.4119/JSSE-432

3.7 Appendix

3.7.1 Interview guideline questions

- How often do you use logical pictures in teaching economics?
- What graphs/charts are you using in teaching economics?
- What is your goal when using logical pictures in economic education?
- In what part of the lesson are you using logical pictures?
- How important are logical pictures for teaching economics?
- What competencies are necessary for students to work with logical pictures in economic class?
- What difficulties do students have when working with logical pictures?
- How do these problems manifest?
- In your opinion, how could the work with logical pictures be improved?
- Do you have additional thoughts regarding logical pictures in teaching economics?
- Do you have any open questions?

3.7.2 Sources of anchor examples

- **Descriptive graph:** Burghardt et al., 2018, 80, originally in: Christin, J., S. Schultz. 2016 "Flixbus kauft Postbus - Was die Fernbusfusion für Fahrgäste bedeutet", SPIEGEL ONLINE, Accessed October 21, 2019. https://www.spiegel.de/wirtschaft/service/flixbus-kauft-postbus-was-die-fernbus-fusion-fuer-sie-bedeutet-a-1105929.html. data source: destatis
- **Graph with domain terms:** Burghardt et al., 2016, 217 orginally in: Statista research department (2016). Accessed October 21, 2019 https://de.statista.com/statistik/daten /studie/2112/umfrage/veraenderung-des bruttoinlandprodukts-im-vergleich-zum-vorjahr. data source: destatis)
- **Graph with domain principle:** Burghardt et al., 2018, 75
- **Descriptive chart:** Burghardt et al., 2018, 33
- **Chart with domain terms:** Burghardt et al., 2018, 229
- **Chart with domain principle:** Burghardt et al., 2018, 38 originally in Schmitz, U., B. Weidtmann. (2000) Handbuch der Volkswirtschaftslehre, Stuttgart: Klett

Acknowledgements

We would like to thank the teachers who participated in the interviews, the textbook publishers and the original right holders for the permission to display the graphs and charts used as anchor examples. Our gratitude also extends to the student assistants who helped with the interview transcriptions.

Malte Ring is a doctoral student at the LEAD Graduate School & Research Network [GSC1028], which is funded within the framework of the Excellence Initiative of the German federal and state governments.

4 How to support text-graph integration: Comparing the effects of passive and active signaling on learning outcomes

Ring, M., Brahm, T., Richter, J., Scheiter, K., & Randler, C. (2020). How to support text-graph integration: Comparing the effects of passive and active signaling on learning outcomes. *Manuscript submitted for publication*.

Abstract

Integrating multiple representations into a coherent mental model is one of the challenges when learning with text and pictures. In this quasi-experimental study (N = 173), we examined how highlighting corresponding information in text-graph learning material can help learners to make the necessary connections and, thus, improve learning outcomes in two domains. We compared a control condition with no highlights to a signaling condition with given highlights and an active signaling condition where students were asked to visually highlight corresponding text and graph information themselves while learning. There was no overall benefit of active signaling; rather, it improved learning for high prior knowledge learners in one of the domains. A mediation analysis revealed that the effect of prior knowledge was mediated via the quality of learner-generated correspondences. Our findings suggest that different methods of supporting text-graph integration might be useful for different students and different learning material.

Keywords: Multimedia learning; signaling; graph-text-integration; highlighting

4.1 Introduction

Printed as well as digital learning materials in most school subjects comprise multiple external representations such as combinations of text, pictures, and tables. Graphical representations are used to further explain or to illustrate verbal content. There is a plethora of empirical evidence suggesting that graphical representations support learners to better understand the learning content. In particular, learners benefit from multiple representations (e.g., text and graph) compared to text alone – a finding known as the multimedia effect (Mayer, 2014). To take advantage of multiple representations, learners need to integrate their information into a coherent mental model (Schnotz, 2014). However, this may be challenging for learners due to high cognitive demands. Different strategies to help learners integrate representations have been developed and shown to improve learning outcomes (e.g., Bodemer & Faust, 2006; Seufert, 2003; van der Meij & de Jong, 2011). Renkl and Scheiter (2017) distinguish between material-oriented and learner-oriented interventions: passive signaling or cuing, where corresponding information in both text and graph is highlighted (e.g., by marking it using the same color), is an example for a material-oriented intervention as it pertains to improving the design of the instruction. In this paper, we refer to this traditional form of signaling as passive since the learner does not contribute to generating the signals. We distinguish it from active signaling where learners are asked to create signals on their own (e.g., by marking correspondences in the materials or adding annotations to it), which would be an example of a learner-oriented intervention.

Optimized learning materials can aid learning; however, they may entail the danger of learning becoming overly passive-receptive, as learners are no longer required to deeply process the contents to figure out how to integrate new knowledge into existing knowledge structures. Accordingly, constructivist learning theories imply that better learning performance is achieved when the learner actively engages with the learning material. This results in active-productive processing of learning content and, in turn, better learning outcomes (Chi & Wylie, 2014). Against this backdrop, one might expect active signaling to be superior to passive signaling of correspondences. At the same time, making connections between representations is what is most challenging for learners studying multimedia material. Thus, it is yet an open question whether learners will be able to cope with the cognitive demands resulting from having to actively generate text-graph correspondences.

To contribute to theories of learning from multiple representations as well as to the body research on instructional support, the main aim of the present study was to investigate whether active signaling (learners highlighting the corresponding information in text and graph while learning from the materials) would be more effective (i.e., improve recall and comprehension learning outcomes) than material with pre-given signals (passive signaling) and material without signals.

Instructional support, however, is not a “one-size-fits-all” approach. For passive signaling, it is well documented that prior knowledge moderates its effectiveness (Richter, Scheiter, & Eitel, 2018). Moreover, similar to other generative learning strategies such as drawing (e.g., Schmeck, Mayer, Opfermann, Pfeiffer, & Leutner, 2014), effects of active signaling are likely to depend on the quality by which the strategy is implemented (i.e., the accuracy of correspondences marked by learners). Consequently, prior knowledge and quality of active signaling were investigated as potential boundary conditions in the present study.

4.2 Theoretical background

4.2.1 Supporting text-picture-integration

According to contemporary models of multimedia learning (e.g., the cognitive theory of multimedia learning [CTML], Mayer, 2014; the integrated model of text and picture comprehension [ITPC], Schnotz and Bannert, 2003), integration of verbal and graphical information into one mental model is a necessary prerequisite for successful learning. That is, learners need to identify correspondences between text and pictures and draw connections between them. Prior knowledge plays an important role as it helps to form an integrated mental model and store the information in long-term memory. Building an integrated mental model, however, is not a simple task and imposes high cognitive demands. It is, therefore, no surprise, that learners have been shown to benefit from instructional support that helps them to identify correspondences across the different representations.

4.2.1.1 Material centered support: Passive Signaling

Material-centered interventions refer to changes in the design or display of the learning material. For example, with passive signaling or cuing, relevant information in different representations is highlighted, without changing their content. It is assumed that passive signaling is effective because it can reduce learners’ cognitive demands by helping learners to select relevant information and to easily identify correspondences between multiple representations (van Gog, 2014). Examples of signals are color coding - where equivalent information in both representations is displayed in the same color – as well as deictic references in the text to the picture (e.g., “see Figure X”). Eye-tracking studies show that signals draw visual attention towards signaled elements earlier and more frequently (Ozcelik, Arslan-Ari, &

Cagiltay, 2010; Richter & Scheiter, 2019; Scheiter & Eitel, 2015). In general, signaling (i.e. highlighting relevant information) improves learning as well as motivation, learning time, and visual attention with small to medium effects (cf. meta-analysis by S. Schneider, Beege, Nebel, and Rey, (2018). Focusing on signals that aim specifically at highlighting correspondences between text and pictures (multimedia integration signals), a meta-analysis by Richter and colleagues (2016) yielded a small to medium effect on transfer/comprehension. The latter effect was evident only for learners with lower levels of prior knowledge, while there was no effect for learners with higher levels of prior knowledge. This finding – namely, that the effect of passive signaling is moderated by prior knowledge – was recently replicated by another meta-analysis (Alpizar, Adesope, & Wong, 2020).

In an experimental field study, Richter et al. (2018) analyzed the interaction of signaling with prior knowledge. 8th graders were distinguished into three groups having small, medium or high prior knowledge. They, then, studied authentic chemistry learning material that either contained multimedia signals (color coding and deictic references) or had signals only in the text (e.g., bold typeface for relevant terminology). Contrast analysis revealed that students with high prior knowledge learned more from material with text-only-signals whereas students with low-prior knowledge profited more from material with multimedia-signals (expertise-reversal effect, cf. Kalyuga, Ayres, Chandler, & Sweller, 2003). As a possible explanation, the authors argue that while low prior knowledge students receive the support that they need, those with high prior knowledge suffer disadvantages from processing potentially redundant information and being nudged into an overly passive processing mode that suppresses their cognitive potential to deeply engage with the content. In a lab study (Richter & Scheiter, 2019), the same material was used again, however, only a partial expertise reversal effect was found: high prior knowledge learners performed equally well independent of the signaling condition, while low prior knowledge learners profited from multimedia signals. According to eye-tracking data, signaling affected the visual attention patterns of low prior knowledge learners, who looked earlier at the diagram when signals were presented; there was no effect for high prior knowledge learners. To conclude, while signals aid students with low prior knowledge, the exact nature of effects for students with high prior knowledge is still less obvious.

4.2.1.2 Learner-centered support: Active signaling

Learner-centered interventions include, among other things, measures that require learners to actively engage with the learning material by generating new artefacts. These artefacts may consist of novel external representations as is the case for learner-generated drawings (Fiorella & Zhang, 2018) or augmentations of the existing representations (e.g., highlighting of text segments). In general, generative learning has been shown to promote learning (Fiorella & Mayer, 2016). A possible explanation for the success of generative learning can be derived from the ICAP-Framework, which states that learners benefit more when they actively engage with the learning material and even construct parts of it compared to the passive reception of pre-given materials (Chi & Wylie, 2014). However, there are certain boundary conditions for the generative learning effect to occur (Fiorella & Zhang, 2018). In particular, what is generated by the learners must be of high quality to represent the content to-be-learned accurately (cf. prognostic drawing effect, Scheiter, Schleinschok, & Ainsworth, 2017; Schmeck et al., 2014).

In this vein, for highlighting texts, some studies show a positive effect (Yue, Storm, Kornell, & Bjork, 2015), whereas most studies find effects that are comparable to reading a text without highlighting and as such describe highlighting as ineffective (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). A possible explanation is the quality of highlights: especially students with low-reading skills either mark irrelevant information or select too much information so that the highlights fail to emphasize important pieces of information. In consequence, highlighting might even have negative effects (Bell & Limber, 2009).

For learning from multiple representations, generative learning activities are aimed at supporting integration. For instance, Bodemer and colleagues studied the effects of active integration (Bodemer, Ploetzner, Feuerlein, & Spada, 2004). In their experiment, students learned how a tire pump works in one of three conditions: text and graph elements were a) physically separated, b) integrated or c) had to be actively connected by drag-and-drop activities. When actively integrating corresponding elements of the representations, students achieved higher learning outcomes, compared to the other two groups. The authors explain this effect with a more optimized use of cognitive resources (Bodemer et al., 2004). Other learner-centered interventions where learners were asked to perform integration only in their minds without creating an external representation were less effective (Bartholomé & Bromme, 2009; van der Meij & Jong, 2011). Moreover, also effects of active integration seem to depend on the quality of the learner-generated artifacts. Leopold, Doerner, Leutner, and Dutke (2015) asked students to study chemistry learning material consisting of text and pictures (and a control with only text). The text-and-picture-groups had to either a) write down important concepts of the text next to the picture (separate from it), b) write down important concepts into the picture (integrated) or c) they received no additional instructional support. While the writing interventions were ineffective at a group-comparison level, learning outcome was higher when learners constructed more accurate referential connections.

If the quality of learner-generated artifacts is crucial for the positive effect of the generative learning strategy, it might be argued that learners benefit from high prior knowledge, as it might help them to create higher quality artifacts (just as research that shows that high skilled readers are better at highlighting, cf. Bell & Limber, 2009). This could be the case, in particular, for tasks which require intensive use of cognitive resources, such as integrating multiple representations. For integrating information from text and pictures, prior knowledge might help both to identify correspondences and to process them at a deep structural level to construct an integrated mental model. (Schnotz & Bannert, 2003). In this vein, Seufert (2019) found benefits of training different reading strategies aimed at identifying text-picture correspondences only for learners with higher levels of prior knowledge. She suggests that these differential effects are due to the fact that high prior knowledge learners are better able to implement the strategies and benefit from the deeper processing induced by them.

4.2.2 Overview of study and hypotheses

In this study, we investigated the effect of different instructional support measures for integrating multiple representations on learning outcomes. To this end, students learned with text-graph material that either offered no instructional support (control) or additionally included pre-given passive signals highlighting text-graph correspondences (passive signaling); alternatively, students were asked to actively highlight corresponding text-graph information during learning (active signaling). Furthermore, we analyzed how prior knowledge and the quality of learner-generated signals influence the effect of the instructional support on learning outcomes. All hypotheses were examined within two learning domains, namely, economics and biology, to test for the robustness of effects.

Based on the positive effects of passive signaling in multimedia learning studies (Alpizar et al., 2020; Richter et al., 2016; S. Schneider et al., 2018), we expected students to show higher learning outcomes from passive signaling compared to unsignaled materials (H_1; passive signaling effect).

Based on the positive effects of learner-centered instructional support (Bodemer et al., 2004), we furthermore expected students to show higher learning outcomes when they have to actively signal corresponding information compared to learning from unsignaled material (H_2; active signaling effect).

Moreover, as a first boundary condition, we investigated the role of prior knowledge. In the light of previous research (Arslan-Ari, 2018; Richter et al., 2016, 2018; Richter & Scheiter, 2019), we expected learners with low levels of prior knowledge to show a stronger passive signaling effect than learners with higher levels of prior knowledge (H_3). Finally, as learners with higher levels of prior knowledge are more likely to accurately implement an active-signaling strategy than those with lower levels of prior knowledge, we assumed that the active signaling effect would be stronger for the former than for the latter (H_4). We argue that a broader comprehension of the content is more relevant than factual knowledge of the material and thus, expected these effects to occur only for knowledge assessing prior comprehension (and not for prior recall knowledge).

As the second boundary condition for active signaling, we explored the role of signaling quality as a proxy for the quality of strategy implementation. Accordingly, within the active-signaling group, we expected the quality of learner-generated signals to be positively correlated with learning outcome in that learners who correctly connected text and graph would show higher learning outcomes than those creating false or no connections (H_5). Moreover, we explored whether the quality of strategy implementation would be positively related to the learners' prior knowledge, as this would further explain why learners with higher levels of prior knowledge might be more likely to benefit from active signaling (H_6). Accordingly, in a final step, we tested whether signaling quality would mediate the effects of prior knowledge on learning outcome (H_7).

4.3 Methods

4.3.1 Participants and design

Based on the two meta-analyses regarding signaling (Richter et al., 2016; S. Schneider et al., 2018), an *a priori* power analysis with a power of .80 and an estimated effect size of .25 was used to determine the necessary sample size of 179 for an ANCOVA[4]: 173 university students (129 females, M_{age} = 24.46 years, SD = 5.93) with different academic backgrounds participated in the experiment. They were paid to participate on an hourly basis. Following our pre-registration, we excluded students who studied economics or biology (4 participants) categorically, resulting in 169 participants for data analyses.

The experiment consisted of five phases: (1) Reading test and pretest, (2) learning phase for domain I (3) posttest domain I, (4) learning phase for domain II, (5) posttest domain II. For both learning phases, participants were randomly assigned to one of three between-subject learning

[4] We used (ANCOVA, df=3, Groups=3, 1 Covariate) in the preregistration since it is the closest approximation to multiple regression as both methods are based on the general linear model. For the main analysis, we only use regression to allow interaction with prior knowledge. The preregistration can be downloaded for review at: https://aspredicted.org/blind.php?x=4qv8yj

conditions: material was not signaled (n = 56 participants), up to three correspondences between text and graphs on each page of the learning materials were signaled (passive signaling condition; n = 54), or material was not signaled but participants were asked to highlight correspondences in the materials (active signaling condition; n = 59).

4.3.2 Materials

The learning material in economics dealt with the principles of supply and demand and consisted of seven pages, with an instructional text of between 81 and 144 words and a line graph on each page (in total, 794 words and 7 graphs). On the first two pages, the relationship between price and economic demand was introduced and the effect of additional factors such as household income was discussed. The next two pages showed the relationship between price and economic supply as well as other influencing factors, e.g., changes in prices for production factors. Lastly, price was modeled as result of supply and demand (equilibrium principle) and it was explained how prices are subject to periodic fluctuations in the cobweb-model (for an example page see Figure 1). The learning material was based on a textbook unit from the upper secondary level (Bauer, Hamm-Reinöhl, Podes, & Riedel, 2012).

The learning material in biology explained the principles of island biogeography and consisted of 6 pages, with an instructional text between 58 and 134 words and a line graph on each page (in total, 688 words and 6 graphs). First, the material introduced the relationship between the immigration rate of new species and the number of species already present on an island and other factors that influence the immigration, such as the distance of an island to the source region. The next two pages illustrated the effect of already existing species on extinction and it was stated that extinction is further dependent on factors such as the size of the island. On the fifth page, the equilibrium point was explained as the number of species, for which the immigration rate equals the extinction rate. As a conclusion, two effects that further influence immigration and extinction were shown (for an example page see Figure 1. The materials are based on an ecology college textbook (Begon, Howarth, & Townsend, 2017).

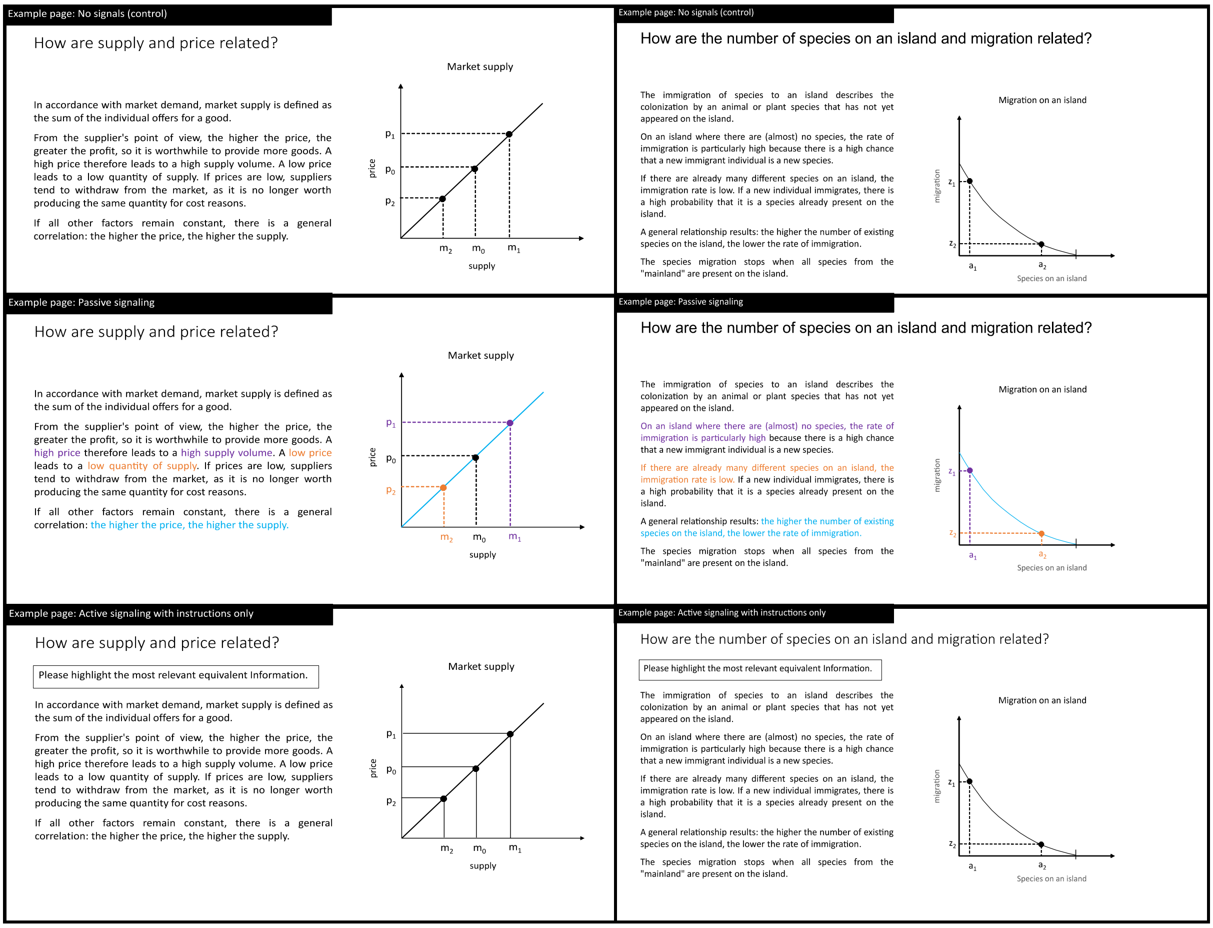


Figure 1. Example pages of the learning material (left side: economics, right side: biology; upper panel: control, middle panel: passive signaling, lower panel: active signaling)

In all three conditions, the text was structured in paragraphs and the graphs were supplemented with text labels (headings, labels for axis, labels for lines and single points). In the passive signaling condition, two or three correspondences in the text (e.g., a part of a sentence) and graph (e.g., the corresponding line) were color-coded. The signals focused on the central relationships between the different variables that are shown on the page. In total, there were 19 (economics) and 14 (biology) signals contained in the materials of the passive signaling condition.

4.3.3 Measures

4.3.3.1 Pretest: Reading comprehension, prior knowledge and motivation

In a first step, reading comprehension was assessed with a German standardized reading test (LGVT 6–12; W. Schneider, Schlagmüller, & Ennemoser, 2007): Learners read a 3.5 pages long text that contains 23 gaps. Within a time limit of four minutes, learners read the text and decided which of the potential words had to be filled in. Based on the number of correctly identified filler words, reading comprehension was determined (max. 23 points).

Prior knowledge was assessed as recall and comprehension knowledge with a subset of the items that were used in the posttest. For biology, we used 4 recall items and 3 comprehension items. In economics, 4 recall items and 4 comprehension items were used. The items were a mix of multiple-choice and open items (for examples see knowledge test section), therefore, for all 4 prior knowledge constructs, participants could achieve 5 points respectively. We report the ratio of achieved points to all possible points (percentage score) as predictors for the analysis.

Furthermore, subject-specific motivation for economics and biology was assessed (4 items per domain, e.g., “I like reading about biological topics” adapted from Wilde, Bätz, Kovaleva, & Urhahne, 2009). The items were answered on a 4-point Likert scale ranging from “I do not agree at all” to “I completely agree”. Cronbach’s alpha was 0.90 (biology) and 0.89 (economics), respectively.

4.3.3.2 Mental effort and task difficulty ratings

Directly after each learning phase, students were asked for their effort (“How much effort did you invest during the learning session?”, adapted from Paas, 1992) and the difficulty of the material (“How easy was it for you to follow the learning unit?”, adapted from Cierniak, Scheiter, & Gerjets, 2009). Both items were rated on a 9-point Likert-type scale ranging from 1 (no effort at all/very easy) to 9 (very much effort/ not easy at all).

4.3.3.3 Knowledge tests

For economics, the posttest contained 8 retention items and 8 comprehension items. Except for one item, which was taken from an economic literacy test (Walstad, Rebeck, & Butters, 2001), all items were created by the authors. The items were either multiple-choice items with one correct answer, three distractors and one “I don’t know”-category or open items. An example of an open recall item is: “Name two factors that influence supply in addition to demand and price”. Comprehension items asked participants to transfer the knowledge to a new situation, for instance: “Professional soccer players are paid better than farmers. Explain this dynamic with the help of the supply-demand-mechanism”. Multiple-choice items scored one point; free-response items scored one, two or three points, depending on the complexity of the question. A maximum of 11 points in recall and 10 points in comprehension could be achieved in economics.

The biology posttest consisted of 8 recall and 7 comprehension items. All items were created for this study. The items were either multiple-choice items or open items. An example of a multiple-choice recall item is “When does species migration to an island stop?” [(1) If all species on an island died out, (2) if migration rate and extinction rate is equal, (3) if extinction rate is higher than migration rate, (4) if all species from the mainland are on the island (5) I don't know]. Comprehension items asked them to transfer the knowledge to a new situation, e.g., "A traveler brings ten new species to an island, considering that the species are similar to the species that are already on the island, which of the following outcomes would you expect?" [(1) Extinction rate rises (2) In the long run there will be more species on the island (3) in the long run there will be fewer species on the island (4) Extinction rate drops (5) I don't know]. Again, multiple-choice items scored one point; free-response items scored one, two or three points, depending on the complexity of the question. A maximum of 10 points in recall and 9 points in comprehension could be achieved in economics.

All free-response items were rated by two trained raters. Across all open items, an average Fleiss' *Kappa* of 0.74 (*SD* 0.16) was obtained after the first rating. Two items with a kappa below 0.5 were rated again after a discussion, yielding a final average Fleiss' *Kappa* of 0.81 (*SD* 0.10). For all open items, the mean of the two ratings was used. For all knowledge tests, we report the ratio of achieved points relative all possible points (percentage score) as a dependent variable. For the percentage score measure, we excluded participants that answered less than 50% of the questions (between 1 and 3 participants depending on the test).

4.3.4 Quality of strategy implementation: signaling quality

Based on the assumption that signaling is effective because it helps learners to identify the relevant information in representations as well as equivalent information across multiple representations (van Gog, 2014), we rated the learner-generated signals in the active-signaling condition according to these two criteria: relevance and equivalence. For every signaled element, a participant received two points when all of the relevant information in the text or graph was highlighted and one point for most of the relevant information. The passive signaling condition served as a reference point in that relevant information was determined as the information that had been signaled in the respective material.

In addition to the points for relevance, a participant received two points if the highlighted information in text and graph was fully equivalent, and one point if it was mostly equivalent. Since finding connections (even if they were not the most central connections) between representations might help learners to integrate both representations, participants also received points for equivalence if they marked corresponding information in text and graph that was not relevant (e.g., the title of the graph in text and the title of the graph). Due to the fact that some items could be answered with information from one representation alone, we decided to award points for relevance (but not for equivalence) even if students applied highlighting in only one of the representations (only text or only graph).

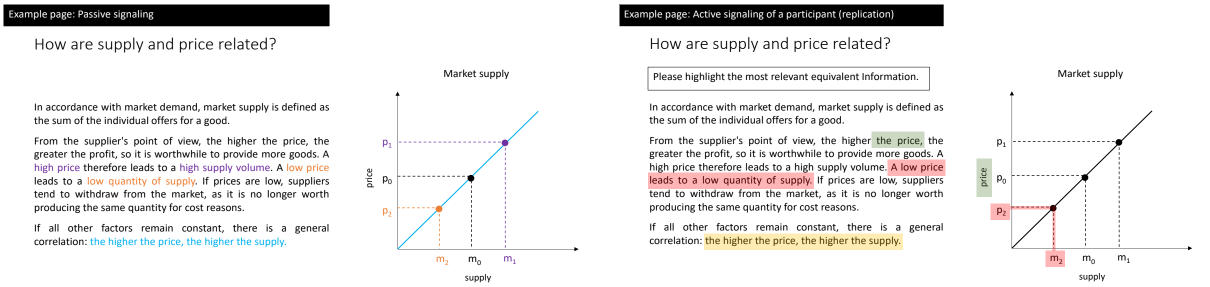


Figure 2. Example of learner-generated signals (left panel: material with passive signals, right panel: active signaling produced by a participant)

Figure 2 shows a reproduction of learner-generated signals (right panel) in comparison to the passive signal material page (left panel). In the right panel in Figure 2, the first signal (green color) refers to equivalent information, it is, however, not regarded as relevant as it is not the most important information, namely, the relationship between two variables (which is also signaled in the signaled condition, purple highlighting in the left panel). The participant, therefore, achieved two points for equivalence, but none for relevance. The second marking in red indicates fully equivalent information; moreover, it also highlights the relevant text information that had been also emphasized in the passive signaling condition. The learner received two points for equivalence as well as two points for relevance. The third signal in orange was only placed in the text and not in the graph. Since it is the most important information, the learner received two points for relevance, but none for equivalence. For the three signaled statements on this page, the learner earned a total of 8 out of 12 possible points. With four possible points for every signaled statement (19 in economics, 14 in biology), a participant could achieve a maximum of 76 points in economics and 56 points in biology. We report the ratio of achieved points relative to all possible points (percentage score) as signaling quality. Participants were excluded if more than half of all possible signals were missing (four participants in economics, three in biology). In a first rating, an average of Fleiss' *Kappa* of 0.58 (*SD* = 0.23) resp. of 0.56 (*SD* = 0.20) was obtained for all pages in the economic resp. biology learning material. After a first discussion, pages with a kappa under 0.5 were rated again and 0.69 (*SD* = 0.18) and 0.65 (*SD* = 0.14) could be achieved in economics and biology, respectively. The disputed cases were resolved in a second discussion among the two raters.

4.3.5 Procedure

All the study and test materials were paper-based. After the pretest (reading comprehension, prior knowledge, domain motivation), each participant completed the learning task and the corresponding test (difficulty, effort, recall, comprehension) for both domains, whereby the two domains were presented in random order. Up to 8 students participated per session, which took around 90 minutes. In all conditions, students were asked to study at their own speed (self-paced) and to note their start and end time for each learning phase. In the learning phase, before the actual learning material, one instructional page (with general and specific instructions) and one example page were shown. In the control and passive signal condition, students were asked not to use any tools (like a highlighter) to study the material. In the active condition, it was explained that while studying the material, it is the students' task to highlight (with 3 colors) the most relevant corresponding information in text and graph. The example page explained a basic principle related to the topic that was comparable to the actual material. While the example page was not signaled in the control condition, the example page consisted of text and graphs with three signaled correspondences, both in the passive signaling condition and in the active condition (to illustrate the principle of signaling to the students).

4.3.6 Data analyses

In a first step, we tested whether the groups were comparable regarding learning prerequisites. For this, reading comprehension, domain-specific motivation (for economics and biology), and prior knowledge were used as dependent variables in a MANOVA. Second, we analyzed differences between the experimental groups, namely, learning time, effort, and task difficulty in a MANOVA. Third, to analyze the effect of passive and active signaling (H_1, H_2) we used multiple regressions for every learning outcome measure (economics recall and comprehension, biology recall and comprehension). The experimental groups were dummy-coded (passive: 1/0, active: 0/1) with the control group as the reference point. Since we assumed a moderating effect of prior comprehension knowledge (H_3, H_4), we allowed for interactions with z-standardized prior knowledge scores. In the last step, we investigated signaling quality as a boundary condition of active signaling (H_5, H_6, H_7). In a mediation analysis, we addressed the question of whether the effect of prior knowledge is mediated via the signaling quality when students actively generate signals while learning. All analyses were conducted using the R-Studio software (R Core Team, 2020).

4.4 Results

4.4.1 Learning prerequisites

Table 1 shows the mean scores and standard deviations for the three groups regarding the learning prerequisites, namely, prior knowledge (recall and comprehension), domain motivation and reading comprehension. A MANOVA revealed no significant differences between the three groups (F (16, 294) = 0.49, p = 0.95). Hence, all groups were comparable concerning the relevant learning prerequisites.

Table 1

Means and standard deviations (in parentheses) for learning prerequisites

Economics	N	Reading Comprehension	Prior Knowl. Recall	Prior Knowl. Comprehension	Domain Motivation
Control	56	19.1 (7.0)	0.47 (0.3)	0.29 (0.2)	2.2 (0.7)
Passive Signals	54	19.5 (6.5)	0.45 (0.3)	0.33 (0.3)	2.3 (0.7)
Active Signals	59	19.4 (7.3)	0.53 (0.3)	0.34 (0.2)	2.1 (0.8)
Biology	N	Reading Comprehension	Prior Knowl. Recall	Prior Knowl. Comprehension	Domain Motivation
Control	56	19.1 (7.0)	0.11 (0.1)	0.34 (0.3)	2.9 (0.6)
Passive Signals	54	19.5 (6.5)	0.14 (0.2)	0.30 (0.3)	2.8 (0.7)
Active Signals	59	19.4 (7.3)	0.15 (0.2)	0.32 (0.3)	2.7 (0.8)

Note: Since reading comprehension was only assessed once and irrespective of the learning domain, values are the same for Economics and Biology.

4.4.2 Mental effort, task difficulty and learning time

Table 2 shows the mean scores of mental effort and difficulty ratings as well as learning time. Here, we found effects of the experimental groups on difficulty in biology [ANOVA; F (2,164) = 8.28, $p < .001$], as well as learning time in both domains [economics: F (2,162) = 44.93, $p < .001$, biology: F (2,164) = 39.5, $p < .001$]. Students rated the active condition as more difficult in biology and in both domains took almost twice the time compared to the other conditions. There were no differences between groups in self-reported effort for both domains.

Table 2

Means and standard deviations (in parentheses) for mental effort, difficulty and learning time

Economics	N	Effort (1-9)	Difficulty (1-9)	Time (in min)
Control	56	6.69 (1.36)	4.89 (1.99)	8.45 (3.40)
Passive Signals	54	6.75 (1.28)	5.06 (2.07)	8.31 (3.32)
Active Signals	59	6.86 (1.22)	5.69 (1.96)	15.30 (6.00)
Biology	N	Effort (1-9)	Difficulty (1-9)	Time (min)
Control	56	6.54 (1.24)	3.88 (2.02)	7.01 (2.93)
Passive Signals	54	6.74 (1.46)	4.50 (2.07)	6.72 (2.77)
Active Signals	59	6.40 (1.46)	5.39 (1.86)	12.40 (5.15)

4.4.3 Effect of passive signaling and active signaling on learning outcome measures (H_1, H_2)

The multiple regressions revealed no significant positive effect of passive signals in both domains (Table 3 and 4). Moreover, active signaling had no effect on learning outcomes in economics. There was, however, a negative effect of active signaling for biology – thus, in contrast to our hypothesis, students received fewer points in their posttests, when they actively highlighted correspondences between text and graph.

Table 3

Means and standard deviations (in parentheses) for learning outcome variables

Experimental group	N	Economics		Biology	
		Recall	Comprehension	Recall	Comprehension
Control	56	0.71 (0.18)	0.43 (0.19)	0.65 (0.18)	0.58 (0.20)
Passive Signals	54	0.70 (0.19)	0.47 (0.19)	0.62 (0.18)	0.52 (0.22)
Active Signals	59	0.67 (0.20)	0.41 (0.20)	0.59 (0.20)	0.48 (0.23)

Table 4

Multiple regression predicting learning outcomes with experimental groups (dummy-coded) and prior comprehension knowledge (z-standardized) [Coefficient-Estimates (SE)]

	Econ Recall	Econ Compr	Bio Recall	Bio Compr
(Intercept)	0.093 (0.13)	-0.011 (0.126)	0.235 (0.134)	**0.256*** **(0.126)**
Dummy: Signaled	-0.052 (0.184)	0.192 (0.179)	-0.268 (0.192)	-0.25 (0.18)
Dummy: Active Sign.	-0.241 (0.18)	-0.171 (0.177)	**-0.426*** **(0.19)**	**-0.442*** **(0.179)**
Prior knowledge – compr.	0.221 (0.146)	**0.382**** **(0.142)**	-0.14 (0.133)	**0.262*** **(0.125)**
Interaction: Prior knowledge × Passive signaling	0.033 (0.189)	-0.05 (0.184)	0.121 (0.193)	0.173 (0.181)
Interaction: Prior knowledge × Active signaling	0.262 (0.191)	0,085 (0.187)	**0.427*** **(0.188)**	0.197 (0.177)

* $p < .05$, ** $p < .01$, *** $p < .001$

4.4.4 Prior knowledge as moderator (H_3, H_4)

Regarding the assumed expertise reversal effect, there was no significant moderating effect of prior comprehension knowledge for passive signaling. Contrary to our hypothesis, students with low prior knowledge did not benefit more from already-signaled material. For active signaling, however, prior comprehension knowledge served as a significant moderator for biology recall outcomes: As hypothesized, high prior knowledge learners showed better recall in the active signaling condition than in the control condition regarding biology (see Figure 3, left panel).[5]

Following a significant interaction, we conducted simple slope analyses at -1 standard deviation and +1 standard deviation relative to the mean of the continuous variable (see Figure 3, right panel). This allows us to estimate the size of the effect of conditions by comparing different points of the continuous moderator variable (prior knowledge) without division of the sample in two groups (e.g., in a median split). Results show that for students with low prior knowledge (bars on the left), active signaling was counterproductive for biology recall (*Beta* = -0.85, $p < 0.01$), whereas signaling strategy did not matter for high prior knowledge learners (bars on the right, *Beta* = 0.001, $p = 0.99$). There was no comparable moderation for comprehension as a dependent variable.

[5] We also tested whether there are any interaction effects for prior recall knowledge (and a combined prior knowledge measure), this was not the case.

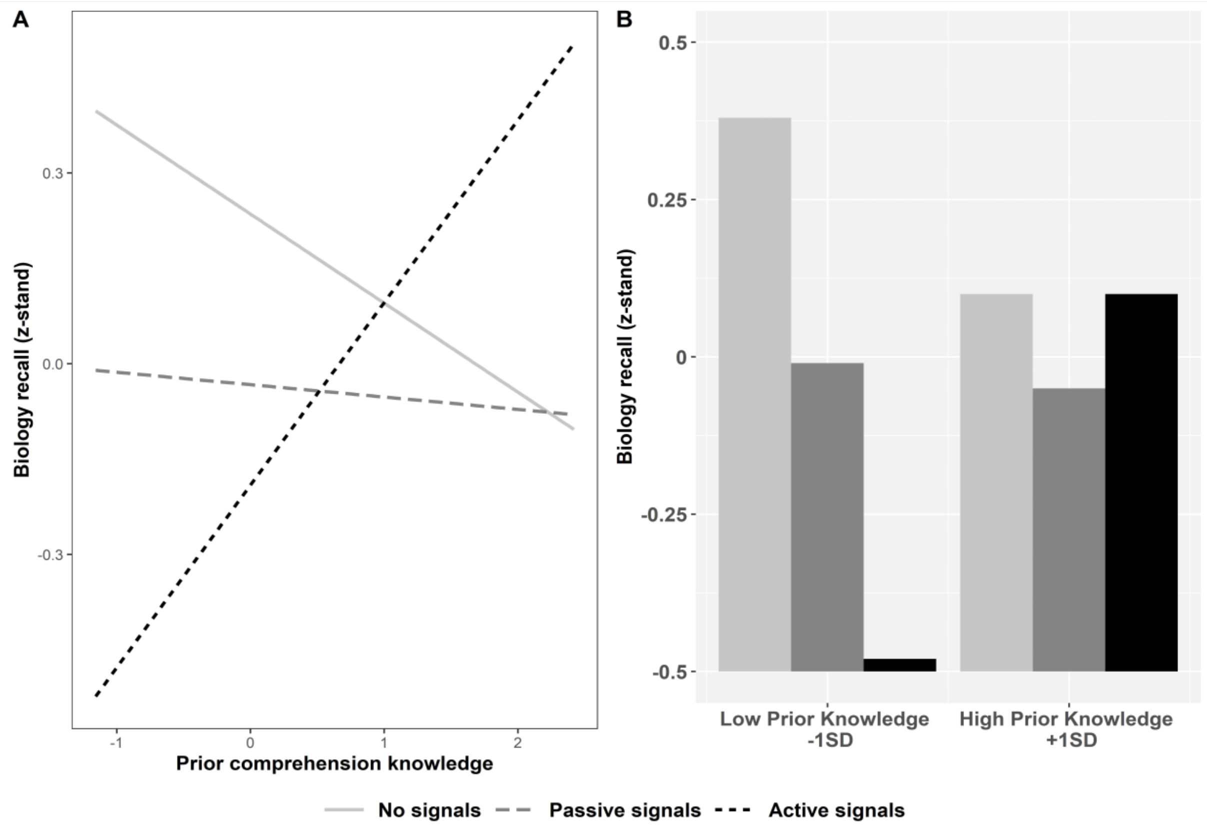


Figure 3. Interaction of prior comprehension knowledge and recall learning outcome in biology (left panel: continuous prior comprehension, right panel: low- and high prior knowledge learners)

4.4.5 Signaling quality as a boundary condition (H_5) and as a mediator for prior knowledge (H_6, H_7)

We assumed that signaling quality might be a boundary condition of active signaling. In total, the quality of signals was rather high in both domains. In economics, the participants received on average 74% (*SD* = 16) for marking the relevant and equivalent information on all pages. In biology, the participants achieved 84% on average (*SD* = 14). The quality of learner generated signals was significantly related to outcome performance in economics ($r(56) = 0.46$, $p < 0.001$ for recall outcomes, $r(52) = 0.49$, $p < 0.001$ for comprehension outcomes) but not related to learning outcomes in biology ($r(53) = -0.01$, $p = 0.93$ for recall outcomes, $r(52) = 0.05$, $p = 0.68$ for comprehension outcomes).

Therefore, we further investigated whether the effect of prior comprehension[6] knowledge was mediated by the signaling quality within the active signaling group for economic learning outcomes. For this, we used z-standardized scores for prior knowledge, signaling quality, and learning outcomes. We report unstandardized regression coefficient estimates (*B*) as they are

[6] We also tested whether prior recall knowledge is a mediator, but this was not the case.

the preferred metric according to Hayes (2013). In a first step, we confirmed that prior comprehension knowledge is a significant predictor for learning outcome, separated into recall (*B* = 0.46***, *SE* = 0.12) and comprehension (*B* = 0.48***, *SE* = 0.13). The supposed mediator (signaling quality) also shows isolated effects on recall (*B* = 0.46***, *SE* = 0.11) and comprehension (*B* = 0.49***, *SE* = 0.12) outcome. In a third step, we tested whether prior knowledge is a predictor for signaling quality, which was the case (*B* = 0.33*, *SE* = 0.14, path a).

We then used mediation analysis (see Figure 4) to test whether effects of prior comprehension knowledge would be mediated by signaling quality. According to the mediation, signaling quality affected recall (*B* = 0.36**, *SE* = 0.11, path b) and comprehension (*B* = 0.39**, *SE* = 0.12, path d) learning outcomes when controlling for prior knowledge. The indirect effect is the product of the effect of prior knowledge and the effect of signaling quality in the model (0.33×0.36 = 0.12 for recall and 0.33×0.39 = 0.13 for comprehension, path c/e). We tested the significance of this indirect effect using bootstrapping procedures as recommended in Hayes (2009). Unstandardized indirect effects were computed for each of 1,000 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 0.12 for recall (*p* = 0.018, the 95% confidence interval ranged from .01 to .28) and 0.13 for comprehension (*p* = 0.010, the 95% confidence interval ranged from .02 to .28). Thus, about 25% of the total effect of prior knowledge on both learning outcomes measures was mediated by signaling quality in economics.

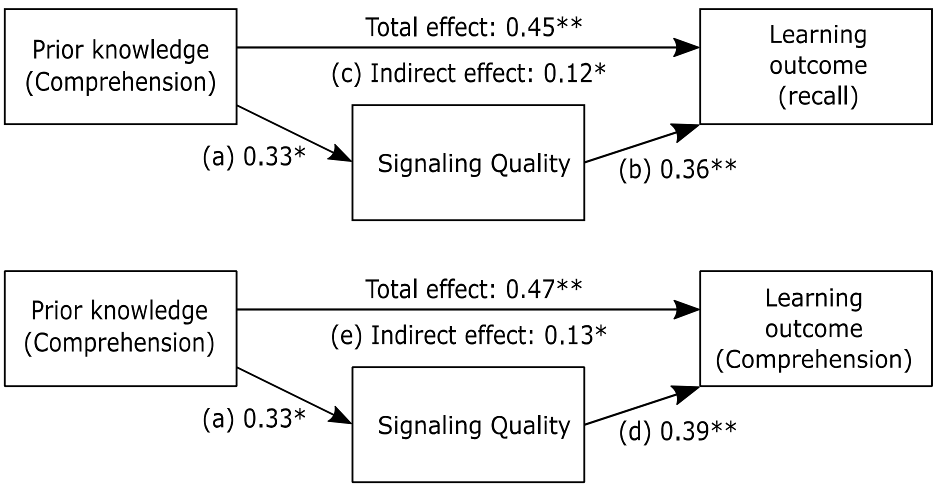


Figure 4. Mediation model regarding the effects of prior knowledge on learning outcomes (upper panel: recall; lower panel; comprehension) as mediated via signaling quality

4.4.6 Exploratory analysis: Differences between domains

Since we unexpectedly encountered differences between the domains, we decided to further analyze the differences in the learning material as a possible explanation. The biology material was not only shorter, but students also rated it as easier, $t(331) = 2.81$, $p < .01$, and were able to generate more relevant and equivalent correspondences between text and graph, $t(108) = -3.37$, $p < .01$. Although they did not invest more effort, $t(329) = 1.46$, $p = 0.1443$, they achieved higher comprehension scores $t(325) = 4.02$, $p < .001$ and spent less time with the material compared to economics, $t(319) = 3.45$, $p < .001$. Only regarding recall learning outcomes, the students achieved fewer points in biology than in economics, $t(333) = -3.66$, $p < .001$ (Table A.1 in appendix).

4.5 Discussion and conclusion

Contributing to the literature on instructional support for multimedia learning environments, this study focused on the effects of passive and active signaling on learning outcomes; furthermore, the role of prior knowledge and the quality of the learner-generated signals were investigated as potential boundary conditions.

Based on the positive effect of passive signaling in multimedia learning studies, we expected especially students with low prior knowledge to benefit when learning from signaled material (H_1 and H_3). However, in contrast to previous studies (Koning, Tabbers, Rikers, & Paas, 2007; Mautone & Mayer, 2007), learners independent of their prior knowledge did not profit when learning from material where correspondent information was highlighted. This finding is astonishing given the abundant evidence in favor of the signaling effect in learning from multiple representations especially for novice learners (Alpizar et al., 2020; Richter et al., 2016). One may wonder whether the lack of replication has to do with the fact that graphs rather than cause-effect diagrams or realistic depictions were used as learning material. Acarturk, Habel, Cagiltay, and Alacam (2008) found that signals aimed at supporting text-picture integration even hampered learning when applied to graphs; in contrast, Authors (submitted) found a positive effect of multimedia integration signals also in the case of text-graph combinations. To the best of our knowledge, these are the only two studies that have investigated the effects of providing signals to support integration of text and graphs so far. Thus, from these mixed results, it is too early to dismiss the differences between prototypical multimedia materials and the present ones as a possible explanation for not replicating the signaling effect. Moreover, in the present study, the learners were quite able to generate relevant and corresponding signals as suggested by the analysis of signaling quality in the active-signaling condition, which at least in economics predicted learning outcome. Hence, it might be that even for learners with low prior knowledge the pre-given signals were redundant to what they were able to infer from the materials on their own. These findings suggest that future research should address the relative effects of passive and active signaling to identify the conditions under which one or the other may be more helpful for learning.

Regarding the effects of active signaling, based on the ICAP framework (Chi & Wylie, 2014) and the positive effects of learner-centered instructional support (Bodemer et al., 2004), we had expected students to benefit from actively searching for and highlighting correspondent information (H_2). Although students took more learning time and rated learning as more difficult in the active-signaling than in the control condition, they did not benefit from active signaling. In contrast to our expectations, at an overall level, learners even achieved fewer points in the learning outcome tests in biology, whereas the differences were not significant in economics. However, a look at how the effects of active signaling are moderated by prior knowledge reveals that for learners with low prior knowledge, active signaling was counterproductive, which at least partly confirmed our assumptions (H_4). Interestingly, the quality of the signals that were produced by the learners was unrelated to learning outcomes in biology. Potentially, high prior knowledge learners took the instruction to generate signals as a prompt to reason about possible text-graph correspondences without necessarily externalizing them on paper. Maybe once they had identified the correspondences, they found them self-evident and did not bother to explicitly highlight them. This explanation, while clearly posthoc, would be in line with the finding that biology was seen as the easier domain compared to economics, where signaling quality was related to learning outcome.

The fact that signaling quality was related to learning outcomes in economics is in line with our assumption (H_5). It at least partly confirms previous findings on generative learning suggesting that the benefits of active learning strategies will depend on the quality with which the strategy is implemented (Schmeck et al., 2014). Additionally, within the active signaling group, signal quality could be identified as a mediator of prior comprehension knowledge in economics. One part of the effect of prior knowledge on learning outcome in economics (around 25%) is mediated via signaling quality. As hypothesized (H_6, H_7), students with higher prior knowledge were able to generate high-quality signals, which, in turn, increased their learning outcomes. This result could be interpreted as an empirical argument for the important role that prior knowledge plays in the generation of an integrated mental model as described both by the Cognitive Theory of Multimedia Learning and the Integrated Model of Text-Picture Comprehension (Mayer, 2014; Schnotz, 2014).

As mentioned previously, domain differences in the findings are possibly due to differences in the complexity of the learning material. Whereas the economics material was “difficult enough” that connecting text-and-graph was challenging (but effective if done right), the biology material was too easy compared to the posttest. Thus, almost all students were able to actively connect text and graph (as indicated by high signaling quality) but only students with high prior knowledge benefited from making these connections. Since the finding was unexpected and the materials were not designed to test this hypothesis, further research is warranted, which takes the complexity of learning material into account. Importantly, a recent study on another generative learning strategy, namely, learning by explaining, has yielded similar findings: Jacob, Lachner, and Scheiter (2020) showed that the instruction to explain a previously read text to a fictitious other student improved learning only when text complexity was high, but not when it was low. In general, such findings are well in line with Cognitive Load Theory (CLT) by suggesting that instructional manipulations aimed at fostering elaborative processing (i.e., stimulating Germane Load) pay off mostly when the to-be-learned content is of sufficient (intrinsic) complexity. Thus, in newer versions of CLT, the notion of Germane Load is given up and Intrinsic Load is seen both as a function of task complexity and the effort invested to handle high-complexity contents (Sweller, van Merriënboer, & Paas, 2019).

To conclude, in the present study on enhancing learning from text-graph comprehensions, we were unable to replicate previous benefits of providing signals to learners that highlight text-graph correspondences. Possible explanations refer to the assumption that graphs may demand different processing affordances than other types of depictions and to the overall ease of the materials that may have rendered support for text-graph integration unnecessary. However, we acknowledge that both of these explanations are posthoc and require further systematic investigations with materials purposefully designed to test these explanations. Moreover, asking learners to generate signals on their own proved ineffective; however, there is some evidence suggesting that these effects depend on the quality with which learners implemented the active-signaling strategy. Accordingly, future research should investigate whether students can be trained in their ability to use active signaling as a strategy and apply their prior knowledge to highlight information that is relevant to text-graph integration. Importantly, whether such a strategy is effective may furthermore depend on the material’s complexity, as the current study showed signaling quality to be strongly related to learning outcomes in the case of material that was presumably more difficult to learn from.

4.6 References

Acarturk, C., Habel, C., Cagiltay, K., & Alacam, O. (2008). Multimodal Comprehension of Language and Graphics: Graphs with and without annotations. *Journal of Eye Movement Research, 1*(3), 1–15. https://doi.org/10.16910/JEMR.1.3.2

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development*, *104*(1), 250. https://doi.org/10.1007/s11423-020-09748-7

Arslan-Ari, I. (2018). Learning from instructional animations: How does prior knowledge mediate the effect of visual cues? *Journal of Computer Assisted Learning, 34*(2), 140–149. https://doi.org/10.1111/jcal.12222

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*(2), 282–293. https://doi.org/10.1037/a0014312

Bauer, M., Hamm-Reinöhl, A., Podes, S., & Riedel, H. (2012). *Unterrichtswerk für die Oberstufe Wirtschaft: Märkte, Akteure und Institutionen [Textbook for higher grade level: Economics - Markets, Actors and Institutions]* (1. Auflage). *Kolleg Politik und Wirtschaft*. Bamberg: Buchner, C C.

Begon, M., Howarth, R. W., & Townsend, C. R. (2017). *Ökologie*. [Ecology] Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49906-1

Bell, K. E., & Limber, J. E. (2009). Reading Skill, Textbook Marking, and Course Performance. *Literacy Research and Instruction*, *49*(1), 56–67. https://doi.org/10.1080/19388070802695879

Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior*, *22*(1), 27–42. https://doi.org/10.1016/j.chb.2005.01.005

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*(3), 325–341. https://doi.org/10.1016/j.learninstruc.2004.06.006

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*(2), 315–324. https://doi.org/10.1016/j.chb.2008.12.020

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Fiorella, L., & Mayer, R. E. (2016). Eight Ways to Promote Generative Learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fiorella, L., & Zhang, Q. (2018). Drawing Boundary Conditions for Learning by Drawing. *Educational Psychology Review*, *30*(3), 1115–1137. https://doi.org/10.1007/s10648-018-9444-8

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. *Methodology in the social sciences*. New York, NY: Guilford Press. Retrieved from http://lib.myilibrary.com/detail.asp?id=480011

Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text complexity matters. *Learning and Instruction*, *68*. https://doi.org/10.1016/j.learninstruc.2020.101344

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Koning, B. B. de, Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology*, *21*(6), 731–746. https://doi.org/10.1002/acp.1346

Leopold, C., Doerner, M., Leutner, D., & Dutke, S. (2015). Effects of strategy instructions on learning from text and pictures. *Instructional Science*, *43*(3), 345–364. https://doi.org/10.1007/s11251-014-9336-3

Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, *99*(3), 640–652. https://doi.org/10.1037/0022-0663.99.3.640

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Ozcelik, E., Arslan-Ari, I. & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, *26*(1), 110–117. https://doi.org/10.1016/j.chb.2009.09.001

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

R Core Team (2020). *A language and environment for statistical computing (Version 3.4.4) [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Renkl, A., & Scheiter, K. (2017). Studying Visual Displays: How to Instructionally Support Learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4

Richter, J., & Scheiter, K. (2019). Studying the expertise reversal of the multimedia signaling effect at a process level: evidence from eye tracking. *Instructional Science*, *47*(6), 627–658. https://doi.org/10.1007/s11251-019-09492-3

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Richter, J., Scheiter, K., & Eitel, A. (2018). Signaling text–picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology*, *110*(4), 544–560. https://doi.org/10.1037/edu0000220

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why Sketching May Aid Learning From Science Texts: Contrasting Sketching With Written Explanations. *Topics in Cognitive Science*, *9*(4), 866–882. https://doi.org/10.1111/tops.12261

Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, *39*(4), 275–286. https://doi.org/10.1016/j.cedpsych.2014.07.003

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, *23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *LGVT 6-12: Lesegeschwindigkeits-und -verständnistest für die Klassen 6-12 [Reading speed and comprehension test for grades 6 to 12]*. Göttingen: Hogrefe.

Schnotz, W. (2014). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 72–103). 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

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*(2), 227–237. https://doi.org/10.1016/S0959-4752(02)00022-1

Seufert, T. (2019). Training for Coherence Formation When Learning From Text and Picture and the Interplay With Learners' Prior Knowledge. *Frontiers in Psychology*, *10*, 193. https://doi.org/10.3389/fpsyg.2019.00193

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Van der Meij, J., & de Jong, T. (2011). The effects of directive self-explanation prompts to support active processing of multiple representations in a simulation-based learning environment. *Journal of Computer Assisted Learning*, *27*(5), 411–423. https://doi.org/10.1111/j.1365-2729.2011.00411.x

Van Gog, T. (2014). The Signaling (or Cueing) Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 263–278). Cambridge: Cambridge University Press.

Walstad, W. B., Rebeck, K., & Butters, R. B. (2001). *Test of economic literacy*. New York: National Council on Economic Education.

Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM): [Testing a short scale of intrinsic motivation]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *15*, 31–45.

Yue, C. L., Storm, B. C., Kornell, N., & Bjork, E. L. (2015). Highlighting and Its Relation to Distributed Study and Students' Metacognitive Beliefs. *Educational Psychology Review*, *27*(1), 69–78. https://doi.org/10.1007/s10648-014-9277-z

4.7 Appendix

Table A.1

Differences between learning outcomes, effort, difficulty and time in economics and biology

	Recall learning outcome *Mean* (*SD*)	Comprehension learning outcome *Mean* (*SD*)	Difficulty (1-9) *Mean* (*SD*)	Effort (1-9) *Mean* (*SD*)	Time *Mean* (*SD*)
Economics	0.69 (0.19)	0.44 (0.19)	5.22 (2.02)	6.77 (1.28)	10.73 (5.50)
Biology	0.62 (0.19)	0.53 (0.22)	4.59 (2.07)	6.55 (1.39)	8.81 (4.61)
P-Value (t-test)	$p < 0.001$	$p < 0.001$	$p < 0.01$	n.s.	$p < 0.001$
Effect Size (Cohen‘s *d)*	0.40	0.44	0.31	-	0.38

Acknowledgement

We would like to thank the experts of the Institute of Evolution and Ecology in Tübingen for fact-checking the learning material as well as the items in biology. Furthermore, our gratitude extends to the research assistants who supported us in rating the open item as well as the quality of learner generated signals. Malte Ring is a doctoral student at the LEAD Graduate School & Research Network [GSC1028], which is funded within the framework of the Excellence Initiative of the German federal and state governments.

5 General discussion

The ability to understand graphical representations is an important skill in the 21st century as it enables students to comprehend (everyday) data (such as election results) as well as domain models (such as the goods market in economics). Large-scale and in-depth studies have demonstrated that learners show a variety of difficulties when they work with representations; these difficulties can be ascribed to the special affordances of graphical representations as well as the lack of understanding of the underlying (domain) principles (e.g., Boels, Bakker, van Dooren, & Drijvers, 2019; Glazer, 2011; Lai et al., 2016; McKenzie & Padilla, 1986).

In the first part of the dissertation, research regarding visual representations was summarized: first, from the perspective of the domain of economic education, where prior research has primarily focused on the understanding of macroeconomics through graphs and charts; second, from the perspective of graph comprehension models, where learners' ability to use graphical representations is modeled as an outcome of task, learner and representational characteristics; and lastly, from the perspective of multimedia and multiple representation research, where the research focus is on the effect of the learning environment design on learning outcomes. Especially for the domain of economic education, the latter two perspectives can offer important insights since they have not yet been comprehensively investigated. By integrating all three perspectives, the following can be analyzed: how (well) learners can work with representations, how representations are used in the domain and consequently how learners can be supported in understanding a domain through visual representations. These overarching dissertation objectives were addressed in three studies.

In Study 1, we closely examined the ability of learners to read graphs that are commonly used in textbooks, newspapers or on the internet and that visualize descriptive data. Then, in Study 2, we analyzed how graphs and charts are used in textbooks, and we described both the use of logical pictures and the relevant abilities and challenges that learners face from a teacher's point of view. Lastly, in Study 3, we examined how text–graph integration can be fostered by highlighting the central correspondences in text and graphs.

The overarching goal of this dissertation was to connect different research fields to better understand how learners and teachers use (multiple) visual representations and how they can be supported in that regard. Based on the literature, the influencing factors for the use of graphical representations were derived as learner characteristics, design characteristics and task characteristics. Each of the three studies contributed in part to an overall understanding of the use of graphical representations in economic education, focusing on the relationship between learner ability and task characteristics (Study 1), the influence of the domain to the form and task characteristics (Study 2) and the influence of design and task characteristics on the effectiveness of learning with graphical representations (Study 3).

In the following section, the results of the three studies are summarized and presented in a broader research context. Thereafter, the strengths and limitations of this dissertation are discussed. In the last two sections, implications of the findings for future research and educational practice are explored.

5.1 Discussion of general findings

5.1.1 Summary of findings

Study 1, which was designed to investigate the graphical literacy of secondary school students in the interdisciplinary field of sustainable development, revealed that eighth graders in the highest German school track are mostly able to read single data points and trends and can perform extrapolations. There was no systematic relationship between what an item asked for and the item difficulty for the learners in this sample. Furthermore, the ability to read graphs correlated with language abilities (performance in German class), math abilities (performance in math class) as well as content knowledge and prior engagement with sustainable development.

In Study 2, two empirical approaches were combined to analyze the use of representations in secondary economic education. First, graphical representations in textbooks were analyzed and categorized according to their form (graph/chart) as well as to the extent to which they visualize a domain principle. The categories were used to explain typical representations and to investigate their use in textbooks. Second, based on the conducted teacher interviews, graphs and charts are used regularly in teaching, not only to visualize economic principles but also to train students in critically analyzing graphical representations. Among the challenges for learners are math- and data-related issues (e.g., when learners are unable to differentiate between absolute and relative numbers) and a lack of integration of representation and domain (e.g., when learners cannot identify the relevant information for a domain question or are unable to connect the graphical information to other external representations such as texts).

Finally, in Study 3, a quasi-experimental design was used to analyze how learners can be supported in learning with text and graphs in two domains. Overall, neither signaled material nor an active signaling task promoted learning in economics and biology for all learners compared to the control group. Students with high prior knowledge, however, benefited from actively integrating both representations in biology. In economics, the relationship between prior knowledge and learning outcome was found to be partially mediated by the quality of learner-generated signals (i.e., learners with high prior knowledge were better at connecting graphs and text, which in turn was associated with higher learning outcomes).

5.1.2 Modeling graph literacy

Different scholars have modeled the ability to read and interpret graphs (Shah, Freedman, & Vekiri, 2005). Those models expand the understanding of where exactly learners succeed and where they struggle when encountering graphical representations. For this, researchers employ tests in which they present graphs and ask learners to perform typical operations, such as reading a data point. A graph can be used to answer different types of questions, which are distinguished according to how much a learner is required to abstract from the information seen in the graph. To this end, most researchers utilize a three-level framework: First-level or "read the data" questions require learners to directly read data points, while the second level – "read between the data" – is used to describe questions where learners need to identify trends, and the last level – "read beyond the data" – consists of questions where learners extrapolate from graphs (Curcio, 1987). In most empirical studies, the difficulty of an item increases with the question level (Curcio, 1987) but this is not always the case (Lachmayer, 2008). One problem with this framework is that researchers operationalize the levels differently; for example, a Level-3 question requires a small extrapolation for a data graph in one study (Curcio, 1987), whereas in another study, it requires connecting the graphical representation to the domain principle (Lai et al., 2016).

The first study in this dissertation extended this line of research by analyzing the ability of eighth graders of the highest school track (Gymnasium) in Germany to read authentic data graphs. In our study, we aimed to clearly operationalize the three levels, along with questions that could be answered with the graph alone and without the help of prior knowledge. Overall, the items were easily solved by the students, and no meaningful differences between questions that ask for single data points, trends or extrapolations were found. The central conclusion regarding the modeling of graphical literacy is that operationalizing a learner's task along specific requirements such as reading simple data points or trends might not be as relevant as previously assumed. If it is irrelevant for learners whether an item requires reading of single data points, trends or extrapolations (and all tasks are equally easy for eighth graders in the highest school track), then another – broader – framework is necessary to conceptualize the aforementioned ability for learners at the beginning of economic education.

In the first study, the graphicacy framework was chosen exactly for its unspecific (as in, domain-unspecific) perspective to allow general conclusions across domains regarding graph reading for visualizations of quantitative data used in textbooks and popular media. A central conclusion from the first study is that this “basic” ability is not a challenge for students at the beginning of economic education. To be able to conceptualize this ability more appropriately, however, it is necessary to have knowledge of how graphical representations are used in learning material and teaching. This was the focus of the second study.

5.1.3 The use of graphical representations in secondary economic education

Based on empirical evidence that learners have the relevant prerequisites for more complex graph tasks, an explorative perspective was used in the second study to analyze how graphical visualizations are used in secondary economic education. Research across domains indicates that learners have difficulties with connecting visual representations to more complex domain principles (Lai et al., 2016; Stieff, Hegarty, & Deslongchamps, 2011; Strober & Cook, 1992). For economic education, most research regarding visual representations has focused on the use of graphs and charts in higher education (e.g., Cohn, Cohn, Balch, & Bradley, 2001; Marangos & Alley, 2007). Concerning secondary education, studies have focused on the quality of visual representations from a design perspective (Aprea & Bayer, 2010) or the influence of format on the understanding of economic principles such as price-building (Jägerskog, 2020; Wheat, 2007). The latter studies have analyzed the use of visual representations from a top-down perspective (i.e., they have focused on how graphical representations should be used based on the effect they have on learning outcomes).

In contrast, the second study in this dissertation analyzed the use of graphical representations from a bottom-up perspective. For this purpose, a framework for typical representations used in economic textbooks based on 450 visualizations was developed and combined with teacher interviews. The findings from the textbook analysis suggest that graphs and charts differ in the degree to which they visualize a certain economic principle. This categorization helps to attribute different abilities and challenges to different typical representations. For example, the graph-reading ability assessed in Study 1 would be primarily relevant for data graphs, whereas conceptual difficulties with price-building as described by Jägerskog (2020) are usually analyzed with graphs or charts that visualize domain principles. Teachers describe visual representations as being highly important for economics and use them in multiple ways in their teaching. As previous research suggests, teachers not only utilize visualizations as part of explanations (Findeisen, 2017; Schopf, Raso, & Kahr, 2019) but also expect their students to be able to critically analyze, construct and translate between graphical representations and other external representations such as texts. In line with previous research from other domains, the teachers also describe the connection between graphical representation and the domain as challenging for students (Lai et al., 2016). Furthermore, critical evaluation seems to be difficult as learners are not able to evaluate the quality of a graph or to draw accurate general conclusions from the data (Aoyama, 2007). Although the teachers noted that learners sometimes struggle with "simple" data extraction as measured in Study 1; they claim that those errors are rare and likely a result of missing concentration rather than a lack of conceptual understanding. This is consistent with the results of the first study, which demonstrated that learners are generally able to read data points and trends.

All in all, the findings are in line with previous studies from other domains and confirm that – at least from a teacher's point of view – the difficulty of graph tasks is associated with the amount of abstraction that is necessary for the task. Thus, in secondary economic education, learners struggle when they are required to draw general conclusions and connect a graphical representation to a domain question. Connecting back to the concepts described in Section 1.3, these differences can be attributed to missing statistical literacy (failing to draw general conclusions) and a lack of representational competence (not being able to understand and use representations of the domain). The abilities and challenges described by the teachers might be used as an empirical foundation of a competence model that describes the ability to work with graphical representations in economic education (see Subsection 5.3.3).

5.1.4 Integrating text and graph

In various domains, research results indicate that integrating text and visual representation is one of the prerequisites for the effectiveness of a learning environment consisting of multiple representational formats. The integration of the representations, however, is often challenging for learners. Study 2 revealed that this is (at least from the teachers' perspective) also a challenge for learners in economic education.

To support learners in connecting text and pictures, different instructional support approaches have been used successfully in the past (Renkl & Scheiter, 2017). Especially design-centered support such as signaling has produced small overall benefits to learners (Richter, Scheiter, & Eitel, 2016), with a stronger effect for novices compared to experts (Richter, Scheiter, & Eitel, 2018). Findings from the third study contradict these claims as even for low prior knowledge learners, no effect of signaling was found in both domains. The outcomes of the active signaling group concerning biology learning outcomes – namely, that learners with high prior knowledge benefit, whereas overall, there is no effect – are in line with other studies that have analyzed the effect of learner-centered interventions (e.g., Seufert, 2019) and provide further support for the moderating role of prior knowledge regarding instructional support. The differences between the domains are also relevant when the relationship between the process data (quality of learner-generated signals), prior knowledge and learning outcomes is analyzed within the active signaling group. In biology, learners were able to find connections more easily, and the quality of their signals was not related to either prior knowledge or learning outcomes. Finally, in economics, a greater variance was observed in the quality of signals, and this is related to prior knowledge and learning outcomes as the effect of the former on the latter is mediated by the signal quality.

To summarize, the study adds to the research of multiple representations and can be used as an empirical argument for the closer examination of the boundary conditions of instructional support. According to the findings, not only learner characteristics, such as prior knowledge, but also material characteristics, such as the complexity of the learning material and the format of the graphical representation, must be taken into account when analyzing the effects of instructional support. As these factors might vary from one domain to another, the generalizability of instructional support across domains should be called into question.

5.2 Strengths and limitations

One outstanding strength of this thesis is the interdisciplinary integration of different research traditions concerning visual representations. Throughout the present dissertation, findings from multiple disciplines are used as an empirical and theoretical foundation to analyze the ability of eighth graders to read graphical representations, to discuss the use of graphical representations in secondary economic education and lastly to measure the effect of instructional support for the integration of text and graphical representations. Therefore, this thesis is (a small) part of a bridge between economic education, science education and educational psychology, and it integrates different perspectives concerning visual representations. This research project is one of the first attempts to thoroughly examine graphical representation in secondary (economic) education in Germany; and it might thus be used as a foundation for future research regarding this highly relevant topic.

Along with the comprehensive theoretical and empirical foundation of this work, a variety of empirical methodological approaches were used in the three studies to do justice to the different research interests, which are a result of the complex requirements of this interdisciplinary research field. The focus of the first study was on the influence of graph-reading task characteristics. Therefore, item response theory was used as it is the statistical approach which is currently used to connect item characteristics with item difficulty (Embretson & Reise, 2013). In the second study, mainly qualitative content analysis was used as it supported the exploratory nature of the research questions (Kuckartz, 2016). Lastly, a quasi-experimental study (Study 3) with a pre- and posttest design was conducted to test the effect of different instructional methods that aim to integrate text and graphical representations on learning outcomes. Both of these quantitative studies were based on appropriate sample sizes that met the requirements of the methods used. In addition, in the third study, a prior power analysis and the main hypothesis were preregistered to promote reproducibility and transparency (Nosek et al., 2019).

All in all, the use of different quantitative and qualitative methods is another major strength of this dissertation as it allowed to (partially) compensate for the disadvantages that result from the use of only a single method such as the restricted generalizability of qualitative research (Kelle, 2006). In the first study, the findings provided a baseline as they empirically confirmed that learners at the beginning of economic education can read graphs rather well; nevertheless, it was not able to identify where learners struggle with graphical representations. In the second study, we consequently extended the framework and identified core issues and challenges that learners faced (from a teacher's perspective), one of which – namely, the integration of graphical representation and text – was then further examined in a subsequent quantitative study. The thesis was thus able to partly overcome the limited transferability of qualitative research as well as the narrow framework and the problems concerning theory-building that resulted from the use of a quantitative methodology (Kelle, 2006).

Although all studies used state-of-the-art methodological approaches, each study comes with a set of limitations. In the first study, the focus was on the ability of eighth graders in the highest school track in south-western Germany, and the results should not be generalized to students of other age groups or lower school tracks. Although the dichotomous multiple-choice test format was appropriate to analyze the relationship between item characteristics and item difficulty according to item response theory, it might not be the best method to model the complex process of graph reading and interpretation (Berg & Smith, 1994). Since the items did not match the ability of the learners, because they were too easy, the analyzed relationships between the graph-reading ability with other measures should be interpreted with caution.

In the second study, the focus was extended to not only graphs but also charts based on the discourse around their use in economic education (M. Davies, 2011; Reingewertz, 2013; Wheat, 2007). Although this was helpful, as it allowed for a more holistic view of logical representations in textbooks, it is still limited to these representations since other realistic illustrations (e.g., pictures and cartoons) were not considered. Furthermore, the inter-rater reliability regarding the domain specificity category was moderate (Cohens Kappa of .58), indicating that the categories of domain specificity might not be as strictly defined and that the distinction between "everyday language" and "economic terminology" is not clear in many cases. Furthermore, the teacher interviews were based on a convenience sample of teachers that most likely overrepresented educators who are interested in the use of graphs and charts in their teaching.

Whereas the first two studies focused on students and teachers in secondary education in schools, the third study used secondary education material but analyzed the effect on recall and comprehension in the laboratory and for university students. The setting was chosen because it allowed for a clear implementation of the intervention as other influencing factors could be easily controlled. It is important to note, however, that effects might be overestimated compared to studies in the field (Hulleman & Cordray, 2009). University students were chosen as participants because learners in secondary education are usually underage, and recruitment is hence much more strongly regulated. As a result, it would not have been possible to achieve a sufficient sample size within a reasonable time period.

Lastly, although the multiple perspectives are a strength of this dissertation, the broad framework comes with a conceptual disadvantage as well. Each aspect of graphical representations in the dissertation (i.e., the learners' abilities, the use of visual representations in teaching or textbooks and instructional support for integration) could only be considered in one study each. While the studies partly build on one another, no study used measures or replicated findings from other studies directly. Therefore, all in all, by further integrating the studies, the results could enhance the respective discourses but there is still a need for further research to demonstrate the resilience of the findings and address every aspect in more detail.

5.3 Implications for future research

5.3.1 Investigating graph literacy

The first study revealed no differences between the different graph-reading levels for eighth graders in relation to the item difficulty. Future research based on the same framework but other participants, namely, students from lower grades and/or students from other secondary school tracks, would be necessary to validate these findings for a more general learner population. As other studies find a three-level structure, it might be possible that the difference becomes evident for learners with an overall lower level of graphical literacy. Furthermore, since time on task is negatively related to item difficulty for reading tasks in large-scale assessments

(Goldhammer et al., 2014), it should be included in future research as an additional measure. As an alternative to the dichotomous items, future research should consider using free-response items, think-aloud studies and eye-tracking. All of these considerations would allow a researcher to stay within the framework of graph literacy where the ability to read graphs is independent of domain knowledge. Moreover, instead of following this approach, it would further be possible to develop a competence model where the ability to read graphs is analyzed as part of a larger graph competence in the economic domain (see below, Subsection 5.3.3)

5.3.2 The use of graphical representations in secondary economic education

Due to the focus on logical pictures in the textbook analysis of the second study, other representations such as text, tables, equations, pictures and drawings were excluded from the analysis. Future research could consider including pictorial representations as this would allow for a comparison between the use of realistic and logical representations. Furthermore, a more holistic approach could integrate a measure for the ratio of text and pictorial elements, as done by Slough and McTigue (2013) for science textbooks. Based on the moderate inter-rater reliability regarding the domain specificity category, future research could further develop the systematic approach, for example, by a clearer definition of the categories.

The teacher interviews provided important insights based on how teachers described the use of graphical representations in the classroom and the typical abilities and challenges of their students when working with visual representations in the classroom. As the self-reported usage might be biased by social desirability or inaccurate self-estimation, future research should further examine the actual use of graphical representations with the help of classroom observation. Similarly, the abilities and difficulties identified by the teachers should be extensively investigated, also by including students as respondents. For this purpose, a think-aloud study, where students are confronted with typical graphical representations and domain tasks, might be an appropriate method to validate the findings of the second study. Lastly, to date, the role of graphical representations in teacher explanations and, associated with this, the relevance and extent of teachers' graph competence have not been closely examined and would be another fruitful area for further work.

5.3.3 Modeling and improving graph competence in economic education

Compared to science and math education, where different competence models have been developed concerning the abilities to use graphs (e.g., Lachmayer, 2008), a coherent competence model in economic education is still lacking. Based on the studies in this dissertation, it is evident that understanding graphical representations is part of domain expertise and that graphs might be the threshold to understanding economic principles (P. Davies & Mangan, 2007). Based on the first two studies in this dissertation, future research should therefore consider the development of a graph competence model in the economic domain. The task requirements should then encompass not only reading, but also interpretation, evaluation and construction of graphs. For each aspect of graph competence, it would be necessary to create items that range from a low level of abstraction to those that focus on the relationship between visual representation and domain. For example, graph reading and interpretation should involve not only items that require the reading of single data points or trends but also items in which learners are required to more generally interpret the graph for a domain question (e.g., by providing participants with new information in the question stem, which they must connect to the graphical representation). For this purpose, graphs that visualize domain principles, such as diminishing returns; break-even analysis; or the relationship between supply, demand and price, could be used. In the development of this instrument a useful intermediate step might be the use of the mentioned think-aloud studies, where learners are confronted with the typical graph tasks that teachers described in Study 2. Findings regarding a learner's ability that are based on the teachers' statements could subsequently be empirically validated, and typical misconceptions could be used as distractors in multiple choice items.

In a next step, such an instrument could be used to analyze the relationship between graph competence and learner characteristics such as domain knowledge. Furthermore, the first study revealed relationships between a learner's ability to read graphs and interest as well as motivational constructs. Since the graphicacy test was not a suitable fit for the eighth graders' relatively high graph literacy, these relationships could be analyzed further with the help of a more reliable measure. Furthermore, such a measure would allow for the investigation of graph competence development and could thus be used to analyze the effects of interventions that aim at fostering graph competence.

Such an intervention could, for example, involve a training that supports learners in understanding graphical representations by practicing the reading rules of the visualizations in the domain, thus transferring the approach of Cromley et al. (2013) to economics. For this purpose, it would be necessary to study the graphs and charts that were categorized as highly domain-specific in the second study. Based on these graphs and charts, the next step could be to identify the use of typical visual elements, for example the use of arrows to denote the exchange of goods and services in the circular flow of income model, or the use of parallel shifts to visualize the effect of different factors in the equilibrium graphs. These findings could then be combined to develop a training where the domain-specific conventions are explained to students, and the effect could be tested in an intervention study in the field.

5.3.4 The effect of (active) signaling

The third study in this dissertation concluded that neither passive nor active signaling is overall beneficial for learning outcomes. The former finding was particularly unexpected considering that passive signals are usually effective (Richter et al., 2016; Schneider, Beege, Nebel, & Rey, 2018). Material complexity and the format of a graphical representation were already discussed as possible explanations in Chapter 4, and both should be addressed in future research as potential boundary conditions of the signaling effect. Within the active signaling group, the quality of the learner-generated signals was related to learning outcomes in economics (and mediated the relationship between prior knowledge and learning outcomes). A possible explanation for a missing main effect of active signaling is that especially learners with low prior knowledge were overwhelmed because the task combined learning new content and using an unfamiliar strategy.

This hypothesis could be tested with two possible designs in future research. First, a training phase could be added, where learners have the opportunity to use the active signaling strategy with familiar content prior to the actual testing material with unfamiliar content. When such a training phase is used, it would further be informative to assess long-term outcomes not only regarding recall and comprehension (with a follow-up test) but also to check whether learners use the strategy when they are not prompted (with follow-up learning material without prompts). As an additional process measure, eye tracking could be used to study the effect of the training on gaze behavior. Second, research could investigate whether the active signaling strategy is more helpful when the most important information is already highlighted in one representation. This might reduce the necessary cognitive capacities, as learners would not have to decide which part of the representation is important and could focus solely on how this information is presented in the other representation. Since text might be the guiding representation in text–picture integration (Schnotz & Bannert, 2003), it would be reasonable to use text signals and ask learners to highlight the information in the graphical representation. Lastly, based on the result that high prior knowledge learners benefited from the intervention in biology, it should be investigated, whether the active signaling technique might be an effective rehearsing strategy for material that consists of text and visual representation.

As the connection between representation and domain (knowledge) played a dominant role in all studies, an overarching conclusion is that future research regarding graphical representation should overcome the focus on surface features and simple graph operations (i.e., on the ability to read single data points). Instead, it seems important to further model the ability to use graphs based on the relationship between visual representations and the domain. This is in line with the reasoning from representational competence research, where researchers argue for a domain-specific modeling of the ability to work with (visual) representations in the science domains (Kozma & Russell, 2005; Nitz, Ainsworth, Nerdel, & Prechtl, 2014; Stieff et al., 2011).

5.4 Implication for practice and policy

As already discussed, further research is necessary to replicate and extend the results of this dissertation. Nevertheless, different implications for educational policy and practice can be derived from the findings. Graphs and charts are frequently used to visualize data, ideas and models, and they have the potential to support learners in understanding complex topics. They are important in everyday life, for example when they visualize unemployment rates or the development of share prices, and more complex graphs and charts are used as visual representations of domain models across multiple disciplines. The understanding of graphical representations is therefore important for students not only because it allows them to make sense of typical visual representations in different media (e.g., by being able to understand a graph in a newspaper or on the internet) but also because understanding the visual representations of a domain is associated with domain expertise (P. Davies & Mangan, 2007; Kozma & Russell, 1997).

As graphical representations are not always intuitive, it is reasonable to maintain the comprehension of graphical displays in the school curricula and to systematically develop a learner’s ability during the course of secondary education. This dissertation demonstrates that learners in the highest school track in Germany at the beginning of economic education are well versed in reading simple data graphs, but they have difficulties with graphs that visualize more complex domain principles in economics. One conclusion for practice and policy is that the comprehension of graphical displays should be taught as part of economic education and special emphasis should be placed on visualizations of domain-specific principles. When working with these graphical representations, teachers should not expect their students to understand them without support and, therefore, should take the time to explain how the visual representation is used within the economic domain. Typical misconceptions and difficulties that learners have when using graphs and charts in the economic domain could (after a more thorough investigation) be taught and discussed in teacher training as part of pedagogical content knowledge. Fundamental knowledge of learners’ misconceptions could give teachers an indication of which visual representation and which accompanying task might require special attention in their teaching.

Regarding the design and selection of learning material consisting of text and graphs, for example in textbooks, a general recommendation can be made based on the findings of multiple representation research: To aid learners in the integration of text and visual representations, text-graph material should be designed in a way that text and pictures are in close proximity to each other (Mayer & Fiorella, 2014). Moreover, the effects of using signals, where central information is highlighted in different representations, is overall positive (Richter et al., 2016); therefore, even though no effect was visible for the sample in this dissertation, signals should be used in learning material in economic education. Further recommendations for the design of graphical representations can be found in Aprea and Bayer (2010) and Oestermeier and Eitel (2008). However, as prompting learners to actively integrate different representations has yielded mixed results both in earlier research (Renkl & Scheiter, 2017) and in this dissertation, additional research should be conducted before clear recommendations for practice can be made.

All in all, this dissertation contributed to the understanding of the use of graphical representations in secondary economic education. With multiple methodological approaches and interdisciplinary foundations, it underlines the importance of those representations and analyzed the connection between domain and visual representation. It lays a strong foundation that future researchers can build on when they seek to analyze how graphical representations can be used to their full potential in economic education and beyond.

5.5 References

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education, 2*(3), 298–318.

Aprea, C., & Bayer, D. (2010). Instruktionale Qualität von grafischen Darstellungen in Lehrmitteln: Kriterien zu deren Evaluation: [Instructional quality of graphical representations in teaching materials: Criteria for their evaluation]. *Beiträge Zur Lehrerinnen- Und Lehrerbildung*, *28*(1), 73–83.

Berg, C. A., & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments. *Science Education, 78*(6), 527–554. https://doi.org/10.1002/sce.3730780602

Boels, L., Bakker, A., van Dooren, W., & Drijvers, P. (2019). Conceptual difficulties when interpreting histograms: A review. *Educational Research Review*, 100291. https://doi.org/10.1016/j.edurev.2019.100291

Cohn, E., Cohn, S., Balch, D. C., & Bradley, J. (2001). Do Graphs Promote Learning in Principles of Economics? *The Journal of Economic Education*, *32*(4), 299–310. https://doi.org/10.1080/00220480109596110

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving Students' Diagram Comprehension with Classroom Instruction. *The Journal of Experimental Education*, *81*(4), 511–537. https://doi.org/10.1080/00220973.2012.745465

Curcio, F. R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, *18*(5), 382–393. https://doi.org/10.2307/749086

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? *Higher Education*, *62*(3), 279–301. https://doi.org/10.1007/s10734-010-9387-6

Davies, P., & Mangan, J. (2007). Threshold concepts and the integration of understanding in economics. *Studies in Higher Education, 32*(6), 711–726. https://doi.org/10.1080/03075070701685148

Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory for Psychologists*. *Multivariate applications series*. Hoboken: Taylor and Francis.

Findeisen, S. (2017). *Fachdidaktische Kompetenzen angehender Lehrpersonen: Eine Untersuchung zum Erklären im Rechnungswesen (Pedagogical content knowledge of preservice teachers: an empirical study of explanations in accounting)* [Doctoral Dissertation, Universität Mannheim]. Wiesbaden: Springer Fachmedien Wiesbaden.

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, *47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, *106*(3), 608–626. https://doi.org/10.1037/a0034716

Hulleman, C. S., & Cordray, D. S. (2009). Moving From the Lab to the Field: The Role of Fidelity and Achieved Relative Intervention Strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88–110. https://doi.org/10.1080/19345740802539325

Jägerskog, A.-S. (2020). *Making Possible by Making Visible: Learning through Visual Representations in Social Science* [Doctoral Dissertation, Stockholm University]. Stockholm: Department of Humanities and Social Sciences Education, Stockholm University.

Kelle, U. (2006). Combining qualitative and quantitative methods in research practice: purposes and advantages. *Qualitative Research in Psychology*, *3*(4), 293–311. https://doi.org/10.1177/1478088706070839

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*(9), 949–968.

Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representational Competence. In J. K. Gilbert (Ed.), *Models and Modeling in Science Education: Vol. 1. Visualization in Science Education* (pp. 121–145). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3613-2_8

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung [Qualitative Content Analysis: methods, practice, computer support]* (3., überarbeitete Auflage). *Grundlagentexte Methoden*. Weinheim, Basel: Beltz Juventa.

Lachmayer, S. (2008). *Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht [Development and validation of a graph competence model for biology education]* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. Retrieved from https://d-nb.info/1019667389/34

Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R., & Linn, M. C. (2016). Measuring Graph Comprehension, Critique, and Construction in Science. *Journal of Science Education and Technology*, *25*(4), 665–681. https://doi.org/10.1007/s10956-016-9621-9

Marangos, J., & Alley, S. (2007). Effectiveness of concept maps in economics: Evidence from Australia and USA. *Learning and Individual Differences*, *17*(2), 193–199. https://doi.org/10.1016/j.lindif.2007.03.003

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 279–315). Cambridge: Cambridge University Press.

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (togs). *Journal of Research in Science Teaching*, *23*(7), 571–579. https://doi.org/10.1002/tea.3660230702

Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction*, *31*, 13–22. https://doi.org/10.1016/j.learninstruc.2013.12.003

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van't Veer, A. E., & Vazire, S. (2019). Preregistration Is Hard, And Worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818. https://doi.org/10.1016/j.tics.2019.07.009

Oestermeier, U., & Eitel, A. (2008). *Lernen mit Text und Bild [Learning from text and pictures]*. Retrieved from https://www.e-teaching.org/etresources/media/pdf/langtext_2014_oestermeier-uwe_eitel-alexander_lernen-mit-text-und-bild.pdf

Reingewertz, Y. (2013). Teaching macroeconomics through flowcharts. *International Review of Economics Education*, *14*, 86–93. https://doi.org/10.1016/j.iree.2013.10.004

Renkl, A., & Scheiter, K. (2017). Studying Visual Displays: How to Instructionally Support Learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Richter, J., Scheiter, K., & Eitel, A. (2018). Signaling text–picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology, 110*(4), 544–560. https://doi.org/10.1037/edu0000220

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, *23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

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

Schopf, C., Raso, A., & Kahr, M. (2019). How to give effective explanations: Guidelines for business education, discussion of their scope and their application to teaching operations research. *RISTAL-Research in Subject-Matter Teaching and Learning*, *2,* 32–50. https://doi.org/10.23770/rt1823

Seufert, T. (2019). Training for Coherence Formation When Learning From Text and Picture and the Interplay With Learners' Prior Knowledge. *Frontiers in Psychology*, *10,* 193. https://doi.org/10.3389/fpsyg.2019.00193

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The Comprehension of Quantitative Information in Graphical Displays. In P. Shah & A. Miyake (Eds.), *Cambridge Handbooks in Psychology*. *The Cambridge Handbook of Visuospatial Thinking* (pp. 426–476). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.012

Slough, S. W., & McTigue, E. (2013). Development of the Graphical Analysis Protocol (GAP) for Eliciting the Graphical Demands of Science Textbooks. In M. S. Khine (Ed.), *Critical analysis of science textbooks: Evaluating instructional effectiveness* (pp. 17–30).

Dordrecht, Heidelberg, New York, London: Springer. https://doi.org/10.1007/978-94-007-4168-3_2

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction*, *29*(1), 123–145. https://doi.org/10.1080/07370008.2010.507318

Strober, M. H., & Cook, A. (1992). Economics, lies, and videotapes. *The Journal of Economic Education*, *23*(2), 125–151.

Wheat, D. (2007). The feedback method of teaching macroeconomics: is it effective? *System Dynamics Review*, *23*(4), 391–413.

Declaration on Contributions to Monography

Although the dissertation is written as a monography, it includes contents of three manuscripts which were written together with other authors. The proportional contributions to the manuscripts are presented in the subsequent tables.

Chapter 2 Do difficulty levels matter for graphical literacy? A performance assessment study with authentic graphs

Author	**Author position**	**Scientific ideas %**	**Data generation %**	**Analysis & interpretation %**	**Paper writing %**
Malte Ring	first	60	100	80	70
Taiga Brahm	second	20	0	20	20
Christoph Randler	third	20	0	0	10
Status in publication process:		Published			

Chapter 3 Logical pictures in secondary economic education: Textbook analysis and teacher perception

Author	**Author position**	**Scientific ideas %**	**Data generation %**	**Analysis & interpretation %**	**Paper writing %**
Malte Ring	first	80	100	80	80
Taiga Brahm	second	20	0	20	20
Status in publication process:		Accepted			

Chapter 4 How to support text-graph integration: Comparing the effects of passive and active signaling on learning outcomes

Author	**Author position**	**Scientific ideas %**	**Data generation %**	**Analysis & interpretation %**	**Paper writing %**
Malte Ring	first	70	100	70	60
Taiga Brahm	second	10	0	10	10
Juliane Richter	third	10	0	10	10
Katharina Scheiter	fourth	10	0	10	10
Christoph Randler	fifth	0	0	0	10
Status in publication process:		Under review			