

RESEARCH

Open Access



# Characteristics of learning tasks in accounting textbooks: an AI assisted analysis

Simone Stütz<sup>1\*</sup>, Florian Berding<sup>2</sup>, Sven Reincke<sup>2,3</sup> and Lena Scheper<sup>2,3</sup>

\*Correspondence:  
simone.stuetz@jku.at

<sup>1</sup> Institute of Business and Vocational Education and Training, Johannes Kepler University of Linz, Altenberger Straße 69, 4040 Linz, Austria

<sup>2</sup> Institute of Vocational and Business Education and Training, University of Hamburg, Hamburg, Germany

<sup>3</sup> Carl Von Ossietzky University of Oldenburg, Oldenburg, Germany

## Abstract

Tasks in accounting textbooks play a vital role when it comes to learning processes. However, hardly any empirical evidence on the quality of accounting tasks exists regarding accounting-relevant characteristics. This is why a new category system containing accounting-relevant aspects was developed to analyze a total of 3,361 tasks from 14 different German accounting textbooks. Descriptive analysis and correlation analysis were performed to assess task characteristics and identify relationships between categories. In addition, in light of the large number of tasks to be analyzed, AI assisted the content analysis, and its usefulness was evaluated. The results indicate that tasks are not sufficiently able to instill accounting competencies such as interpreting data, assessing the relevance of information, or identifying and solving underlying accounting problems. The findings further show that AI and human coding yield similar results in most categories, suggesting that AI assistance is useful for content analysis when evaluating a large number of tasks.

**Keywords:** Learning tasks, Accounting textbooks, AI analysis, Accounting education, Modeling cycle, Process orientation, Problem-based learning

## Introduction

“Accounting information is central to the functioning of international capital markets and to managing small businesses, conducting effective government, understanding business processes, and raising and addressing questions about how economic decisions are made” (The Pathway Commission 2012: 21). Accounting can be called “the language of business” (Weil et al. 1999: 3). These statements point out the general relevance of accounting, as well as accounting education in particular. Accounting education is considered a core element of commercial education (Preiß 2015; Weil et al. 1999). In recent years, scholars and public institutions have suggested the need for modification of accounting education due to wide-ranging changes in accounting-related work environments (e.g. Jordanski 2020; Klein and Küst 2020), resulting in a need to shift accounting education from a “preparer orientation” towards a “user orientation”<sup>1</sup> (Albrecht and Sack 2000; Chiang et al. 2014; McKinney et al. 2017; Rebele and St. Pierre 2019; Stanley

<sup>1</sup> The traditional “preparer approach” focuses on the training of procedures and rules of accounting by recording and documenting transactions, whereas the “user approach” emphasizes the overall impact of business transactions on various aspects such as a company’s equity or assets (Chiang et al., 2014, p. 43 et seq.).

and Marsden 2012). Accountants perform a wide variety of functions, including advising businesses and supporting management in planning, understanding, and evaluating business operations. Therefore, up-to-date accounting education not only focuses on learners' abilities to correctly and lawfully report business transactions, but enables them to obtain extensive competencies to adequately use complex accounting systems, gain broad economic understanding, and transfer their knowledge to real business situations (Achtenhagen 1996; Preiß 2015; Preiß and Tramm 1996; The Pathway Commission 2012). Scholars globally emphasize the importance of problem-based learning in accounting education, enabling students to analyze and solve complex real-world problems (Dockter 2012; Hansen 2006). The Pathway Commission created by the American Accounting Association to develop recommendations for the future of accounting education defines problem solving as a "professional foundational competency" (2015: 18), while the Accounting Education Change Commission (1992: 250) states that learners should possess "the ability to confront unstructured problems." In the German-speaking realm, related claims of fostering modeling competencies in accounting education have been the subject of a broader debate (e.g. Berding et al. 2019; Guggemos 2016; Helm 2016). Adapting the modeling cycle from mathematics didactics to accounting, competent accountants need to pass through four critical steps to solve complex accounting-related problems. First, they need to translate business-related situations into an "accounting model" by simplifying and structuring the situation, detecting underlying problems, correctly analyzing documents and, if necessary, searching for additional information. In a next step, mathematical and accounting operations are applied which lead to accounting results. These then need to be re-transferred to the real world by interpreting and validating the results with respect to real business situations. This process is called a modeling cycle, and requires the competencies to translate, operate, interpret, and validate (Berding et al. 2019; Blum 2011). Another recent demand is the focus on business processes and value chains that enable learners to gain a "broad and interdisciplinary view of the work environment" (Walker and Ainsworth 2001: 41).

One important way to foster educational reforms is by assessing and adapting learning materials. Accounting textbooks have a major influence on accounting education, as educators rely extensively on textbooks as instructional tools as well as resources for learning tasks (Tramm and Goldbach 2005). In the past, research focused on the readability (Davidson 2005; Ernst 2011), content (Blix et al. 2021; Ernst 2012; Ferguson et al. 2010; Golyagina and Valuckas 2016; Laksmana and Tietz 2008; Wells 2018), the kind of knowledge presentation and acquisition (Berding and Lau 2018), and the ideological character (Ferguson et al. 2006; van der Kolk 2019) found in accounting textbooks. Studies on the characteristics of accounting tasks in textbooks are generally limited, even though it is widely recognized that learning processes are largely driven by tasks. Second, existing research has mainly used cross-domain criteria to assess accounting tasks rather than the criteria that are especially relevant in and drawn from accounting education research. Third, prior studies analyzed a relatively small number of tasks (e.g. by focusing on tasks from specific book chapters or particularly difficult topics), leading to limitations in terms of representativity.

This study aims to achieve three main research objectives. First, by applying domain-specific categories, it sheds light on task characteristics from an accounting education

perspective. Second, this study evaluates a large number ( $N=3361$ ) of tasks from 14 different German-language accounting textbooks using artificial intelligence (AI) assistance, leading to a more solid and broader data basis for further analysis. In addition, this study aims at critically evaluating the usability of artificial intelligence for the content analysis of (accounting) tasks.

This paper is organized as follows. Chapter 2 provides a theoretical background by discussing relevant task characteristics in accounting education and outlining existing research. Chapter 3 presents the research questions. Chapter 4 gives information on the selected textbooks, outlines the coding system used for content analysis, and describes the process of training and applying AI. Chapter 5 presents the results of statistical analyses. Chapter 6 discusses the study's main findings and limitations.

### **Essential characteristics of tasks in accounting**

Because this study was conducted in Germany, some introductory clarifications of the German educational system will be helpful. The German vocational training system combines theory and training in a real-life work environment.<sup>2</sup> One specific characteristic of its so-called “dual system” is the statutory cooperation between companies and publicly funded, part-time vocational schools, which provides apprentices with market-relevant skills and competencies. This close connection between the labor market and school system is largely restricted to the German-speaking realm (Federal Ministry of Education and Research 2020). At present, 327 different apprenticeships are available (Bundesinstitut für Berufsbildung 2018). Within this wide variety of different apprenticeships, commercial trainings (e.g. industrial clerk, bank clerk, office clerk) comprise an important segment of the dual vocational training system, and include accounting-specific education (Brötz et al. 2015).

Recent years have seen a lively and intense debate about reforming accounting curricula (and subsequently learning material) both in Germany and the United States (e.g. Flood 2014; Preiß 2015). The need for change stems from two main trends. First, current accounting education is said to not sufficiently provide learners with skills required in today's 21st-century accounting work environment, such as the qualified use of ERP systems, the handling of big data, or the appropriate use of accounting systems as complex tools for analyzing and planning business situations (rather than just documenting them). Second, regardless of the specific occupation as an accountant, everyone involved in business transactions is required (or at least encouraged) to obtain a broad economic understanding; accounting-related knowledge lies at the heart of this kind of comprehensive economic competence (e.g. Flood 2014; Jackling and Lange 2009; Jordanski 2020; Kavanagh and Drennan 2008; Preiß 2015). The following paragraphs summarize current and recurring demands in accounting education, helping derive requirements for the design of tasks in accounting. However, it's important to keep in mind that tasks are always a model of reality, and never represent reality itself. Creators of tasks play a crucial role here, as they are responsible for creating the “model of reality” and the complexity of the task. The complexity of a task is higher if a task is very similar to the

---

<sup>2</sup> The German vocational training system is done on the ISCED (International Standard Classification of Education) Level 3 (upper secondary education). Programs at this level are “typically designed to complete secondary education in preparation for tertiary education or provide skills relevant to employment, or both” (Unesco Institute for Statistics (2012: 38). Vocational training in Germany usually takes three years, and its learners are normally 16 to 20 years of age.

real-world situation. Put differently, the complexity decreases if the task represents a simplified model of reality. A creator of a task can increase or decrease its complexity in multiple ways, e.g. by structuring relevant information, providing additional but irrelevant information, or by clearly stating what the problem is.

### Process orientation

One demand that has been frequently addressed in accounting is the need for process orientation (Neuweg 2020; Preiß 2015; The Pathway Commission 2012; Walker and Ainsworth 2001). Organizations have always worked process-oriented, i.e. in a way that “they perform a sequence of activities that consume resources for the production of goods and/or provision of services in order to create value for the customer and for the organization [...]” (Trigo et al. 2016: 988). Accounting systems prepare, document, and evaluate these processes and provide data to initiate new processes. Business processes are “a common factor along all organizations” and are therefore vital when it comes to the education of future (accounting) employees (Trigo et al. 2016: 988). The Pathway Commission (2015: 21) defines process management and improvement as broad management competencies, which set the ability to “describe the value chains that create customer and organizational value, including key processes and technologies commonly used within and between organizations to deliver products and/or services” as a main learning objective. In their 2012 published strategy for the next generation of accountants, they point out the relevance of process orientation in accounting education (The Pathway Commission 2012). However, process chains are often neglected, focusing on separate, accounting-relevant business operations (e.g. calculating a selling price, posting depreciations). By concentrating on single operations and functional areas (e.g. accounting, finance, management), learners do not gain a broad understanding of business in an integrated manner. And when this is the case, accounting cannot be used as a vital tool which documents, evaluates, or initiates real business processes, but only as an abstract and formal system whose use and purpose often remain unrecognized (Neuweg 2020; Tramm et al. 1996).

To stimulate the process thinking of learners, German vocational schools have redesigned their occupational curricula, following the principle of process chains<sup>3</sup> (Preiß 2015). One learning field focusing on accounting content (commercial management) for example pursues the following goal:

*Learners record a company's relationship to customers and suppliers by means of information, money and value flows and carry out evaluations. They process receipts and systematically document data resulting from operational processes complying with relevant legal regulations. Using these records, they outline a company's assets and its financial situation, determine its success and discuss main factors influencing a company's business success. (Staatsinstitut für Schulqualität und Bildungsforschung 2004: 18)*

The importance of process and value chains is highlighted by designing tasks in a process-oriented fashion, making clear that all business situations are part of a process

<sup>3</sup> With a process chain orientation, curricula are not structured by content or subject, but instead by learning fields. For example, with “handling special sales situations,” the content to be covered in this learning field is consumer behavior, warranty, purchase on credit, conflict behavior, etc.

chain and are therefore preceded or followed by other situations or decisions (e.g. before buying a new production machine, different offers were obtained, and investment calculation methods were applied). Promoting process thinking means that a task should first realistically describe business situations and processes (e.g. Which stakeholders are involved? Which goods are produced? Which services are provided?). This can be done for example by using realistic documents such as financial statements, receipts, balance sheets, or contracts. In addition, it seems important when teaching to illustrate that every accounting-relevant process is part of a process chain and is therefore preceded and followed by other business processes, enabling learners to gain a “broad and interdisciplinary view of the work environment” (Walker and Ainsworth 2001: 41).

### **Problem orientation**

Problem orientation is similar to the concept of problem-based learning, a method which has been frequently discussed in accounting education and other disciplines. The impact and usefulness of problem-oriented learning has been proved in many empirical studies. For example, a meta-meta-analysis conducted by Hattie (2012: 85), and including 221 studies, revealed that problem-solving teaching is one of the most successful ways to promote students’ learning outcomes ( $d = 0.72$ ). The ability for effective problem-solving can be found in almost all recent recommendations for improving accounting curricula (e.g. Cunningham 2014; Sundem 2014; Watty 2014). Sundem (2014: 623) compiled a list of competencies needed by accounting graduates, which were included in 20 different statements from a variety of educational and professional institutions in the United States. The “ability to identify and solve unstructured problems in unfamiliar settings and to apply problem-solving skills” was one of the capabilities demanded here. The Pathway Commission (2015: 18) defines problem solving as a professional accounting competency, including the ability to “apply a systematic process of using professional judgment to solve a problem”, “interpret information accurately and objectively”, and “describe the process of gathering information about a situation before making a decision”.

The need to foster learners’ abilities to search for additional information, evaluate the relevance of information, and detect and solve realistic problems were frequently addressed in the past (e.g. Dockter 2012; Hansen 2006; Stanley and Marsden 2012). “Students should have the ability to locate, obtain and organize information, and develop the ability to identify and solve unstructured problems in unfamiliar settings; and to exercise judgement based on comprehension of an unfocused set of facts” (Stanley and Marsden 2012: 268).

As a result, tasks should be designed in a problem-oriented manner. This means that a task should include a problem which has to be detected by the learner. This aspect is crucial, as tasks are usually designed in a way that the problem and the activity which should be performed are clearly stated. In order to further increase problem orientation, a task should additionally not include all the information necessary to solve the problem, and simultaneously include an informational overflow that requires learners to search for additional information and assess the relevance of available information.

### Modeling cycle

Another recent demand placed upon accounting education is closely linked to the request to foster process thinking and problem solving. Adapting the modeling cycle from mathematics education, accounting education scholars in the German-speaking realm promote modeling competencies of learners (e.g. Berding et al. 2019; Guggemos 2016; Helm 2016; Neuweg 2020; Preiß 2005; Winther and Achtenhagen 2008). In mathematics, the first step towards successfully solving a real problem is to correctly evaluate a given situation and translate the real situation into a mathematical model. Mathematical operations are then applied, which lead to mathematical results. In another step, these mathematical results need to be retransferred to a real-world context by interpreting the results. Finally, the solution needs to be checked against reality and, if necessary, modified (Blum 2011; Crouch and Haines 2004; Große 2015).

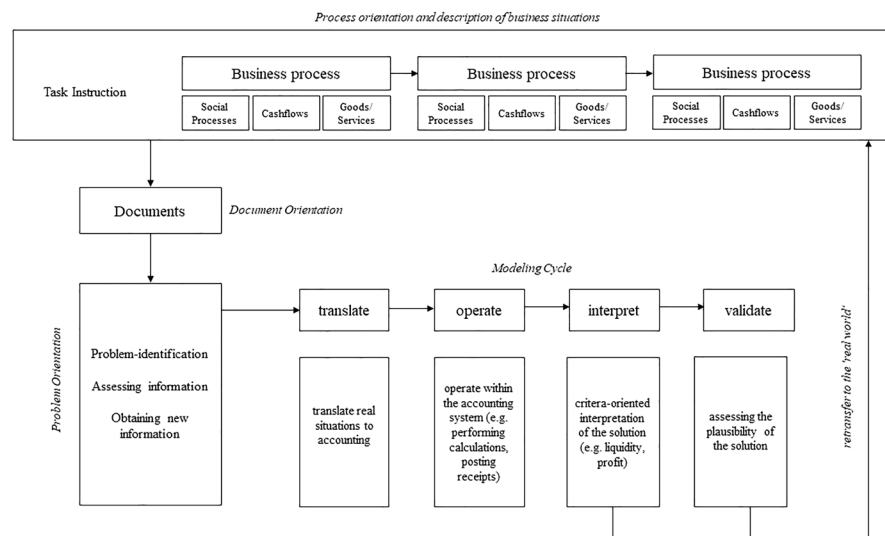
This cyclical process can be transferred to accounting-related problems. Similar to mathematics, accounting is a complex system consisting of abstract rules and procedures, powerful tools that help understand, document, evaluate, and model business processes. This makes it necessary to first transfer business situations to an accounting model. Learners need to identify the problem, structure the situation, assess information and, if necessary, search for additional information. In accounting, documents serve as the “bridge between the real world and the accounting world” (Neuweg 2020: 140) as they represent real processes (e.g. the real process of buying a production machine is represented by an incoming invoice). The correct handling of documents is therefore a vital skill. In a next step, these documents need to be correctly handled (e.g. posting the incoming invoice, calculating a depreciation). The process does not end here. After developing a solution within the accounting system, it needs to be re-transferred to the real world, leading to real consequences (e.g. how does the purchase affect my liquid assets and what can we do about it?). Usually, all processes are evaluated with respect to their impact on (at least) two major business goals: liquidity and profitability (Berding et al. 2019; Neuweg 2020; Preiß 2015).

Regarding modeling competencies, tasks need to allow learners to go through several steps (or ideally all steps) of the modeling process, especially focusing on the transformation from the real world to accounting (translation) and the reverse (interpretation), as “successful mathematical modelling involves an ability to move between the real world and the mathematical world, bearing both in mind” (Crouch and Haines 2004: 199). The same is true for modeling in the accounting domain. The first step of the modeling cycle (transforming situations to accounting) is crucial, because this is often challenging for learners and a frequent cause of learning difficulties (e.g. Tramm et al. 1996; Türling 2014). However, these translation efforts are often not necessary, as tasks are already simplified models of reality which do not require any further translation.<sup>4</sup> In addition, tasks should not only require learners to operate (e.g. calculating, documenting), but should also encourage them to continue to work with the solution by interpreting it in terms of different aspects (e.g. liquidity and profitability aspects, which are often neglected in accounting education, Albrecht and Sack 2000; Preiß 2015), while checking their solution against reality (validate).

---

<sup>4</sup> Strictly speaking, translation efforts are necessary. They are executed by the task designer, and not by the learner.





**Fig. 1** Summarized characteristics for tasks in accounting

Summing up, we believe that accounting tasks should include the following characteristics. Figure 1 provides an additional illustration.

- *Process orientation:* A task should provide realistic, detailed information on real-world business situations and processes (Which stakeholders are involved? Which goods and services are provided?). We call this category “density of real-world information” and include three subcategories for the description of social processes, cash flows, as well as goods and services. We also developed a criterion for the use of documents, referring to it as “document orientation.” A task should furthermore refer to preceding or following processes to allow learners to gain a broad view of occurring business processes. We call this category “**process orientation**.”
- *Problem orientation:* A task should include a realistic and relevant problem which has to be detected and solved by the learner, require learners to search for additional information, and assess the relevance of the information. We name this category “**problem orientation**” and include three subcategories of problem identification, obtaining missing information, and assessing the relevance of information.
- *Modeling cycle:* A task should require learners to translate real-world situations to the world of accounting, operate within the accounting system, and interpret and validate the solution. This criterion is called the “**modeling cycle**” and contains the subcategories of translation, operation, interpretation, and validation.

Regarding existing research, we are not aware of any study (german-speaking or Anglo-American area) which specifically focuses on one of the addressed aspects. However, existing research findings on the quality of accounting tasks regarding other aspects, e.g. the cognitive demand of a task, indicate that accounting tasks tend to be rather low in quality. Davidson and Baldwin (2005) for example analyzed end-of-chapter tasks contained in textbooks for Intermediate Financial Accounting classes and found that only 9.1% of the tasks are located at level five (Evaluation) or level six (Synthesis) of

the taxonomy for learning objectives. Arek-Bawa and Dhunpath (2017) come to similar results, with only 7% of the tasks focusing on the two top levels. Findings from the German-speaking area point in the same direction. Most of the tasks aim at the reproduction of knowledge and are located at levels one to four on the taxonomy. In addition, tasks tend to insufficiently refer to real-world problems and rarely simulate authentic occupational situations (Bloemen 2011; Ernst 2012; Thoma and Schumacher 2018; Wuttke et al. 2022).

### Research questions

As we are not aware of any study examining the characteristics of accounting tasks regarding the categories addressed in Chapter 2, our objective is to provide mainly descriptive information; our study aims to address the following research questions:

Q1: Are accounting tasks in German textbooks process-oriented, and do they provide realistic and detailed information about real business situations?

Q2: Are accounting tasks in German textbooks problem-oriented?

Q3: Do tasks contained in accounting textbooks require learners to go through the steps of the modeling cycle?

Q4: Do accounting textbooks contain typical types of tasks, e.g. do certain characteristics occur in combination with other aspects and how are different task characteristics related?

What is more, in prior research studies, only a small number of tasks, e.g. from specific book chapters or particularly difficult topics, was evaluated. Another objective of this study is therefore to analyze a large number of tasks; this is done by applying AI to content analysis. Assessing the quality of a task based on different aspects is crucial and complex. Therefore, the fifth research question addresses the usability of AI for content analysis.

Q5: How good is AI at assessing the quality of learning tasks compared to human coders?

### Method

#### Sample

To answer the research questions, this study analyzed 14 vocational education and training textbooks in Germany. Except for one, these were published between 2016 and 2019 (see Table 1). The books are used for different apprenticeships (e.g. industrial clerk, retail salesman) and contain a total of 3,361 tasks. Table 1 provides information on the publication year, the number of tasks available, and the industry training the book is used for. The study includes all different types of tasks contained in the textbook (exercises, questions, case studies, etc.).

#### Measurement Instrument

Qualitative content analysis described by Mayring (2015) was applied for analyzing task characteristics from an accounting-specific perspective. A new category system was developed for this purpose based on the theoretical background outlined in section two. First, we included a category for evaluating the process orientation of tasks. The key



**Table 1** Bibliographic Information of the Analyzed Textbooks

	Year	Number of Tasks	Branch
1	2018	777	Retail Sector
2	2016	94	Media
3	2018	116	Taxation
4	2015	122	Office Management
5	2019	654	Industry
6	2017	190	Industry
7	2017	25	Industry
8	2018	157	IT
9	2018	671	Industry
10	2019	128	E-Commerce
11	2016	40	Industry
12	2017	28	Industry
13	2016	233	Industry
14	2017	126	Industry

Textbooks 1 and 5 are well-established in vocational education and training, and are published in their 15<sup>th</sup> and 48<sup>th</sup> editions, respectively

questions here pertained to whether a task is part of a process chain, and whether it is clearly stated that a situation is preceded or followed by other processes. What is more, we included categories for analyzing whether a task realistically describes business situations. The category system also includes whether a task uses documents to represent reality, as documents serve as a representation of it. In addition, we developed categories for evaluating a task's potential to foster modeling competencies (translate, operate, interpret, and validate), especially focusing on the first aspect by implementing three criteria for problem orientation. Concerning the interpretation of solutions, we particularly focused on two aspects: liquidity and profitability, as these are two main goals of every company and therefore relevant criteria for decision making. Developing a category system is always a trade-off between complexity and usability. As category levels, we used either two (0 = not existing, 1 = existing) or three levels (e.g. 0 = no process orientation, 1 = limited process orientation, 2 = developed process orientation). The codebook (see Table 2) comprised a definition of each category, a description of the different category levels, example tasks, and coding rules.

The category system is designed for human coders, and in the current study (due to the large number of analyzed tasks) was applied via artificial intelligence (AI). AI describes the attempt to simulate human actions by a computer (Kleesiek et al. 2020). The application of AI for content analyses provides several promising advantages. First, AI can analyze more data than human coders, leading to a more solid and broader data basis for investigation at lower costs. Second, the saved resources can be allocated to the difficult aspects of content analysis, providing the opportunity for more valid results (Scharkow 2011). Third, trained AI can be used for further research, reducing costs and offering the application of exactly the same coding scheme. This can increase the comparability of different studies. Fourth, trained AI can be embedded into learning analytics software for accounting educators, supporting them in creating high quality learning tasks for

**Table 2** Codebook for content analysis

Category	Key question/definition	Description of levels	Example tasks for different levels	Explanation
Density of Real-World Information and Process Orientation	Is the task part of a process chain? Is it clear that the situation stated in the task is preceded or followed by another situation? Does the task refer to preceding or forthcoming situations?	0 = no reference to preceding or forthcoming situations 1 = limited reference to preceding or forthcoming situations 2 = reference to several preceding or forthcoming situations or to preceding and forthcoming situations	<i>Level 1:</i> On Dec. 31 <sup>st</sup> the invoice from the public utility company for electricity for December is still outstanding. Based on the meter reading, a consumption of 680 euros is estimated a) What is the correct posting on Dec. 31 <sup>st</sup> ? b) How should the invoice amount of 1,309 euros including 19% VAT be booked on January 20 <sup>th</sup> of the following year?	The task involves two steps of the process chain (posting on December 31 <sup>st</sup> , and posting the incoming invoice on January 20 <sup>th</sup> )
Description of social processes	Does the task represent social processes occurring in a particular situation realistically and in a detailed manner?	0 = no representation 1 = abstract representation 2 = detailed and realistic representation (e.g. by using pictures, accurate descriptions, etc.)	<i>Level 1:</i> The managing directors of Sommerfeld Ltd. plan to build a new production hall (estimated costs: 400,000 euros). The managing director, Mrs. Farthmann, proposes to finance the hall exclusively with equity capital. Mr. Feld is of the opinion that it would be more advantageous to take out a loan from the bank. a) Present the advantages of financing with equity capital in an overview for Mrs. Farthmann. b) Collect arguments in favor of Mr. Feld's proposal	The task description outlines a social process/situation (two opposing opinions of managing directors), but not in great detail

**Table 2** (continued)

Category	Key question/definition	Description of levels	Example tasks for different levels	Explanation
Description of cash flows	Does the task represent the real-world context realistically and in a detailed manner when it comes to liquid funds?	0 = no representation 1 = abstract representation 2 = detailed and realistic representation (e.g. by using pictures, accurate descriptions, etc.)	<i>Level 1:</i> The vehicle tax for five company cars (2,610 euros) is paid to the tax authority by bank transfer on September 1 <sup>st</sup> for the period from September 1 <sup>st</sup> to August 31 <sup>st</sup>	The task contains a cash flow, but does not provide a realistic and detailed representation (e.g. by using an account statement)
		0 = no representation 1 = abstract representation 2 = detailed and realistic representation (e.g. by using pictures, accurate descriptions, etc.)	<i>Level 1:</i> Purchase of a photocopier for 1,000 euros + 19% VAT. Correctly document the transaction	No detailed description of the purchased product available (e.g. by using a picture)
		0 = no document included in the task 1 = document included, but irrelevant for solving the task 2 = document included and relevant for solving the task	<i>Level 2:</i> Correctly post the following receipt. (Note: incoming invoice is displayed)	The document included is necessary to come to a solution as no information (e.g. concerning the supplier or prices) is available except for the receipt
Problem Orientation	Does the task contain realistic documents which need to be used for solving the task?	0 = task with no problem (e.g. basic knowledge questions) 1 = problem is given 2 = problem needs to be identified by the learner	<i>Level 2:</i> On December 31 <sup>st</sup> , a liability of 12,000 euros was recognized for the legal costs of a pending lawsuit. Before the balance sheet is drawn up on March 10 <sup>th</sup> , the lawsuit unexpectedly ends in our favor. Evaluate the situation and take the necessary steps	The task includes a problem which is not clearly stated for the learner. The learner has to identify the problem and take further steps to solve it
		Is the learner required to identify the problem, or is it already given within the task instruction?		

**Table 2** (continued)

Category	Key question/definition	Description of levels	Example tasks for different levels	Explanation
Obtaining missing information	Is the learner encouraged to research additional, relevant information in order to solve the task, or is all relevant information already included?	0 = no research necessary 1 = research within the task necessary (e.g. from balance sheets, income statements, contracts) 2 = research from other resources (e.g. internet, textbook or other learning material) or from more than one source within the task is necessary	<i>Level 2:</i> a) Calculate the profitability for Fasty Fruit. b) Check whether this ratio has changed by determining the previous year's value and compare it to the current value. c) Compare the ratio from a) with the values of other companies in the same industry. To do this, carry out an internet search	The task requires the learner to search for additional information from other resources (internet) and within given material within the task
Assessing the relevance of information	Does the task contain additional information so that learners must distinguish relevant from irrelevant information?	0 = all information is needed to complete the task, no selection necessary 1 = task contains irrelevant information; learner must differentiate between relevant and irrelevant information	<i>Level 1:</i> The textiles privately withdrawn by family Schiffberger in the last quarter have not yet been recognized. Cost Price: 1.478 euros, Replacement Price: 1.350 euros Sale Price incl. 20% VAT: 2,778 euros	The learner needs to assess which price is relevant for correctly solving the task
Modeling Cycle	Does the task require translation of an economic phenomenon into accounting? (e.g. by correctly using specific terms, simplifying and structuring information, etc.)	0 = no translation into an accounting model necessary 1 = translation necessary	<i>Level 0:</i> The food retailer Krug purchases organic wine from Italy. He is charged 144 euros freight for 480 bottles. How much does the freight cost him for 720 bottles? <i>Level 1:</i> Assist Rudolf in preparing the protocol by assigning the investment requests mentioned at the department manager meeting to one of the investment types <i>Note: detailed conversation process is displayed</i>	Task 1: No translation necessary, as this is not an accounting problem, but only a mathematical problem Task 2: A situation (department manager meeting) needs to be transferred to accounting (identifying investment types) by structuring the information, using specific terms, etc

**Table 2** (continued)

Category	Key question/definition	Description of levels	Example tasks for different levels	Explanation
Operate	Does the task require performing calculations or applying formal accounting rules?	0 = no calculations or posting documents necessary 1 = calculations or posting documents necessary	<i>Level 1:</i> For a current lawsuit, court costs are estimated at 5,800 euros. Correctly post the situation on the balance sheet day (Dec. 31 <sup>st</sup> )	Learners are required to apply formal accounting rules, as they need to correctly document the situation (accrued liabilities for litigation)
Interpret	Does the task require interpreting the solution?	0 = no interpretation of the solution necessary 1 = interpretation of the solution necessary	<i>Level 1:</i> Calculate for Juices & More for the month of July a) the gross profit b) the gross margin c) the return on sales. Assess whether the company achieves satisfactory results	Learners are required to interpret the solution, as they need to assess whether gross margin, gross profit, and return on sales are satisfactory
Interpret: liquidity aspects	Is the learner required to interpret the solution of the task or a transaction/situation in terms of its impact on liquidity?	0 = no interpretation regarding the impact on liquidity necessary 1 = interpretation regarding the impact on liquidity necessary	<i>Level 1:</i> Assist Tina Fröhlich with the following questions regarding the evaluation of the results for the month of September: a) Calculate gross profit, gross margin, and return on sales for the month of September. b) Evaluate their respective trends over the three months of business. c) Review the development of liquidity. Compare the development of cash and short-term receivables on the one hand and short-term liabilities on the other. d) Consider the consequences of taking out the loan for liquidity in the coming months	The task requires learners to interpret a situation (e.g. taking out a loan) regarding its impact on liquidity

**Table 2** (continued)

Category	Key question/definition	Description of levels	Example tasks for different levels	Explanation
Interpret profitability aspects	Is the learner required to interpret the solution of the task or a transaction/situation in terms of its impact on profitability/profits?	0 = no interpretation regarding the impact on profits necessary 1 = no interpretation regarding the impact on profits necessary	<p><i>Level 1:</i> Check whether the following business transactions increase, decrease, or have no effect on the company's profit</p> <p>c) Receipt of December 23<sup>rd</sup>: payment of social insurance contributions for the month of December</p> <p><i>Note: only one transaction stated as an example</i></p>	The task requires the learner to interpret a transaction regarding its impact on profitability aspects
Validate	Does the task require the solution to be evaluated?	0 = no validation of the solution necessary 1 = validation necessary	<p><i>Level 1:</i> A manufacturer of household appliances produced three types of juicers in the past accounting period. Determine from the information listed below using break-even analysis a) the operating profit and the contribution margin per unit b) Critically comment on your calculation results</p>	Learners have to critically evaluate their developed solution



their students. This study attempted to analyze the tasks based on AI-generated information. The next sections briefly describe the development of the AI.

### Generating data for training the AI

An important element of AI is machine learning, which occurs when a computer develops the solution to a problem by generating a computer program itself (Alpaydin 2019; Lanquillon 2019). In the special case of supervised machine learning, AI tries to generate a prediction model, which transforms input data into output data, e.g. a target variable. Here, AI needs a data collection of input *and* output data for identifying the relationship between the two data types (Lanquillon 2019). Thus, in a first step it is necessary to generate the input and the output data for training the AI. When this is done, human coders have to rate a number of tasks by applying the category system to ensure the validity of the ratings.

Determining the quality of ratings within content analysis is an important but difficult task. As Zhao et al. (2013) state, most of the available indices for measuring inter-coder reliability are based on problematic assumptions, leading to paradox results and abnormalities. Hove et al. (2018) show that different indices reach very different values for the same data. Against this background, Krippendorff's  $\alpha$  (Krippendorff 2004) is the most recommended coefficient (Hayes and Krippendorff 2007), even though the validity of  $\alpha$  is the subject of controversial scientific discussion (e.g., Feng and Zhao 2016; Krippendorff 2016; Zhao et al. 2018). The current study follows the *best available for a situation* approach, proposed by Zhao et al. (2018), and reports the following measures:

The *percentage of agreement* measures the number of observation units that all coders rated the same (tolerance is zero), or that all except one coder rated the same (tolerance is one). The percentage of agreement (zero tolerance) does not correct for guessing, but produces the fewest paradoxes and abnormalities (Zhao et al. 2013).

*Krippendorff's  $\alpha$*  corrects for guessing and is applicable for scales of any level and any number of coders (Hayes and Krippendorff 2007). It ranges between -1 and +1. Values close to 1 indicate a perfect reliability. According to Krippendorff (2004: 240), a value of at least 0.67 indicates an acceptable coding. However, it suffers from the greatest number of paradoxes and abnormalities. For example,  $\alpha$  decreases with larger sample sizes, and can produce low values although the coders highly agree, while pure guessing can produce some values above zero (Zhao et al. 2013).

Whereas the percentage of agreement is a more liberal index, Krippendorff's  $\alpha$  is a very conservative one (Zhao et al. 2013). The percentage of agreement produces unfairly high reliability values in situations with a low agreement, whereas Krippendorff's  $\alpha$  produces unfairly low scores for uneven distributions (Feng 2013; Zhao et al. 2013). Since most empirical studies of textbooks in vocational education and trainings show extremely uneven frequency distributions (e.g. Arek-Bawa and Dhunpath 2017; Berding and Lau 2018; Bloemen 2011; Davidson 2005),  $\alpha$  could systematically generate too-low reliability scores. This problem increases by applying AI because the number of analyzed tasks rises significantly, meaning the "truth" may be located between percentage agreement and Krippendorff's  $\alpha$ , leading to reporting both measures in the current study.

Table 3 reports the quality of the generated classifications  $\alpha$ . The percentage agreement for human ratings were computed with the package *irr* for *R* (Gamer et al. 2019).

**Table 3** Inter-coder reliability measured by Krippendorff's alpha

Construct	Human rating						Training of AI			
	K = 5, N = 56			K = 3, N = 56			I = 100			
	Initial analysis			Validation analysis			Neural Net		Random Forest	
	% <sup>a</sup>	% <sup>b</sup>	$\alpha$	% <sup>a</sup>	% <sup>b</sup>	$\alpha$	% <sup>b</sup>	$\alpha$	% <sup>b</sup>	$\alpha$
Process orientation	91	52	0.095	98	68	0.457	85	0.397	87	0.110
Description of social processes	82	75	0.089	98	71	0.285	93	0.460	94	0.383
Description of cash flows	94	48	0.409	95	84	0.579	90	0.394	92	0.310
Description of goods and services	94	58	0.529	96	60	0.441	92	0.410	94	0.350
Document orientation	89	76	0.483	95	95	0.709	96	0.383	96	0.474
Identification of problems	94	46	0.393	93	54	0.445	71	0.439	78	0.554
Obtaining missing information	74	43	0.287	91	75	0.732	72	0.432	76	0.303
Assessing the relevance of information	100	83	0.057	100	65	0.166	99	0.153	99	-0.006
Translate	100	21	0.064	100	38	0.157	67	0.220	75	0.245
Operate	100	60	0.618	100	91	0.872	83	0.646	88	0.742
Interpret	100	65	0.543	100	61	0.382	78	0.339	83	0.280
Interpret: liquidity aspects	100	87	0.325	100	96	-0.009	97	0.236	97	0.142
Interpret: profitability aspects	100	89	0.504	100	93	0.407	92	0.507	93	0.341
Validate	100	85	0.070	100	88	0.258	100	0.980	100	1.000

K Number of raters, N number of analyzed tasks, I iterations of performance assessment

<sup>a</sup> Estimation with Tolerance 1

<sup>b</sup> Estimation with Tolerance 0

After developing a first draft of the category system, five people randomly rated 56 tasks. As seen in Table 3, the percentage of agreement (tolerance one) is very high, but percentage of agreement (tolerance zero) is low. Furthermore, the cutoff level for  $\alpha$  could not be reached for all variables during the first application. The coding team discussed deviations and revised the category system as a result.

In a next step, three persons of the coding team rated 56 new and randomly selected tasks. The alpha values increased here, and three scales yielded good results (obtaining missing information, operate, document orientation). The percentage agreement (tolerance 0) also increased in most cases. As the alpha values of the other scales remained below the cutoff value, the coding team discussed the deviations again and revised the category system a second time. After the second revision, two people from the author team rated 1,080 tasks. If the raters felt uncertain about a rating, the author group scored the task together. Because the assessment of the didactical quality of a task is very difficult and complex, this process is understood as ensuring the best validity possible. In every case, human raters used the concrete texts of a task as well as the corresponding graphics, illustrations, and tables. In contrast, AI could only use the textual data and was unable to use graphics and pictures.

### Preparation of the data for AI application

The information on the 1,080 tasks provided the basis for developing the AI. Although several studies comparing the performance of different AI algorithms exist (e.g., Lorena et al. 2011; Scharkow 2011), only a few studies compare the performance of classification tasks based on textual information. For example, the study conducted by Hartmann et al. (2019) using text messages, comments on Facebook, and Amazon reveals that random forest and naive Bayes perform best. The study by Berding et al. (2021) with textual data produced by students and apprentices reveals a superior performance of bagging, Glmnet, and decision trees. Thus, it is currently not clear which method works best for analyzing textbooks. Against this background, the current study applies random forest because tree-based algorithms are promising for classification tasks. Additionally, the study considers neural nets as a different class of AI.

Inspired by the structure and functionality of the human brain, training a neural net means establishing links between one or several “neurons” (Ertel 2016). The training process optimizes the links in a way that the neural net transforms the input data into correct output data (Lanquillon 2019). Another algorithm found within AI is decision trees. A special form of decision trees is random forest, where the algorithm generates a huge number of decision trees, and the category level that the most trees imply is allocated to the analyzed unit (Lorena et al. 2011). The concrete computations rely on the work by Lang et al. (2019) with the R package *mlr3*, Wright and Ziegler (2017) (R package *ranger*), and Venables and Ripley (2007) (R package *nnet*). While *mlr3* represents the interface for using different kinds of machine learning in R, both *ranger* and *nnet* support multiclass data. In terms of machine learning, multiclass data means that the output data has more than two categories for each construct.

For both random forest and neural net, the textual data was prepared for training the AI in several steps. First, the document-term matrix (DTM) was generated. The DTM reports the cases in the rows and the frequencies of the words in the columns. Second, the words were reduced to verbs, adjectives, adverbs, and nouns because these language elements carry most of the information (Papilloud and Hinneburg 2018). In a next step, all of the words were lemmatized. Lemmatization transforms the words to their root form (e.g. from “he goes” to “go”). Finally, stopwords and all words with a very small frequency were deleted. All steps are very important because they reduce the dimensionality of the input data and concentrate on the most important words in terms of content. The preparation was done with *quanteda* (Benoit et al. 2020) and *UDPipe* (Wijffels et al. 2019).

While the data preparation for random forest was finished with this step, neural net required an additional transformation. For neural nets, it’s important that the input data and the output data use only values between  $-1$  and  $+1$ . Thus, all data was transformed by applying the following function for training and predicting the data:

$$f(x_i) = 0.95 * \frac{x_i - \min(x)}{\max(x) - \min(x)} + 0.01$$

The function  $f$  ensures that the data is in the range between  $-1$  and  $+1$ , and that the extreme values are never reached in order to avoid computational problems. The predicted data was re-transformed by applying the inverse function of  $f$ .

The synthetic minority oversampling technique (SMOTE) was applied to ensure a correct classification of categories with low frequencies. These kind of techniques are important because most machine learning algorithms do not perform well with imbalanced data (Haixiang et al. 2017).

### Estimating the AI performance

The complete sample was randomly split into training data (75% of the 1,080 tasks) and test data (25% of the 1,080 tasks) to assess the performance of both AIs. According to the simulation study of Berding et al. (2021), a sample size of about 300 training units can be sufficient for training and estimating the performance of AI for text analyses if the AI is applied for predicting pedagogical and didactical variables on similar input data.

In a first step, the AI learned to rate tasks by analyzing the training data. Following this, the AI rated the remaining 25% of the tasks. In a third step, the ratings of the AI for the test data were compared to the expert ratings by computing Krippendorff's  $\alpha$  and percentage of agreement. This procedure was repeated 100 times to adjust the performance estimation from sample effects. These samples represent bootstrap samples. The mean of the 100 alpha values was used as a performance estimate. Table 3 reports the results of this bootstrap approach. Both algorithms achieved a high percentage of agreement with human coding. However, AI reached the cut-off value of 0.67 proposed by Krippendorff in only two of the categories (2004). To achieve the best results possible for each category, the best AI was selected based on the alpha values. Table 4 shows which construct was predicted by both AI algorithms.

The quality of the training data is critical for generating valid results as the simulation study by Song et al. (2020) showed. In this study, the level of intercoder-reliability measured by Krippendorff's Alpha explains about 62% of the mean absolute prediction error. While the mean absolute prediction error for validation data with Krippendorff's Alpha about 0.50 is 0.550, the error is about 0.0357 for an alpha of 0.70. and about 0.0262 for an alpha of 0.90. The results also indicate that AI tends to underestimate the characteristics of a category.

### Training the AI and generating data for the study

After estimating the performance of the AI, the complete sample of 1,080 tasks was used for training the algorithm to take all available information into account, and the data was prepared as described in Sect. 4.4. The trained learner predicted all available 3,361 tasks, which form the basis for the following analyses. The implementation of the AI was based on the work by Benoit et al. (2020), Lang et al. (2019), and Wijffels et al. (2019) for R.

## Results

### Descriptive data of task characteristics in accounting

To answer the research questions "Are accounting tasks in German textbooks process-oriented, and do they provide realistic and detailed information on real business

**Table 4** Relative and Absolute Frequencies of the Category Levels Generated by AI

Construct	Category level			Applied AI method
	0	1	2	
Process orientation	1914 (57%)	915 (27%)	532 (16%)	Neural Net
Description of social processes	2812 (84%)	491 (15%)	58 (2%)	Neural Net
Description of goods and services	2784 (83%)	577 (17%)	0 (0%)	Neural Net
Description of cash flows	2713 (81%)	648 (19%)	0 (0%)	Neural Net
Document orientation	3271 (97%)	0 (0%)	90 (3%)	Random Forest
Identification of problems	2292 (68%)	1059 (32%)	10 (0%)	Random Forest
Obtaining missing information	1902 (57%)	963 (29%)	496 (15%)	Neural Net
Assessing the relevance of information	3361 (100%)	0 (0%)		Neural Net
Translate	2866 (85%)	495 (15%)		Random Forest
Operate	1260 (38%)	2101 (62%)		Random Forest
Interpret	2431 (72%)	930 (27%)		Neural Net
Interpret: liquidity aspects	3187 (95%)	174 (5%)		Neural Net
Interpret: profitability aspects	2618 (78%)	743 (22%)		Neural Net
Validate	3359 (100%)	2 (0%)		Random Forest

situations?,” “Are accounting tasks in German textbooks problem-oriented?” and “Do tasks contained in accounting textbooks require learners to go through the steps of the modeling cycle?”, Table 4 reports the relative and absolute frequencies generated by the AI for all available tasks. The table shows that most tasks only reach the lowest level. In fact, none or only a few tasks require the learners to assess the relevance of information, or analyze the validity of the generated solution for explaining and shaping business situations (validate).

Regarding process orientation, about 16% of the tasks are embedded in business processes and refer to *several* forthcoming or previous steps of the process chain, leading to a high process orientation. Another 27% of the tasks focus on the previous *or* the forthcoming process step. The majority of the tasks (57%), however, do not include previous or forthcoming steps of the process chain, and are therefore low in process orientation. Table 2 entails a sample task for level 1, as it includes two steps of a process. Concerning the description of goods and services, 17% of the tasks provide an abstract description of the value creation process (e.g. which goods and services are produced/provided). No task contained a detailed and realistic description, e.g. describing the situation in detail. The results are similar for real-world information on cash flows. When it comes to social processes, the tasks yield slightly better results regarding level 2. About 58 tasks (2%) refer to a realistic and detailed description of

the underlying social process. Another 15% provide at least an abstract representation. In line with this result, most tasks do not require using any documents to develop a solution (97% of the tasks). Table 2 again contains example tasks for level 1 regarding the description of social processes, cash flows and goods and services. For example, the task “Purchase of a photocopier for 1000 euros + 19% VAT. Correctly document the transaction.” describes a product, which was purchased. However, it is not described in great detail or by using a picture to display the photocopier authentically.

Regarding problem orientation, the vast majority of the tasks are tasks which do not contain a problem (68%). If the task does in fact contain a problem, it is clearly stated what the problem is and how it should be solved (32%). Thus, with the exception of ten tasks, learners do not need to identify the problem themselves. Table 2 outlines an example task for level 2. The task describes a business situation, but does not include the exact problem and how the learner should solve the problem. Instead, the task instructs the learner to evaluate the situation and take all necessary steps.

When it comes to searching and assessing information, most tasks provide all the information necessary for generating a solution (57%). Only 29% of the tasks require learners to search for information *within* the provided learning materials. 15% of the tasks demand learners to use additional information sources for generating a solution (e.g. the internet or the textbook, for an example task see Table 2). No tasks contain irrelevant information.<sup>5</sup> As a result, learners are not required to differentiate relevant from irrelevant information, or select the data necessary for problem solving. A fictional task for level 1 is provided in Table 2. Besides written descriptions, the information is represented within tables ( $n = 508$ ), T-accounts ( $n = 279$ ), receipts ( $n = 149$ ), other documents ( $n = 114$ ), and graphics ( $n = 38$ ).

When it comes to modeling business situations, Table 4 shows that tasks tend to focus on certain steps of the modeling cycle, while neglecting other steps. If a task requires modeling parts of the process, it mostly concentrates on applying formal accounting rules (62%). Table 2 shows a task which requires applying accounting rules. Transforming a real-world problem to accounting (translate) and re-transferring a solution from accounting to reality (interpret) is required less frequently, and is necessary in only 15% to 27% of the analyzed tasks. For example, the task presented in Table 2 requires learners to check whether the stated business transactions increase, decrease or have no effect on the company's profit. Therefore, the task was rated at level 1 regarding the interpretation of profitability aspects. Only two tasks require learners to critically assess their solution (validate). Regarding analyzing the effects of business situations on corporate goals, most tasks deal with profitability, and only 174 tasks (5%) focus on liquidity aspects.

In order to further illustrate the descriptive results, a sample task is outlined below in addition to the examples provided in the codebook. Regarding the process orientation, the task refers to a previous process (the purchase of the truck) and is therefore rated at

<sup>5</sup> A task could only be rated at level 1 (informational overflow), if irrelevant information was added deliberately by the task instructor. Therefore, tasks including documents (e.g. an invoice) were rated at level 0, as they necessarily contain irrelevant information (e. g. an invoice includes information such as the address or the bank account number of the supplier).



level 1. By referring to the purchase of the truck, the task provides an abstract description of a good (rated at level 1), but does not include cash flows or social processes (rated at level 0). Regarding problem orientation, the task contains a problem, but it is clearly stated for the learners what activity they have to perform in order to solve the problem. Therefore, the task is rated at level 1. However, the task does not require to search for additional information or differentiate relevant from irrelevant information (rated at level 0). When it comes to the modeling cycle, the task requires learners to apply formal accounting rules (therefore to operate within the accounting system). However, an interpretation and validation of the solution is not necessary.

*On April 12, 20.. a truck was purchased by the beverage delivery company Lisa e.K.. The acquisition costs amounted to 72,820.00 euros, the asset depreciation range is nine years. The total mileage of the truck is estimated at 900,000 km. In the year 20... 66,000 km were driven. Calculate and post the depreciation according to the performance*

#### **Types of tasks in accounting and correlations between different task characteristics**

To provide answers on research question “Do accounting textbooks contain typical types of tasks?” a latent class analysis is applied to the data. This approach attempts to identify groups of tasks (“classes”) based on their values on the different categories. Initially, models with one up to five classes were fitted to the data with the help of a robust maximum-likelihood estimator taking the cluster structure of the data into account (the tasks are clustered in textbooks). For all five models the  $p$ -value of the Pearson chi-square test is nearly one, indicating an absolute fit of the models to the data (Geiser 2013: 258–259). However, the Vuong-Lo-Mendell-Rubin Likelihood ratio test is not significant for all estimated models. Indeed, the  $p$ -value ranges from 0.414 to 0.661, indicating that all models with two up to five classes do *not* significantly fit the data better as a single class (Geiser 2013: 266–267). Thus, latent class analysis implies that there are no different types of accounting tasks, which is in line with the high concentration of specific categories shown in Table 4. Therefore, the small variation in the categories cannot be traced back to different groups of tasks.

In order to provide an alternative view on the nature of accounting tasks, a correlation table is estimated. This approach does not try to cluster tasks but to identify the relationships between the different categories. With the help of *Mplus*, the standard errors were adjusted to the clustered structure of the data (Muthén and Muthén 2017). In order to achieve the most valid estimates of the correlations, all variables are modeled as categorical. Furthermore, the WLSMV estimator is used. This estimator allows the estimation of valid standard errors without assuming a specific distribution (Bentler and Dudgeon 1996; Finney and Di Stefano 2013) and is recommended for categorical variables (Finney and Di Stefano 2013). Table 5 presents the results. Referring to Cohen’s (1988) rules for product-moment correlation as a rule of thumb for this analysis, a value of at least 0.10 indicates a small, of at least 0.30 a medium and of at least 0.50 a strong relationship.

Regarding the description of the real world and the orientation on business processes, all categories (except document orientation) are positively related. This relationship is in conformance with the expectations, as the description of business processes usually

**Table 5** Intercorrelation between different task characteristics

	Density of real-world information and process orientation					Problem orientation		Modeling cycle			
	Process orientation	Description of social processes	Description of cash flows	Description of goods and services	Document orientation	Identification of problems	Obtaining missing information	Translate	Operate	Interpret	Interpret: liquidity aspects
Density of real-world information	.220**										
Description of social processes											
Description of cash flows	.441**	.185**									
Description of goods and services	.277**	.245**	.632**								
Document orientation	-.093	-.052	-.148	.305							
Identification of problems	.293**	.378**	.107	.028	.147						
Obtaining missing information	.003	-.084*	-.008	-.005	.298**	-.043					
Translate	.327**	.038	.340**	.244**	-.027	.016	.003				
Operate	.556**	.247**	.407**	.281**	.339*	.580**	-.023	.152			
Interpret	.045	.022	-.070	-.009	.234	.138**	.072*	-.095	.211**		
Interpret: liquidity aspects	.037	.241**	-.094	-.280**	-.156	.180*	.090	.040	.101	.080	
Interpret: profitability aspects	.060	.037	-.139**	-.364**	-.165	.217**	-.03	-.097	.400**	.168**	.183**

Two categories ("assessing the relevance of information" and "validate") could not be included in the analysis due to a lack of variance in the data

refers to past or future cash flows, social processes or good flows. In particular, the description of cash flows as well as the description of goods and services are strongly correlated, indicating that they are often addressed together within a task. Unexpectedly, document orientation is not significantly related to the other categories, implying that documents do not lead to higher levels of process orientation or more complex real-world descriptions. However, the use of documents is positively related to the category “obtaining missing information”. This relationship is plausible, as relevant information is normally obtained from documents included in a task (e.g. from an invoice or from a balance sheet). In addition, the use of documents and operating within the accounting system are significantly correlated, implying that if a task includes a document (e.g. an incoming invoice), these documents need to be processed in the accounting system.

Regarding problem orientation, the category “assessing the relevance of information” could not be included in the analysis due to a lack of variance in the data. Interestingly, the categories “identification of problems” and “obtaining missing information” are not significantly correlated. Therefore, the requirement to search for additional information does not lead to a higher level of problem orientation. Vice versa, a task can be high in problem orientation without the demand to explicitly research information.

Referring to the different steps of the modeling cycle, only some categories show a significant correlation. The category “validate” could not be included in the analysis due to a lack of variance in the data. The first step of the modeling cycle (translate) is not significantly related to any other subsequent step, whereas the second step of the modeling cycle (operate) is related to the third step (interpret). Hence, tasks demanding students to operate within the accounting system often also demand an interpretation of the developed solution. This seems both plausible and desirable, as abstract results need to be given meaning by interpreting the results in terms of specific aspects (e. g. the impact of a process on the company’s profit). Taking a closer look, the correlation table shows that interpretation efforts mostly concern profitability aspects and less likely liquidity aspects. Thus, if tasks demand an interpretation, they very often demand the interpretation in terms of profit, but not necessarily in terms of liquidity. The lack of correlations between the other steps of the modeling cycle further indicate that the tasks do not demand learners to pass a complete modeling cycle.

Table 5 further outlines some interesting connections between the three main categories (density of real-world information, problem orientation and the modeling cycle). First, results indicate that process orientation as well as the description of the real world are related to problem orientation. Tasks with higher levels of process orientation and more detailed descriptions of social processes show higher levels of problem orientation. However, this is not true for the description of cash flows and goods and services.

Second, regarding the interaction between the real-world description and the modeling cycle, results show that tasks containing a more detailed description of the situation and a high level of process orientation tend to demand learners to translate the described situation into a formal accounting model and to operate within the accounting system. However, high levels of process orientation and detailed descriptions of a situation are not correlated with interpretation efforts. This implies that tasks, although they address and describe real business situations, do not necessarily require to re-transfer abstract results to the real world. Interestingly, more detailed

descriptions of social processes are positively related with interpretations regarding liquidity aspects, whereas the description of goods and services is negatively related to interpretations regarding profitability and liquidity aspects. In addition, the description of cash flows is also negatively related to the interpretation in terms of profitability. Whereas some of these correlations seem coherent (e. g. the negative relationship between cash flows and profitability), others are highly counter-intuitive. For example, the negative correlation between the description of goods and services and the interpretation regarding profitability and liquidity is not comprehensible, as the purchase and sale of goods and services directly affects a company's profit and liquid funds.

Third, tasks with higher levels of problem identification tend to demand learners to operate within the accounting system and to interpret the results in terms of liquidity and profit. However, they do not necessarily require learners to transfer the problem into the abstract system of accounting. The category "obtaining additional information" is not significantly correlated to the different steps of the modeling cycle, except to the category "interpret".

### Comparison of AI and human ratings

After presenting the results generated by AI, Table 6 gives information on the reasonableness of the current findings to answer the research question "How good is AI at assessing the quality of learning tasks compared to human coders?" The Wilcoxon tests comparing the ratings of experts and AI coders reveal significant results for all scales except "validate," making a closer inspection of the deviations between human and AI coders necessary. For a more detailed view, Table 6 presents the effect size  $r$  which Cohen calls  $w$ . According to his study, an  $r$  between 0.10 and 0.30 indicates a small, between 0.30 and 0.50 indicates a medium, and an  $r$  of at least 0.50 indicates a strong deviation (Cohen 1988). The deviation between human and AI coders is practically relevant for only two categories.

Table 6 reports a medium deviation between human and AI coders for "profitability aspects" and "process orientation." In both cases, the AI is less conservative than the human coders. In other words, the AI more frequently assigns a task to a higher level than the human coders. For example, human coders rated 4.3% of the tasks at level two of "process orientation", whereas AI rated 12.7% of the tasks at level two. The same is true for "profitability aspects" (8.7% and 20% at level two). In all other cases, the deviation is only small. A possible explanation for this deviation is that human raters are able to take additional information into account when assessing a task, e.g. a task's illustration, tables, and pictures for their ratings as well as information provided within the textbook (but not within the task), while AI has to rely only on the textual information. The next chapter discusses the reported findings.

### Discussion and conclusions

The current study analyzed 3361 tasks from 14 accounting textbooks with the help of AI. The results indicate that not all accounting-relevant characteristics proposed in this study are well integrated. In terms of process orientation, the tasks yielded



above-average results compared to the other categories. This result can be attributed to the redesign of German occupational curricula, which follow the principle of process chains. Textbook authors have most likely adapted to this change to comply with the requirements of the curricula. Nevertheless, there is still room for improvement, as the majority of the tasks do not provide information on previous or forthcoming business processes, although the focus on business processes is seen as vital when it comes to the education of future accountants (e. g. Preiß 2015; The Pathway Commission 2012; Walker and Ainsworth 2001). However, the results regarding process orientation are in line with previous research findings, indicating that tasks mainly focus on isolated processes, while neglecting the consideration of complete business processes (Bloemen 2011). What is more, a high level of process orientation is positively related to a number of other relevant task characteristics (e.g. problem identification, description of social processes, cash flows and goods and services, translating a real situation to accounting and operating within the accounting system), highlighting the importance of designing process-oriented accounting tasks.

When it comes to providing real world information, tasks lack a detailed and realistic description of occurring social processes, value processes, and cash flows. Referring to chapter 2, both low process orientation and abstract contextual information are problematic, as learners do not get to see what happens in the real world. This lack of reference to reality can subsequently lead to learning difficulties (e.g. Tramm et al. 1996) and limit learners' ability to perform in the real work environment, as they are not able to link the real world and the accounting world as part of their training. Although descriptions of goods, cash flows and social processes are both positively related with one another and to other categories (e.g. identification of problems, translating and operating within the accounting system), some negative correlations seem problematic. For example, tasks providing a more detailed description of goods and services tend to not promote a learners' ability to interpret these processes with regard to liquidity and profitability aspects. The same is true for cash flows and the interpretation regarding profitability. This could lead to an insufficient illustration of the connection between real business processes and their impact on corporate goals.

Considering that accountants work with (digital) documents on a daily basis, textbook tasks do not foster the important skill of correctly handling them, with only 3% of the tasks containing realistic documents. Tasks with no documents, low levels of process orientation and abstract descriptions of business situations are not suitable for enabling learners to gain a "broad and interdisciplinary view of the work environment" (Walker and Ainsworth 2001: 41) and a profound understanding of complex business processes.

A further cause for concern is the lack of problem orientation. The vast majority of the tasks do not contain a problem or clearly outline the problem including potential solution paths, leading to the assumption that many tasks focus on the reproduction or application of knowledge. This indicates a knowledge/skill gap in training, because identifying, analyzing, and solving unstructured accounting problems are vital skills demanded by potential employers (e.g. Jackling and Lange 2009; Kavanagh and Drennan 2008). These findings are also in line with existing results, as for example the findings of Bloemen (2011: 135) show that more than half of the tasks do not include a problem.



As outlined in chapter 2, the ability to gather and assess information is seen as vital skill when educating future accountants. However, findings indicate that tasks do not (or in only limited fashion) promote skills regarding the evaluation of information, as no task requires assessing the relevance of included information and the majority of the tasks do not require to search for additional information.

Regarding modeling competencies, results further indicate that accounting tasks split the complete modeling cycle into separate steps and mainly focus on applying formal accounting rules (operate), whereas the translation of an economic phenomenon into accounting as well as interpreting and validating the solution is required less frequently. This focus on individual steps and separation of tasks does not allow learners to go through the complete modeling cycle. Tasks especially emphasize operating within the accounting system. Thus, tasks in accounting textbooks focus on skills that allow accountants to acquire routine in dealing with accounting rules and concepts. Choosing the right concept for a business situation and validating the usefulness of the chosen concepts seem to be less important. The focus on operations within the accounting system (e.g. calculating a cash discount, posting an invoice, etc.) is problematic and out-of-date, as many calculations and postings are performed automatically by modern accounting systems and therefore do not need the assistance of human accountants (Hmyzo and Muzzu 2020; Klein and Küst 2020). These advances in accounting systems allow accountants to “spend less time on mundane tasks [...] and more time on activities such as evaluating outliers, improving business processes, making judgments and presenting findings” (Blix et al. 2021: 2). Detecting mistakes and outliers, and improving business processes means learners need to know if and how business situations affect variables such as a company’s profitability or liquidity. Hence, they need to interpret data and business situations, rather than just document them. These kinds of analytical skills are demanded by employers and accounting educators alike (e.g. Kavanagh and Drennan 2008), even though they currently remain insufficiently promoted by learning tasks. When interpretation is necessary, tasks focus on profitability aspects and neglect liquidity aspects. The emphasis on profit also seems problematic, as liquidity is crucial for the survival of a company (Albrecht and Sack 2000; Preiß 2015).

Although the findings of the presented study are mainly descriptive, they outline some key weaknesses of accounting tasks and provide practical implications for the design of tasks in textbooks:

- The need for modifying and adapting accounting education evolves from a rapidly changing work environment. From existing research we know that digitalization and new technologies massively influence both “typical” work activities of accountants and relevant skills, competencies and knowledge (e.g. Hmyzo and Muzzu 2020; Kavanagh and Drennan 2008; Klein and Küst 2020; McKinney et al. 2017). Increasing data availability, new possibilities for data preparation and visualization, the automation of processes (e.g. robotic process automation, AI) as well as new business models (Klein and Küst 2020) change the skills necessary for successfully working in the accounting area. For example, in a largely automated workplace, bots or AI complete large parts of routine activities (Hmyzo and Muzzu 2020), massively reducing repetitive work (e.g. posting records) for accounting employees. However, results of our

study indicate that tasks largely focus on promoting formal accounting techniques and on applying accounting rules, leading to learners who are capable of routinely recording and documenting basic business situations, but fail to cope with more challenging and complex situations. Therefore, responding to a changed work environment and referring to the modeling cycle, task creators are encouraged to put less emphasis on skills regarding the recording and documentation of business situations in the accounting system. Reducing the number of (routine) tasks, which focus on applying accounting rules, provides time to address other relevant skills, such as the interpretation of data, the validation of solutions or the search and evaluation of information. According to our research results, these skills are insufficiently promoted by existing accounting tasks.

- When it comes to interpreting data, results indicate that tasks focus on profitability aspects rather than liquidity aspects. In order to ensure a broad economic understanding, liquidity aspects should be equally addressed.
- Accountants work in a world of symbols, figures, data, documents and information. Their work activities are characterized by the fact that they do not have an immediate and direct influence on the “real world”. This results in two consequences. First, accounting work is never an end in itself, but its purpose is to prepare, document or evaluate real business processes. Second, interventions in the “real world” can only be made indirectly (Neuweg 2020; Tramm 2009, 2010). A lack of connection to real business processes is problematic and challenging for learners (and often reason for learning difficulties, e.g. Berding et al. 2019; Tramm et al. 1996), as accounting is then seen as abstract a formal system with no relevance for real business situations. Therefore, accounting tasks should be designed in a way that they allow learners to gain insight into business processes and connect the “accounting world” to the “real world”. This can be done by designing tasks process-oriented, provide illustrations and pictures, include relevant documents or describing a situation in more detail. We once again highlight the importance of documents, as they are a representation of reality (Berding et al. 2019; Neuweg 2020).
- The importance of problem orientation in accounting education was addressed frequently and strongly in the past (e.g. Cunningham 2014; Stanley and Marsden 2012; The Pathway Commission 2015). However, results indicate that “the ability to identify and solve unstructured problems” (Sundem 2014: 623) is insufficiently promoted by accounting schoolbook tasks. Therefore, we encourage task designers to increase problem orientation by a) designing tasks, which contain a realistic accounting problem b) designing tasks in an open manner (no precise description of the problem) and not defining clear solution paths c) creating an information surplus (irrelevant information) while leaving out necessary information and thereby demanding information research.

However, it is important to keep several aspects in mind when interpreting and discussing the findings and drawing conclusions. First, we selected specific criteria for evaluation which are proposed in accounting education research, neglecting other possible relevant criteria for analyzing accounting tasks. Therefore, the findings from this study do not report on task quality in general, but only if and to what extent accounting

tasks meet the criteria proposed in this study. Second, we acknowledge that one single task cannot (and should not) meet all criteria at the highest level. It is indeed necessary and useful that tasks focus on specific competencies (e.g. operating within the accounting system or interpreting data), especially when beginners are involved. Considering the totality of the tasks, all criteria should be met to a sufficient degree. For example, although it is not necessary (or useful) that one task requires a learner to pass through all four steps of the modeling cycle, in the course of training, learners should in fact be required to translate, operate, interpret, and validate. The same is true for other criteria. Third, we acknowledge that task designers are limited by page numbers when designing tasks for schoolbooks and face specific requirements by publishers, school authorities and curricula. In addition, our study only included tasks from textbooks. However, of course teachers can and will use other sources for tasks.

Regarding inter-coder reliability for human codings, the results show that values are partly unsatisfactory. This might be due to several reasons. First, the results indicate that values are insufficient, especially for complex categories such as “translate,” “validate,” or “assessing the relevance of information.” Rating these categories is a difficult and challenging task, leading to lower values of inter-coder reliability. Other reasons for poor values might be an inaccurate codebook or insufficient rater training. New measures for calculating inter-coder reliability are necessary to correctly and adequately analyze the quality of the ratings. As can be seen in Table 3, the percentage of agreement and Krippendorff’s  $\alpha$  vary tremendously regarding some categories. For example, the percentage of agreement (tolerance is 0) is 88% for “validate”, whereas Krippendorff’s  $\alpha$  is 0.258. This might be due to the uneven distribution of this category (nearly 100% of the tasks are rated at level 0). Regarding the interpretation based on liquidity aspects, the percentage of agreement (tolerance is 0) is 96%, and Krippendorff’s  $\alpha$  is -0.009. Therefore, the information value of these reliability measures is limited. We tried to ensure the best validity possible by conducting two rater trainings, and two people from the author team rated 1,080 tasks in total. In fact, the quality of AI predictions heavily depends on the quality of the human coding, which represents the input data of the training. As the simulation study by Song et al. (2020) indicates, the prediction error of AI can be reduced if human coders highly agree on the judgments of a coding unit.

In terms of using AI assistance to perform content analysis, the results indicate that AI is a useful help when analyzing a large number of tasks. However, rating tasks is a complex endeavor, and the findings show deviations between human and AI coding, especially (again) when considering complex categories such as process orientation, or if a task requires the identification of problems or the translation of situations into an accounting model. Regarding less complex categories (e.g. document orientation, operate, or interpret), AI and human coders yield similar results. Therefore, in future studies AI assistance can and should be used for categories that are low in complexity, giving researchers the opportunity to focus on more challenging ones.

This paper closes with potential for future research. First, it might be useful to investigate cross-textbook differences, as task characteristics may vary across different textbooks. Second, as the study did not detect whether task characteristics vary across topics, or if the complexity of the tasks (e.g. more parts of the modeling cycle contained in a task) increases over the course of a book, further analyses might be helpful. Third,

the study does not differentiate between different types of tasks (e.g. exercises, knowledge questions, case studies, etc.), which necessarily focus on different aspects. Therefore, future studies should specifically address differences between different task types.

#### Acknowledgements

Not applicable.

#### Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

#### Funding

Not applicable.

#### Availability of data and materials

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

#### Declarations

##### Competing interests

The authors declare that they have no competing interests.

Received: 15 June 2022 Accepted: 31 October 2022

Published online: 12 November 2022

#### References

- Accounting Education Change Commission (1992) The first course in accounting: position statement no. two. *Issues Account Educ* 7:307–312
- Achtenhagen F (1996) Entwicklung ökonomischer Kompetenz als Zielkategorie des Rechnungswesenunterrichts. In: Preiß P, Tramm T (eds) *Rechnungswesenunterricht und ökonomisches Denken: Didaktische Innovationen für die kaufmännische Ausbildung*. Gabler, Wiesbaden, pp 22–44
- Albrecht WS, Sack RJ (2000) Accounting education: charting the course through a perilous future. *Accounting education series*, vol 16. American Accounting Association, Sarasota, FL
- Alpaydin E (2019) *Maschinelles Lernen*. De Gruyter, Oldenburg
- Arek-Bawa O, Dhunpath R (2017) Assessment and cognitive demand in higher education accounting textbooks. *Alternation* 24:140–166. <https://doi.org/10.29086/2519-5476/2017/v24n2a8>
- Benoit K, Watanabe K, Wang H, Müller S, Perry PO, Lauderdale B, Lowe W (2020) *quanteda.textmodels: scaling models and classifiers for textual data* [Computer software]. <https://CRAN.R-project.org/package=quanteda.textmodels>
- Bentler PM, Dudgeon P (1996) Covariance structure analysis: statistical practice, theory, and directions. *Annu Rev Psychol* 47:563–592. <https://doi.org/10.1146/annurev.psych.47.1.563>
- Berding F, Lau I (2018) Epistemic messages in textbooks for vocational education and training. *J Educ Media Memory Soc* 10:39–63. <https://doi.org/10.3167/jemms.2018.100203>
- Berding F, Beckmann A, Kürten V (2019) Modellieren mit dem Rechnungswesen: Entwicklung eines rasch-konformen Messverfahrens für erfolgswirksame Vorgänge und didaktische Implikationen. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik* 115:567–602
- Berding F, Jahncke H, Holt K (2021) Learning Analytics in der Wirtschaftspädagogik: Eine Simulationsstudie für die Anwendung überwachter maschineller Lernens für Inhaltsanalysen am Beispiel von Grundvorstellungen und (Selbst) Reflexionskompetenz. *Zeitschrift für Berufs- und Wirtschaftspädagogik, Sonderheft*:237–291
- Blix LH, Edmonds MA, Sorensen KB (2021) How well do audit textbooks currently integrate data analytics? *J Account Educ* 55:100717. <https://doi.org/10.1016/j.jaccedu.2021.100717>
- Bloemen A (2011) *Lernaufgaben in Schulbüchern der Wirtschaftslehre: Analyse, Konstruktion und Evaluation von Lernaufgaben für die Lernfelder industrieller Geschäftsprozesse*. Schriften zur Berufs- und Wirtschaftspädagogik. Rainer Hampp, Augsburg
- Blum W (2011) Can modelling be taught and learnt? Some answers from empirical research. In: Kaiser G, Blum W, Borromeo Ferri R, Stillman G (eds) *Trends in teaching and learning of mathematical modelling*. Springer Netherlands, Dordrecht, pp 15–30
- Brötz R, Kock A, Annen S, Schaal T (2015) Gemeinsamkeiten und Unterschiede der kaufmännischen Ausbildungsberufe. In: Brötz R, Kaiser F (eds) *Kaufmännische Berufe: Charakteristik, Vielfalt und Perspektiven*. Bertelsmann, Bielefeld, pp 91–106
- Bundesinstitut für Berufsbildung (2018) *Verzeichnis der anerkannten Ausbildungsberufe 2018*, Bonn
- Chiang B, Nouri H, Samanta S (2014) The effects of different teaching approaches in introductory financial accounting. *Account Educ Int J* 23:42–53. <https://doi.org/10.1080/09639284.2013.833724>
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*. Psychology Press, New York
- The Pathway Commission (2015) *In Pursuit of Accounting's Curricula of the Future*
- Crouch R, Haines C (2004) Mathematical modelling: transitions between the real world and the mathematical model. *Int J Math Educ Sci Technol* 35:197–206. <https://doi.org/10.1080/00207390310001638322>

- Cunningham BM (2014) Developing critical thinking in accounting education. In: Wilson RMS (ed) *The Routledge companion to accounting education*. Routledge, London, pp 399–419
- Davidson RA (2005) Analysis of the complexity of writing used in accounting textbooks over the past 100 years. *Account Educ Int J* 14:53–74. <https://doi.org/10.1080/09639280410001695241>
- Davidson RA, Baldwin BA (2005) Cognitive skills objectives in intermediate accounting textbooks: Evidence from end-of-chapter material. *J Account Educ* 23:79–95. <https://doi.org/10.1016/j.jaccedu.2005.05.001>
- Dockter DL (2012) Problem-based learning in accounting. *Am J Bus Educ* 5:547–554. <https://doi.org/10.19030/ajbe.v5i5.7211>
- Ernst F (2011) Lesbarkeit von Rechnungswesenbüchern an kaufmännischen Berufsschulen. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik* 107:408–423
- Ernst F (2012) Fachdidaktische Analyse von Lehrbüchern für den Rechnungswesenunterricht in Deutschland und den USA
- Ertel W (2016) *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*, 4th edn. Springer, Wiesbaden
- Federal Ministry of Education and Research (2020) *Report on Vocational Education and Training 2019*
- Feng G (2013) Factors affecting intercoder reliability: a Monte Carlo experiment. *Qual Quant* 47:2959–2982. <https://doi.org/10.1007/s1135-012-9745-9>
- Feng G, Zhao X (2016) Do not force agreement. *Methodology* 12:145–148. <https://doi.org/10.1027/1614-2241/a000120>
- Ferguson J, Collison D, Power D, Stevenson L (2006) Accounting textbooks: exploring the production of a cultural and political artifact. *Account Educ Int J* 15:243–260. <https://doi.org/10.1080/09639280600850679>
- Ferguson J, Collison D, Power D, Stevenson L (2010) The views of ‘knowledge gatekeepers’ about the use and content of accounting textbooks. *Account Educ Int J* 19:501–525. <https://doi.org/10.1080/09639281003594294>
- Finney S, Di Stefano C (2013) Nonnormal and categorical data in structural equation modeling. In: Hancock GR, Mueller RO (eds) *Structural equation modeling: a second course*, 2nd edn. Information Age Publ, Charlotte, N.C., pp 439–492
- Flood B (2014) The case for change in accounting education. In: Wilson RMS (ed) *The Routledge companion to accounting education*. Routledge, London, pp 81–101
- Gamer M, Lemon J, Fellows Puspendra Singh I (2019) irr: various coefficients of interrater reliability and agreement. <https://CRAN.R-project.org/package=irr>
- Geiser C (2013) *Data analysis with Mplus*. Guilford Press
- Golyagina A, Valuckas D (2016) Representation of knowledge on some management accounting techniques in textbooks. *Account Educ Int J* 25:479–501. <https://doi.org/10.1080/09639284.2016.1211942>
- Große CS (2015) Fostering modeling competencies: benefits of worked examples, problems to be solved, and fading procedures. *Eur J Sci Math Ed* 3:364–375. <https://doi.org/10.30935/scimath/9444>
- Guggemos J (2016) Modellierung und Messung von Kompetenz im externen Rechnungswesen. Dr. Hut, München
- Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, Bing G (2017) Learning from class-imbalanced data: review of methods and applications. *Expert Syst Appl* 73:220–239. <https://doi.org/10.1016/j.eswa.2016.12.035>
- Hansen JD (2006) Using problem-based learning in accounting. *J Educ Bus* 81:221–224. <https://doi.org/10.3200/JOEB.81.4.221-224>
- Hartmann J, Huppertz J, Schamp C, Heitmann M (2019) Comparing automated text classification methods. *Int J Res Mark* 36:20–38. <https://doi.org/10.1016/j.jresmar.2018.09.009>
- Hattie J (2012) *Visible learning for teachers: Maximizing impact on learning*. Routledge, London, New York
- Hayes AF, Krippendorff K (2007) Answering the call for a standard reliability measure for coding data. *Commun Methods Meas* 1:77–89. <https://doi.org/10.1080/19312450709336664>
- Helm C (2016) Welche Denkschritte durchlaufen Schüler/innen beim Erstellen von Buchungssätzen? *Wissenplus* 16/17:38–41
- Hmyzo E, Muzzu A (2020) Technologie im Rechnungswesen—Wenn die Maschine besser und schneller bucht. In: Berding F, Jahncke H, Slopinski A (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer, Wiesbaden, pp 99–116
- Hove D, Jorgensen TD, van der Ark LA (2018) On the usefulness of interrater reliability coefficients. In: Wiberg M, Culpepper S, Janssen R, González J, Molenaar D (eds) *Quantitative psychology*, vol 233. Springer International Publishing, Cham, pp 67–75
- Jackling B, Lange P (2009) Do accounting graduates’ skills meet the expectations of employers? A matter of convergence or divergence. *Account Educ Int J* 18:369–385. <https://doi.org/10.1080/09639280902719341>
- Jordanski G (2020) Kaufmännische Steuerung und Kontrolle im 4.0 Arbeitsumfeld—Anforderungen an duale Ausbildungsberufe. In: Berding F, Jahncke H, Slopinski A (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer, Wiesbaden, pp 59–82
- Kavanagh MH, Drennan L (2008) What skills and attributes does an accounting graduate need? Evidence from student perceptions and employer expectations. *Accounting & Finance* 48:279–300
- Kleesiek J, Murray JM, Strack C, Kaissis G, Braren R (2020) Wie funktioniert maschinelles Lernen? A primer on machine learning (A primer on machine learning). *Radiologe* 60:24–31. <https://doi.org/10.1007/s00117-019-00616-x>
- Klein J, Küst C (2020) Wie die Digitalisierung im Rechnungswesen die Aufgaben und Anforderungen an die Mitarbeiter/-innen verändert. In: Berding F, Jahncke H, Slopinski A (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer, Wiesbaden, pp 83–98
- Krippendorff K (2016) Misunderstanding reliability. *Methodology* 12:139–144
- Krippendorff K (2004) *Content analysis: an introduction to its methodology*. SAGE Publications
- Laksmiana I, Tietz W (2008) Temporal, cross-sectional, and time-lag analyses of managerial and cost accounting textbooks. *Account Educ* 17:291–312. <https://doi.org/10.1080/09639280701788745>
- Lang M, Binder M, Richter J, Schratz P, Pfisterer F, Coors S, Au Q, Casalicchio G, Kotthoff L, Bischl B (2019) mlr3: a modern object-oriented machine learning framework in R. *JOSS* 4:1903. <https://doi.org/10.21105/joss.01903>
- Lanquillon C (2019) Grundzüge des maschinellen Lernens. In: Schacht S, Lanquillon C (eds) *Blockchain und maschinelles Lernen: Wie das maschinelle Lernen und die Distributed-Ledger-Technologie voneinander profitieren*. Springer, Wiesbaden, pp 89–142

- Lorena AC, Jacintho LF, Siqueira MF, de Giovanni R, Lohmann LG, de Carvalho AC, Yamamoto M (2011) Comparing machine learning classifiers in potential distribution modelling. *Expert Syst Appl* 38:5268–5275. <https://doi.org/10.1016/j.eswa.2010.10.031>
- Mayring P (2015) *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, 12th edn. Beltz Pädagogik. Beltz, Weinheim
- McKinney E, Yoos CJ, Snead K (2017) The need for ‘skeptical’ accountants in the era of Big Data. *J Account Educ* 38:63–80. <https://doi.org/10.1016/j.jaccedu.2016.12.007>
- Muthén LK, Muthén BO (2017) *Mplus user's guide*. Muthén & Muthén, Los Angeles, CA
- Neuweg GH (2020) Das Linzer Ebenen-Modell als Instrument zur Ausbildung des denkenden Buchhalters. In: Greimel-Fuhrmann B, Fortmüller R (eds) *Wirtschaftsdidaktik - den Bildungshorizont durch Berufs- und Allgemeinbildung erweitern: Festschrift für Josef Aff. Facultas, Wien*, pp 135–144
- Papilioud C, Hinneburg A (2018) *Qualitative Textanalyse mit Topic-Modellen: Eine Einführung für Sozialwissenschaftler*. Springer Fachmedien, Wiesbaden
- Preiß P (2005) Entwurf eines Kompetenzmodells für den Inhaltsbereich Rechnungswesen/Controlling. In: Gonon P, Klausner F, Nickolaus R, Huisinga R (eds) *Kompetenz, Kognition und neue Konzepte der beruflichen Bildung*. VS Verl. für Sozialwiss, Wiesbaden, pp 67–85
- Preiß P, Tramm T (1996) Die Göttinger Unterrichtskonzeption des wirtschaftsinstrumentellen Rechnungswesens. In: Preiß P, Tramm T (eds) *Rechnungswesenunterricht und ökonomisches Denken: Didaktische Innovationen für die kaufmännische Ausbildung*. Gabler, Wiesbaden, pp 222–323
- Preiß P (2015) Kaufmännische Steuerung und Kontrolle als Kernqualifikation kaufmännischer Ausbildung—von der Dokumentation zur Steuerung der Geschäftsvorfälle als Arbeitsprozesse im Rahmen von Geschäftsprozessen. In: Brötz R, Kaiser F (eds) *Kaufmännische Berufe: Charakteristik, Vielfalt und Perspektiven*. Bertelsmann, Bielefeld, pp 189–205
- Rebele JE, St. Pierre EK, (2019) A commentary on learning objectives for accounting education programs: the importance of soft skills and technical knowledge. *J Account Educ* 48:71–79. <https://doi.org/10.1016/j.jaccedu.2019.07.002>
- Scharkow M (2011) Zur Verknüpfung manueller und automatischer Inhaltsanalyse durch maschinelles Lernen. *Medien & Kommunikationswissenschaft* 59:545–562. <https://doi.org/10.5771/1615-634x-2011-4-545>
- Song H, Tolochko P, Eberl J-M, Eisele O, Greussing E, Heidenreich T, Lind F, Galyga S, Boomgaarden HG (2020) In validations we trust? The impact of imperfect human annotations as a gold standard on the quality of validation of automated content analysis. *Polit Commun* 37:550–572. <https://doi.org/10.1080/10584609.2020.1723752>
- Staatsinstitut für Schulqualität und Bildungsforschung (2004) *Lehrplanrichtlinien für die Berufsschule*
- Stanley T, Marsden S (2012) Problem-based learning: does accounting education need it? *J Account Educ* 30:267–289. <https://doi.org/10.1016/j.jaccedu.2012.08.005>
- Sundem GL (2014) Fifty years of change in accounting education: the influence of institutions. In: Wilson RMS (ed) *The Routledge companion to accounting education*. Routledge, London, pp 611–631
- The Pathway Commission (2012) *Charting a national strategy for the next generation of accountants*, London
- Thoma M, Schumacher V (2018) Lernaufgaben in Schulbüchern – Empirische Befunde zum kognitiven Aktivierungspotenzial im Fach Rechnungswesen. *bwp@ Spezial AT-1: Wirtschaftspädagogische Forschung und Impulse für die Wirtschaftsdidaktik – Beiträge zum 12. Österreichischen Wirtschaftspädagogikkongress*:1–18
- Tramm T, Hinrichs K, Langenheilm H (1996) Lernschwierigkeiten im Buchführungsunterricht. In: Preiß P, Tramm T (eds) *Rechnungswesenunterricht und ökonomisches Denken: Didaktische Innovationen für die kaufmännische Ausbildung*. Gabler, Wiesbaden, pp 158–221
- Tramm T, Goldbach A (2005) Gestaltungsprinzipien und theoretische Grundlagen innovativer Schulbücher zur ökonomischen Berufsbildung—am Beispiel der „prozessorientierten Wirtschaftslehre“. *Wirtschaft und Erziehung*:203–213
- Tramm T (2009) Von der Geschäftsprozess- zur Lernprozessperspektive: Das Zusammenspiel von Prozessorientierung, systemischer Perspektive und prozessübergreifender Kompetenzentwicklung im lernfeldstrukturierten Berufsschulunterricht. In: Pongratz H (ed) *Prozessorientierte Wirtschaftsdidaktik und Einsatz von ERP-Systemen im kaufmännischen Unterricht*. Shaker, Aachen, pp 77–101
- Tramm T (2010) Berufliche Kompetenzentwicklung im Kontext kaufmännischer Arbeits- und Geschäftsprozesse. In: Schapfel-Kaiser F (ed) *Anforderungen an kaufmännisch-betriebswirtschaftliche Berufe: Berichte zur beruflichen Bildung*. Bertelsmann, Bielefeld, pp 65–88
- Trigo A, Belfo F, Estébanez RP (2016) Accounting information systems: evolving towards a business process oriented accounting. *Proc Comput Sci* 100:987–994. <https://doi.org/10.1016/j.procs.2016.09.264>
- Türling J (2014) *Die professionelle Fehlerkompetenz von (angehenden) Lehrkräften: Eine empirische Untersuchung im Rechnungswesenunterricht*. Springer, Wiesbaden
- Unesco Institute for Statistics (2012) *International standard classification of education (ISCED) 2011*. UNESCO Institute for Statistics
- van der Kolk B (2019) Ethics matters: the integration of ethical considerations in management accounting textbooks. *Account Educ* 28:426–443. <https://doi.org/10.1080/09639284.2018.1543602>
- Venables WN, Ripley BD (2007) *Modern applied statistics with statistics and computing*, 4th edn. Springer
- Walker KB, Ainsworth PL (2001) Developing a process approach in the business core curriculum. *Issues Account Educ* 16:41–66. <https://doi.org/10.2308/iaec.2001.16.1.41>
- Watty K (2014) Generic skills within the accounting curriculum. In: Wilson RMS (ed) *The Routledge companion to accounting education*. Routledge, London, pp 276–293
- Weil RL, O'Brien PC, Maher MM, Stickney CP, Davidson S (1999) *Accounting: The language of business*, 10th edn. T. Horton and Daughters, Sun Lakes, Ariz
- Wells PK (2018) How well do our introductory accounting text books reflect current accounting practice? *J Account Educ* 42:40–48. <https://doi.org/10.1016/j.jaccedu.2017.12.003>
- Wijffels J, Straka M, Straková J (2019) *udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the 'UDPipe' NLP Toolkit [Computer software]*. <https://CRAN.R-project.org/package=udpipe>
- Winther E, Achtenhagen F (2008) Kompetenzstrukturmodell für die kaufmännische Bildung. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik* 104:511–538



- Wright MN, Ziegler A (2017) ranger : a fast implementation of random forests for high dimensional data in C++ and R. *J Stat Soft* 77:1–17. <https://doi.org/10.18637/jss.v077.i01>
- Wuttke E, Seeber S, Geiser C, Turhan L (2022) Zur Problemhaltigkeit von Aufgaben in kaufmännischem Abschluss und Zwischenprüfungen—Ergebnisse aus Aufgabenanalysen. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik* 118:25. <https://doi.org/10.25162/zbw-2022-0002>
- Zhao X, Liu JS, Deng K (2013) Assumptions behind Inter-coder reliability Indices. *Ann Int Commun Assoc* 36:419–480. <https://doi.org/10.1080/23808985.2013.11679142>
- Zhao X, Feng G, Liu J, Deng K (2018) We agreed to measure agreement—redefining reliability de-justifies Krippendorff's alpha. *China Media Res* 14:1–15

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---