

## RESEARCH ARTICLE

# The quality of risk reports: Integrating requirement levels of standard setters into text analysis

Nils Helms  | Reinhold Hölscher | Matthias Nelde | Jochen Schneider

Department of Business Studies and Economics, Technical University Kaiserslautern, Kaiserslautern, Germany

## Correspondence

Nils Helms, Department of Business Studies and Economics, Technical University Kaiserslautern, Kaiserslautern, Germany. Email: [helms@wiwi.uni-kl.de](mailto:helms@wiwi.uni-kl.de)

[Correction added on 17 September 2021, after first online publication: The legal statement had been corrected in this version.]

## Abstract

The intention of this paper is to shed light on the analysis of financial disclosure through the integration of requirement levels. This in return will lead to the development of a general applicable evaluation methodology based on Bloom's taxonomy system. Therefore, it will be possible to explicitly consider the relevance of the given information. To underline the appropriateness of our method, we combine the requirement levels with a qualitative content analysis. Based on the German accounting standard DRS 20, we clarify the respective application of the requirement levels in the context of the qualitative content analysis. Hence, we will discuss the limitations of our developed approach. In addition, we analyze further areas of application in the context of qualitative analysis of financial disclosure. All things considered, it is evident that our chosen approach, through the integration of a taxonomy system, contributes to the validity of established text analyzing methods.

## KEYWORDS

DRS 20, qualitative content analysis, requirement level, risk report, text analysis

## 1 | INTRODUCTION

Corporate decisions and actions are future-oriented, and therefore subject to risk. In order to inform stakeholders on an individual level about the corporate risk taking initiative, companies frequently publish risk reports. Based on the importance of risk reporting in the eyes of external addressees, risk reports have become an intensively reviewed subject in business research. Of particular interest are procedures measuring the quality of risk reports.

Henceforth, we will present and discuss a newly developed methodology that integrates the expediency of the information into the assessment of risk report quality. The relevance of the information is derived from the legal requirements and standards on which the risk reports are based upon.

This paper is structured in the following way: In Chapter 2, we describe existing measures used to assess the quality of financial disclosures. The description of existing measures is followed by a summarization of the status quo of the current research. Subsequently, Chapter 2 will be closed by the introduction of the so-called requirement level. The concept underlying the requirement level forms the basis for integrating the varied relevancy of information when evaluating the quality of risk reports. With the purpose of integrating the requirement levels into quality evaluation, we apply Bloom's taxonomy system.

Chapter 3 will begin with the explanation of Bloom's classification system. In this context, we will show how taxonomic levels can be applied in order to identify and delineate various requirement levels. The taxonomy levels are then combined with qualitative content analysis.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Corporate Accounting & Finance* published by Wiley Periodicals LLC

Qualitative content analysis is an established methodology used for the assessment of risk reports. By doing so we can show how varying degrees of information, in regards to their decision usefulness, can be systematically included when measuring reporting quality. The methodology just mentioned will be illustrated in Chapter 3.3 using the DRS 20 as an example. The individual categories to be reported according to DRS 20 are transferred into the taxonomy system with the intention of evaluating the usefulness of the information. Chapter 3 will be rounded off by a discussion about the limitations of our developed approach.

Chapter 4.1 then looks at how a scoring procedure can be used to combine the assessments of the individual categories into a total score. The weighting of the subscore values is again carried out using the taxonomy levels. The limits of this approach are discussed in Chapter 4.1 too. Furthermore, Chapter 4.2 examines whether the concept of requirement levels can also be integrated into other procedures for assessing the quality of the risk report. We show that an application is not only possible in the context of qualitative content analysis, but also in other text analyzing procedures. This article ends with a summary of the central findings.

## 2 | THE STATUS QUO OF QUALITATIVE MEASUREMENT OF RISK REPORT QUALITY

### 2.1 | Methodology and research status of text analysis

Measuring the quality of external corporate reporting and in particular risk reporting has a long tradition in business research (Miihkinen, 2012). A common similarity of all research approaches is that the assessment of the quality underlying the report is observed by the perceived usefulness to the reader. Nevertheless, a considerable variety regarding the design of the different approaches can be found. This variety relates to the measurement of the quality of the report and the conceptual focus of the evaluations. Furthermore, it must be mentioned that the definition of “quality” varies in its detail among the different approaches.

Basically, the different methodologies can be classified into quantitative and narrative-based approaches.

The quantitative analysis uses solely number-based information, for instance, from the balance sheet, the income statement, or key risk figures in the risk report (Pérignon & Smith, 2010). The basic idea of quantitative studies is to enable the addressees to estimate the future development of companies’ performance based on historical information (Ryan, 2012).

The usefulness relating to the decision making process is regarded as high, if meaningful forecasts can be derived out of the current values regarding future characteristics of annual key figures (Lindemann, 2006).

Narrative-based approaches aim at analyzing text fragments of external reports. We will focus on the narrative-based approaches in this paper. This procedure is generally referred to as text analysis (Loughran & McDonald, 2016). Thereby special emphasis is placed on the rule-based text analysis (Beattie et al., 2004).<sup>1</sup> The existing methods are categorized as follows: Figure 1 at the highest level encompasses all methods that can be differentiated according to their type of evaluation (Elshandidy et al., 2018; Hassan & Marston, 2019; Li, 2010).

Manual evaluation of the quality is assessed by reading and interpreting the reports. In addition, there are also text-based quality measures that enable evaluation by means of automated procedures based on algorithms.

Approaches based upon automatic evaluation with algorithms include evaluation methods of readability, lexicographic and document similarity methods. Regarding the readability method: when the published text is easy to read and easy to be understood, then the provided information is of high quality and therefore very useful regarding the decision-making process (Gunning, 1971; Loughran & McDonald, 2016). The parameters to determine the readability originate from US schoolbooks over specific school years (Reck, 2016; Zhou et al., 2017). This concept can be transferred into the context of report quality assessment, by raising the question of whether most readers would understand the given text without being cognitively challenged, and if the intended message is conveyed easily (Hoitash & Hoitash, 2018). The readability is measured by quantitative characteristics: for example, the average number of words in a sentence, average number of syllables in a word, or the average word length (Reck, 2016).

Lexicographical methods identify and count specific key words in order to analyze the sentiment and intonation conveyed indirectly by the text passages (Kaya & Seebeck, 2018; Loughran & McDonald, 2016). If the general sentiment conveyed corresponds to the quantitative facts, the passages therefore are ranked high in terms of their decision usefulness. If many negatively labeled words are used in a certain text section (e.g., the described facts show a negative development in present or future key figures) a high decision usefulness exists. The methods that have been developed differ in the bag

<sup>1</sup> At the same time, analysts, experts, and private individuals also use non-rule-based quality measures, such as questionnaires or interviews. These are not considered, because of their minimal theoretical foundation (Baron et al., 1999; Gerding, 2016; Hassan & Marston, 2019; Koonce et al., 2005).

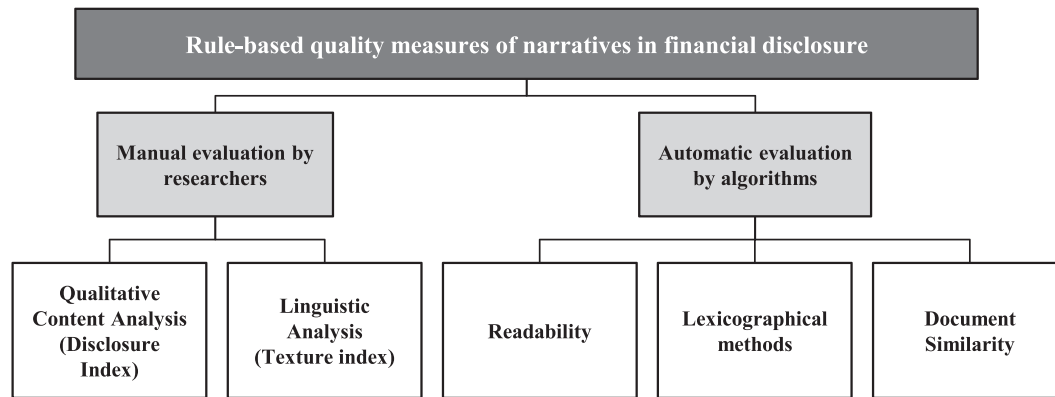


FIGURE 1 Rule-based methods to evaluate the text quality

of words used to define specific sentiment characteristics (Elshandidy et al., 2018). Mathematical-statistical methods like N-Grams, Support Vector Machines, and Bayesian Methods (Loughran & McDonald, 2016) exist in addition to linguistically sound word lists, and subjective compilations by researchers (Campbell et al., 2014; Elshandidy et al., 2013; Loughran & McDonald 2011; Yekini et al., 2016).

The document similarity evaluation method analyzes text passages in order to find substantial similarities concerning prior publications by the entity. Furthermore, it is possible to search for similarities among other entities included in the data sample (Loughran & McDonald, 2016). A strong resemblance presumes that the text is composed of prefabricated standard elements (boilerplate) which have only a minor relation to the actual company situation (Loughran & McDonald, 2016). A high degree of similarity reduces the decision usefulness, and thereby the quality of the report. In the case of high-quality company reports, for example, only slight similarities can be found, since the text components are individually adapted to the current situation of the company.

The manual evaluation methods used by researchers can be divided into linguistic analysis and qualitative content analysis.

The focus of linguistic analysis lies in the evaluation of the text in terms of phonological, morphological, syntactical, semantical, structural, and contextual characteristics (Beattie et al., 2004; Roseberry, 1995). Since these linguistic features are more than just a count of textual elements, as in the case of readability, the features must be manually collected by a researcher (Beattie, 2014). As the texts include a multitude of such criteria, the linguistic features are often condensed into the so-called Texture Index (Sydserff & Weetman, 1999). The level of decision usefulness can be enhanced by increasing the efficiency of the communication with the reader. Thus, if the efficiency is increased, the decision usefulness can be regarded as high and therefore the quality underlying the report.

Increasing the communication efficiency can be achieved by linking meaningful content, allowing a reader-friendly organization of sentence structures, an organized flow of ideas, a high relevance concerning the content and the assumption of little prior knowledge. However, the concepts of communication efficiency vary by analysis. The operationalization of these criteria always takes place via linguistic text features (Sydserff & Weetman, 1999).

The second manual method is the so-called qualitative content analysis. This approach is based on a specific set of criteria operationalized by a coding guideline. This coding guideline determines how financial disclosure meets certain quantitative and qualitative requirements (e.g., those derived from statutory regulations). The data obtained and patterns observed enable researchers to measure the quality of the report (Botosan, 1997; Bowman, 1984; Guthrie et al., 2004). The evaluation can be done through a variety of methods ranging from qualitative verbal methods to complex quantitative-statistical approaches (Beattie et al., 2004; Miihkinen, 2013). Due to its significance for this publication the qualitative content analysis is discussed in Chapter 3.2.

Despite the availability of an extensive toolset of methodologies and advancing integration of computerized text analysis (Kaya & Seebeck, 2018), gaps in the quality measurement can still be found. Especially when considering the significance of published information and compliance with the requirements of a standard setter. In the context of quality assessment, it is noticeable that no current method captures the informational needs of the addressees as well as the requirement levels for specific report components across all procedures. Therefore, in the following discussion, we take a closer look at how to explicitly incorporate those requirements into quality measurement, and which methodological refinements and extensions are deemed necessary for quality evaluation.

To verify the research gap, we conducted a systematic literature review. This was done by defining<sup>2</sup> key search expressions, and combinations thereof, and systematically searched databases of the platforms “Econbiz,” “Business Source Premier,” “Science Direct,” “Springer Link,” “Social Science Research Network,” and “Wiley Online.” In light of the vast amount of literature on disclosure quality available, we additionally performed an in-depth analysis of survey articles in the accounting narrative area.<sup>3</sup> No relevant literature was found to interlink the concept of requirement levels of standard setters to the information demand of readers. Based on the identified research gap, we developed a method that aims at incorporating the requirement levels for narrative elements into disclosure quality measures.

## 2.2 | Necessity to integrate the requirement level

In this chapter, we will introduce the concept of requirement levels. Requirement levels consider the different usefulness of information when assessing report quality.

As part of external reporting, the companies risk reports are subject to the requirements defined by the standard setters. The standard setters have to consider the different interests and informational needs of the addressees (Wagenhofer & Ewert, 2015). In addition, the risk report is part of the management report. Further, the purpose of risk reports lies in the disclosure of important risk-relevant information in order to make decisions based on information not included in other reports or accounting instruments (Filipiuk, 2008). Thus, the management and risk reports close a gap, since readers such as investors are often not interested in primarily ex post financial data, but rather prefer to be informed with non-financial information concerning the future (Nagy & Obenberger, 1994; Pike et al., 1993).

The reports should deliver decision-relevant information to the readers.<sup>4</sup> In order to assess whether the conveyed information is important, it is necessary to evaluate the recipient’s level of knowledge. Additionally, the impact of the information on future decisions must also be assessed (Filipiuk, 2008). The knowledge of an

individual can be simply regarded as a quantity that is increased by received information. This requires that the individual does not yet has this specific information. Information can therefore only be relevant, if the purpose-oriented knowledge changes and the received knowledge is considered. Besides the determination concerning the relevancy of the given information, another question concerns the degree of relevance. The latter can be evaluated by the gained knowledge through the added information. The greater the knowledge gained through the additional information, the greater its impact on the decision. Hence, the higher the relevance of the respective information.

This evaluation is only possible, if the level of knowledge of the respective reader and the decision situation he faces is known. Consequently, the quality of a risk report is evaluated from the perspective of each individual reader. When specifying which (decision-relevant) information from the companies is to be transmitted to the addressees, the standard setter must consider the different levels of knowledge and their informational needs (Stein, 2011). The standard setter can only formulate meaningful regulations, if the regulations are based upon the level of knowledge of a certain target group. The quality of reporting can therefore be measured based on the adopted accounting rules and consequently must not be measured at an individual level. Quality measurement thus implicitly assumes that the standard setter has an idea regarding the information needed and its relevance for the addressees. Under this premise, the quality of risk reporting can be assessed by evaluating the company’s compliance with the requirements of the standard setters.

The financial reporting standards include the requirements that define which decision-relevant information is published in the risk report.<sup>5</sup> However, to satisfy the informational needs, the degree of decision-relevance must be considered. The concept of a requirement level is introduced in order to integrate the degree of decision-relevance in evaluating the quality of risk reports.

The requirement level is derived from the degree of decision-relevance of information determined by the standard setter. When setting forth the scope of further knowledge to be disclosed in each category of the risk report the standard setter must consider the prior knowledge of the readers. Additionally, the standard setter must also consider how this knowledge will influence decisions. If the standard setter recognizes that extra knowledge is needed in a category, or if such extra knowledge will significantly influence the decision of readers, the financial reporting standards must ensure that the risk reports

<sup>2</sup> The search terms are “text analysis,” “disclosure quality,” “information demand,” “standard setter,” “requirement level,” and “Bloom’s taxonomy.”

<sup>3</sup> For this, the following survey articles were considered: Beattie (2014); Core (2001); Elshandidy et al. (2018); Hassan and Marston (2019); Jones and Shoemaker (1994); Li (2010); Loughran and McDonald (2016); Marston and Shrivs (1991); Ryan (2012).

<sup>4</sup> For the definition of information in this context, see Filipiuk (2008) and Stein (2011).

<sup>5</sup> See the general disclosures under national and international financial reporting standards (Coenenberg et al., 2016). Chapter 3.3 deals in detail with the disclosures required by DRS 20.

contain this kind information. It is therefore not enough to simply require that information on this category must be presented in the risk report. Instead, the financial reporting standards must include appropriate requirement levels in order to assure that purposeful knowledge is disclosed. Hence, the requirement levels defined in the financial reporting standards must be considered when measuring the risk report's quality. Only under consideration of the requirement levels can an evaluation be made whether the purpose-oriented knowledge obtained by a risk report is sufficient enough to raise the knowledge to the desired level of the addressees.

German Accounting Standards (DRS 20) contain detailed requirements regarding the information that must be provided and the respective degree of decision relevance. For example, DRS 20.11 defines and delineates the terms presentation and analysis: "Presentation is defined as 'statement of facts or the description of matters' and 'analysis' refers to the identification of origins and the interdependencies between cause and effect." From the definitions, it becomes clear that less purposeful knowledge is conveyed by the "presentation," whereby more purposeful knowledge is conveyed through "analysis." Thus, the requirement level of analyzing elements in risk reports is higher than the requirement level of presenting elements. Such diverse requirement levels must especially be considered when measuring the quality of a risk report.

The International Financial Reporting Standards (IFRS) also contain requirement levels to define the knowledge that must be conveyed. However, these are less differentiated compared to DRS 20. For example, often only amounts or values are required to be disclosed. Pursuant to IFRS 7.36, an entity "shall disclose [...] the amount that best represents its maximum exposure [...]" which means disclosing knowledge to readers only in the form of a specific value. However, entities are also required to provide explanations in certain categories. Pursuant to IFRS 7.41(a), entities must disclose the method used to measure market risk: "The entity shall also disclose an explanation of the method used [...]" It is therefore not sufficient to mention the methodology, but it must be explained by the company. The explanation is intended to provide the addressee with more purposeful knowledge than is necessary when only values or methods need to be mentioned. If the standard setter requires explanations, it can be assumed that the degree of decision-relevance of the information must be higher. Hence, the IFRS calls for a higher requirement level to measure the quality of a risk report.

The above discussion clearly shows that the degree of decision-relevance must be considered in the form of requirement levels when measuring the quality of a risk report. With the help of the requirement levels, the hierarchization of transferable knowledge can be integrated into

the measurement. In the following chapter, we develop a method that is able to explicitly consider requirement levels in the quality measurement of risk reports. This will be done by means of taxonomy levels.

### 3 | DEVELOPMENT OF AN EVALUATION METHOD BASED ON TAXONOMIC LEVELS

#### 3.1 | Classification system according to Bloom as part of quality measurement

Hereinafter, we will develop a method that integrates the degree of decision-relevance described in Chapter 2.2. In doing so, we apply requirement levels to measure the quality of risk reports. By integrating requirement levels into quality measurement, we can evaluate transparently, if the authors of the risk report explicitly comply with the decision-making relevance required by the standard setter.

The reader's intended information level can only be reached, if the authors of the risk report comply with the various hierarchies of information set forth by the standard setter.

In order to ensure the highest degree of objectivity, we use a procedure based on the so-called taxonomic levels to determine the standard setter's requirement level.

In the field of biology the Oxford dictionary defines taxonomy as "the branch of science concerned with classification, especially of organisms; systematics." In 1956, the American psychology professor, Benjamin Samuel Bloom, published a book entitled, "Taxonomy of Educational Objectives – The Classification of Educational Goals" (Bloom et al., 1971). Ever since, this terminology has additionally been used for describing learning goals in educational science. Thereby, taxonomy should be developed logically and internally consistent. Moreover, taxonomy should go beyond a pure form of classification (Bloom et al., 1971). Basically, taxonomy can encompass random elements. But in Bloom's system, the taxonomy must be structured in a way that allows the developed categories to be ranked or assigned to a hierarchy (Bloom et al., 1971).

Bloom is concerned with the order of learning outcomes and therefore with the question of whether the desired learning outcome is achieved. The learning objectives, for example, could consist of pure memorization (low taxonomy level) or in problem solving (higher taxonomy level) (Bloom et al., 1971). Consequently, the objective of his taxonomy system is to define unambiguous terms that classify the intended results of the educational process. The student behavior is thereby represented by the defined unambiguous terms (Bloom et al., 1971). In the following

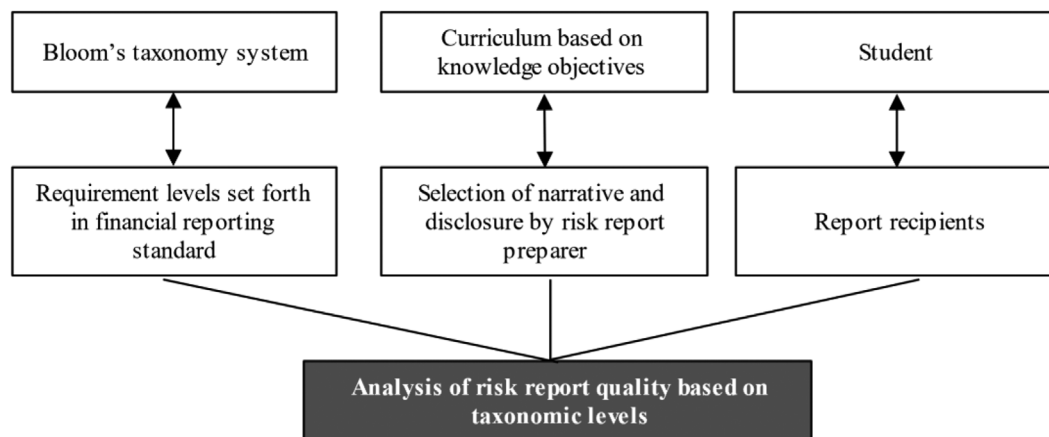


FIGURE 2 Transfer of Bloom's taxonomy system to risk reports

discussion, we want to transfer Bloom's theoretical basis to the assessment of risk reporting quality. This will be done by connecting terms that express a specific level of expertise and the fulfillment of that level to evaluate the quality of a risk report. Figure 2 summarizes the analogies described below.

Based on Bloom's taxonomic levels, learning goals are generally understood as the level of cognitive knowledge that students are expected to achieve. Therefore, in the context of the measurement of risk reporting, standard setters should define precise expectations and requirement levels in the risk reporting standard. The hierarchical requirements (formulated requirement levels in the financial reporting standard) are thereby structured in a way that ensures to express the degree of decision-relevance (formulated in Chapter 2.2) clearly. In the field of cognitive science, the curriculum is thereby contingent on the learning goals. Transferring this to the measurement of the risk report quality: the author of the report must observe the requirements stated by the standard setter (selection of the narrative and disclosures by the preparer). Thus, the resulting implementation of the statements in the risk report should enable the addressees to reach the level of information expected by the standard setter. The expected level of information is thereby derived from the respective requirement level in the form of the specific taxonomy level (Bloom et al., 1971). The research group around Bloom developed six classes of learning, each with an inherent taxonomic level. The logical order of these taxonomic levels range from easy to complex learning objectives, illustrated by the following list (Bloom et al., 1971):

1. Knowledge
2. Comprehension
3. Application
4. Analysis

5. Synthesis
6. Evaluation

These six taxonomy levels support the researcher by clearly defining the respective levels. Based on the respective taxonomy level, it is therefore possible to assign a respective criterion in the accounting standard to the requirement level. It should be noted that the specific taxonomy level must be clearly defined. Further, it must be possible to explicitly link the taxonomy levels to the accounting standards. Additionally, the respective taxonomy level must be clearly defined. The easiest way to assign the taxonomy levels to the requirement levels is if the wording of the requirement level corresponds verbatim to the target level of the respective taxonomy level, or if there are definitions for the respective requirement level in the accounting standard itself.

If only one of these conditions is met, one can refer to commentaries in the literature on financial reporting. Thereby the requirement levels can be determined out of the wording of the standard.

Table 1 illustrates the taxonomic levels and highlights how the requirement level rises with the taxonomic levels.

In summary, in the context of this publication the desired learning outcome of the respective taxonomy level can be understood as the pre-specified goal of the standard setter for risk reporting in companies. This involves checking whether a reader of the risk report can achieve the learning outcome required by the standard setters for the respective taxonomy level. With rising taxonomy levels, the degree of decision relevance increases and therefore the significance for the reader. Also, the requirement for the reporting increases. Integrating taxonomic levels into quality measurement of risk reports creates a theoretical framework that helps to link the requirement level of the respective financial reporting standard to the degree of fulfillment in the respective risk report of the company.

**TABLE 1** Description of Bloom's taxonomic levels (Bloom et al., 1971)

Taxonomic level	Level of expertise	Explanation of taxonomic level
1	Knowledge	Knowledge of facts, such as data, events, and classifications.
2	Comprehension	The ability to interpret, that is, present a situation at any level of generalization and extrapolate this.
3	Application	Apply abstractions (e.g., theories) to concrete situations.
4	Analysis	Analyze organizing principles that determine systematic constituents and structures; discover relationships.
5	Synthesis	Derive the consequence of abstract relationships able to reorganize a situation or combine them.
6	Evaluation	Quantitative and qualitative judgments of the extent to which materials and methods meet specific criteria. Connect narratives.

### 3.2 | Integration of the taxonomy levels into qualitative content analysis

Chapter 2.2 in conjunction with Chapter 3.1 demonstrated the reasonableness to take the requirement levels into account when it comes to the quality measurement of risk reporting. As mentioned in Chapter 2.1, several different methodologies exist in order to measure the quality of the risk report. In the course of the following explanations, the taxonomy levels will be integrated into qualitative content analysis.<sup>6</sup>

Since the latter also involves coding and evaluating texts based on a predefined system of categories, this approach seems to enable a very good link between both concepts (Linsley & Shrivs, 2006).

The central differentiation criterion to other methods of content analysis lies in the origin and the definition of the category system. On the one hand, the category system is derived from the theory of qualitative content analysis (deductive), and on the other hand, by the comparison of the text that must be analyzed (deductive-inductive) and enhanced (Mayring & Fenzl, 2014).

The structured content analysis is suitable for measuring the quality of risk reports.

Each and every risk report is checked whether it meets the qualitative requirements. The requirements, which are grouped into categories, are derived from the accounting standards. For every category, specific quality gradations have been established. The individual gradations of quality are then further specified with a definition and a core example. Thereby, coding the text material by different persons leads to a largely uniform result (Flick, 2016).<sup>7</sup>

A risk report represents a specific document type with comparable characteristics and results for all items analyzed. Similarly, the questions and categories can be applied identically to all risk reports to obtain a comparative evaluation via a structured analysis (Flick, 2016). The quality measurement is done by a structured analysis on an ordinal scale. Coding the specific references in the risk reports results in characteristics with a natural order and an internal logic of sequence, such as no sentence, one sentence, two sentences, and three or more sentences in one category (Mayring, 2015).<sup>8</sup> It must be noted that the additional sentences create an added value for the addressee (Beattie & Thomson, 2007; Beretta & Bozzolan, 2004). Therefore, the decision relevant information is enhanced (see Chapter 2.2). Figure 3 depicts the process sequence of the analysis. Each step is detailed below.

In step 1, we will determine the relevant analysis units for our investigation. Based on the defined data material, the analysis unit consists of the evaluation unit, the context unit, and the coding unit (Mayring & Fenzl, 2014). The objective of this initial process step will be the precise characterization of the starting material. Thereby, the evaluation unit encompasses a strictly delimited amount of risk reports being assessed. The delimitation can be expressed, for example, by an stock market index. One possible evaluation unit would be the companies in the DAX index. The risk report itself can be considered as the context unit. Thus, the risk report represents the largest possible single item for content analysis. Lastly, the coding unit must be defined. The coding unit must be determined as the smallest possible analytical framework. In this study, the analysis is carried out textually at the word level, or in the case of illustrations, at the level of the individual representation.

Step 2 concerns the definition of the expressions. By defining the expressions, we determine the extent and the possible range of the criteria to be analyzed. Two different evaluation scales are used to analyze the risk reports. The so-called binary criteria assume that certain evaluations have a yes/no character and therefore only a statement can be made about their presence or absence. The

<sup>6</sup> Chapter 5 integrates taxonomic levels into other methods of text analysis.

<sup>7</sup> This approach is also referred to as intercoder reliability.

<sup>8</sup> The coding scheme is addressed in detail later.

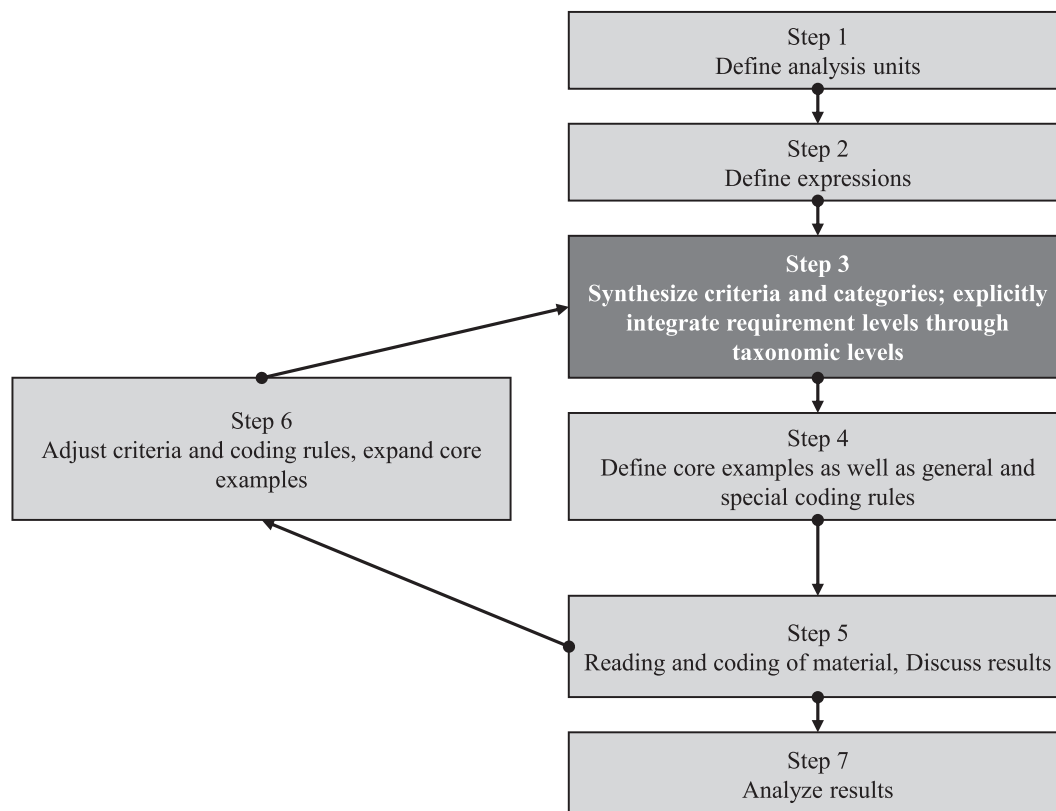


FIGURE 3 Sequence of a selected analysis method (adapted from Mayring, 2015)

TABLE 2 Example of general rules for coding expressions

Criterion type	Case	Code	
Binary	Criterion is not mentioned in risk report.	0	
	Criterion is mentioned in risk report.	1	
Quantifiable	Criterion is not mentioned in risk report.	0	
	Criterion is mentioned in risk report:	Extent: One sentence	1
		Extent: Two sentences	2
		Extent: More than two sentences	3

ordinal nature of the criteria is interpreted in terms of existence being always better than absence of a criterion.

The so-called quantifiable criteria are based on the assumption that more complex situations can be comprehensively presented and thus contain a greater amount of information for the reader. The idea behind this type of criteria is that the extent to which the respective facts are dealt with should be explicitly included in the evaluation (Hoitash & Hoitash, 2018). A requirement for this is that additional information is contained in the respective sentence. If this condition is met, the additional sentences are interpreted as a higher quality fulfillment of the required reporting standard. However, it would not be recorded as an increase in quality, if, for example, the same information is given in three successive sentences. By using a point

system for the quantifiable criteria, the fulfillment of the respective standard is checked operationally. The evaluators of the risk reports have to determine whether there should be a maximum number of points, since the use of additional information decreases with increasing scope. As an example, such a maximum limit was defined in Table 2.

Table 2 depicts that no more points are awarded, if the statement extents more than two sentences. The decision as to whether a certain criterion is measured in binary or quantitative terms results logically from the nature of the criterion or from the description of the claim within the accounting standard. The specified coding rules result from the following evaluation logic:

In the third step, the methodology developed in Chapter 3.1 was considered. Therefore, the established requirement



levels are explicitly taken into account as part of risk reporting. When coding, attention is paid to the taxonomy level at which the standard setter has classified the respective criterion. If the standard is assigned to taxonomy level 1 and the requirement level is therefore very low, the non-fulfillment of this criterion is regarded as less significant than a criterion that is assigned to taxonomy level 6 and therefore has to meet the highest level of requirement.

In step 4, when defining core examples and general and special coding rules, a systematic structure for carrying out the actual evaluation of the analysis objects must be developed. This systematic structure should be made available to the coders in the form of a guide. The coding rules and core examples should clearly illustrate the characteristics of the examination. Thereby, for example, the checked criteria, criteria types, and analysis units should be made accessible to the coding persons. The objective is to increase intercoder reliability, that is, the greatest possible agreement between the coding results of different encoders.

After establishing the coding rules in the fifth step, several analysts evaluate all risk reports, followed by a discussion of any persisting deviances. Based on that, the evaluation reports are prepared and systematically expanded and specified by applying them to the analysis units. In an iterative process, steps 3–6 of Figure 3 should be repeated until the coders cannot detect any significant deviations in the evaluation. This procedure is intended to increase the reliability, objectivity, and validity of the investigation.

In the implementation of the coding, the principles of Krippendorf (Krippendorf, 2004a, 2004b) should be observed. Krippendorf's principles specifically demand a high reliability of the investigation. In particular, this means that the coders must work under the same conditions, with the same coding instructions and independently of one another. By doing so, it is ensured that the process of data generation is reproducible and independent of environmental conditions. The generated data should be compared with each other and in case of discrepancies checked again. Subsequently, the scientists should agree on a coding variant according to the consensus principle. It should be noted that the analyzed documents are always created by a specific person or a group with a specific objective. Therefore, a universal correctness cannot be achieved (Flick, 2016).

In the seventh and last step, after the rule-based coding of the risk reports, the results can be analyzed or further processed. If the qualitative data is further processed in the form of coded text elements, they can, for example, be quantitatively condensed using a scoring model. In the following section, the categories and criteria are exemplified based on the DRS 20.

### 3.3 | Applying the research methodology developed to the DRS 20

In this chapter, the quality of a risk report will be assessed based on the presented methodology. In doing so, the requirement levels are being explicitly considered. The basis of evaluation will be the DRS 20. According to the previous considerations, fulfillment of the DRS 20 requirements determines whether the report is of high quality or not. Therefore, as explained in Chapter 2.2, the standard setter assumes a type of reader with a specific level of knowledge. The German DRS 20 was thoroughly reformed at the end of 2012 and last revised in 2019. The DRS 20 regulates the management reporting for all parent entities that have to prepare a group management report or do it voluntarily. Furthermore, the DRS 20 had to be applied initially at the financial year beginning after December 31, 2012. A risk report is a part of the management report, whereby risks are defined as follows: "Potential future developments or events that could lead to a negative deviation from the group's forecasts or objectives." (DRS 20.11). Additionally, the forecasting period should be at least 1 year (DRS 20.127; DRS 20.156). In sections 135–164, the DRS 20 sets forth precise requirements for risk reports. In order to assess compliance with the requirements of the DRS 20 for risk reporting at company level and thus to be able to measure the quality of a risk report, it is necessary to develop a structured criteria system consisting of a qualitative content analysis in connection with the taxonomy levels. The compliance on a company level in that case is specified in the sub-paragraphs. The individual requirements are generally structured in line with the categories stated in the aforementioned sections of the DRS 20. At the top level, a distinction is made between the risk management system and the individual risks. In the following, this distinction will be called main category. The sections in DRS 20 that fall under these main categories are then combined by themes as subcategories, each of which also includes common subthemes. As illustrated in Table 3, the respective subtheme is assigned to the relevant section of DRS 20. Furthermore, the criteria to be checked are split into binary and quantifiable parameters. A risk report may or may not contain binary criteria. The goal is to check whether the analyzed risk report fulfills the (formal) requirements in compliance to DRS 20.

The quantifiable criteria are verified by a binary perspective, and in addition, more deeply by differentiating statements in the respective risk report. This procedure was explained in Chapter 3.2.

As illustrated in Chapter 2.2, DRS 20 encompasses diverse requirement levels. In the following, we will link the requirement levels to the taxonomic levels that were

TABLE 3 Categories and criteria pursuant to DRS 20

Type of criteria	Description	DRS 20	Requirement level
Main category 1: Risk management system			
Subcategory 1.1: Structure and framework (SF)			
Binary	Generally recognized framework	Sec. P139	Disclose
Quantifiable	Structure of risk management system	Sec. P137	Address
Subcategory 1.2: Goals and strategies (GS)			
Quantifiable	Risk management goals	Sec. P137	Address
Quantifiable	Risk management strategies	Sec. P137	Address
Binary	Risks basically not recorded	Sec. P140	Present
Binary	Risks basically avoided	Sec. P140	Present
Subcategory 1.3: Risk management process (RMP)			
Quantifiable	Identification step	Sec. P144	Discuss
Quantifiable	Quantification step	Sec. P144	Discuss
Quantifiable	Management and control step	Sec. P144	Discuss
Quantifiable	Internal monitoring of risk management processes	Sec. P144	Discuss
Main category 2: Single risks and overall risk exposure			
Subcategory 2.1: Single risks (SR)			
Quantifiable	Evaluation of exposure to risk	Sec. 149	Assess
Quantifiable	Quantifying exposure to risk	Sec. 150, Sec. 152	Present
Quantifiable	Consequences and impacts	Sec. 149, Sec. 157	Analyze
Subcategory 2.2: Risk exposure and overall risk (RE)			
Quantifiable	Changes in single risks	Sec. 159	Discuss
Binary	Risks endangering the status quo	Sec. 148	Designate
Quantifiable	Overall risk exposure	Sec. 160	Summarize
Subcategory 2.3: Clarity and transparency (CT)			
Binary	Explicit pre- and post-mitigation risk exposure	Sec. 157	Present
Binary	Ranking and categorization of risks	Sec. 162	Combine

developed in Chapter 3.1. According to each section, entities must “present, disclose, analyze, discuss, assess, designate, summarize, combine, and address” specific situations concerning the risk management system and single risks. The above-mentioned terminologies will now subsequently be transferred to the hierarchical classification system.

In the analysis, we examine whether the entity’s risk report fulfills the DRS 20 requirement level expressed in the respective taxonomic level. The reader should reach the learning outcome for each taxonomic level set forth in the DRS 20. The higher the taxonomic level, the higher the reporting requirement, and accordingly higher is its significance for the reader. The classification into taxonomic levels is done by using the operators that are identified within the DRS 20 sections and listed in Table 3.

For the purposes of classification, these operators are listed again in alphabetical order: “Address, analyze, assess, combine, disclose, designate, discuss, present and summarize.” This classification facilitates compliance with the standard setter’s intention, in this case for DRS

20. The problem is that not all the DRS 20 requirements are defined within the standard by the significance of their requirement. Hence, in such cases, reference to commentaries in the literature on the management report for defining the significance of the respective requirement is necessary. In cases where a term in the report was not found, we looked it up in Duden (German reference dictionary). For all definitions contained in DRS 20 an unambiguous classification was possible. However, concerning the commentary literature the classification is partly ambiguous. For example, in this paper, we interpret the requirements: “Address” and “Present” as equivalent, based on the commentary in the literature that, “In view of the choice of words in the law (‘address’), a verbal presentation of risks ... satisfies the requirements” (MünchKomm-HGB/Lange, 2013, §289 marg. 104).<sup>9</sup> This

<sup>9</sup> Remarks on other terms, for instance, are also in Beck’s commentary on a balance sheet (Beck’scher Bilanz-Kommentar/Grottel (2016), §289 margin no. 20).

citation states that the words “Address” and “Present” are to be classified at the same level.

The concrete assignment of the requirements mentioned in the DRS 20 to one of the taxonomy levels according to Bloom will be demonstrated by two examples.<sup>10</sup> As evidently depicted in Table 3, DRS 20.P140 requires “Presentation” of whether any risks are basically not recorded or avoided. “Presentation” is the operator for the desired entity-specific explanation of the situation in the entity’s risk report. Pursuant to DRS 20.11,<sup>11</sup> knowledge is presented as “Statement of facts or description of matters.” This is consistent with description of the subcategory of the taxonomic level, knowledge, where “knowledge of dates, events, persons, places, sources of information, etc.” is to be conveyed (Bloom et al., 1971). If this logic of the hierarchy between the taxonomy levels is followed, it is the lowest reporting requirement, since it is only a matter of imparting knowledge.

In contrast, DRS 20.K144 calls for higher-level reporting. This section states that a risk report should “discuss” the “identification, assessment, management and control of risks, as well as the internal monitoring of these workflows.” Pursuant to DRS 20.11, the operator “discuss” encompasses “a detailed explanation, commentary and interpretation of a matter that goes beyond a straightforward presentation.”

According to the descriptions in DRS 20 paragraph 11, the operator “discussion” includes the “further explanation, commentary and interpretation of a situation beyond the pure presentation” (DRS 20 paragraph 11). Especially, a DRS standard should serve with an explanatory character for the “Information on requirements, causes or consequences of facts or measures” (DRS 20, item 11). Here too, one can apply the definition of the operator in DRS 20, for comparison with Bloom’s taxonomy system. Grouping is done in level 5, called “summarize,” described as “putting together of elements and parts so as to form a whole” (Bloom et al., 1971). The anticipated outcome of a “discussion” equates to “production of a plan, or proposed set of operations” (Bloom et al., 1971), or “derivation of a set of abstract relations” (Bloom et al., 1971). The explanation is intended to enable the addressees of the risk report to assess a situation regarding the causes of its occurrence and the possible consequences of its occurrence.

Analogous to the procedure described above, the stated requirements are assigned to matching taxonomic levels.

<sup>10</sup> The classification of learning objective operators in taxonomic levels was not taken from practical guides by diverse education services. A comparison of various leaflets and guides to prepare curricula highlighted significant discrepancies in classification. An independent grouping logic is followed to achieve a compelling classification.

<sup>11</sup> Key definitions of terms in DRS 20.11 are presented for an understanding of the set of rules in the DRS standard.

The assignment is based on the categories developed in Table 1. Table 4 lists the outcome. In order to illustrate on how to match an operator to a taxonomic level, a definition and explanation are given at the level of each classification system.

To check the DRS 20-compliant documentation of the individual risks, an additional differentiation of subcategory 2.1 from Table 3 is necessary. As a result, the criteria mentioned in subcategory 2.1 are checked and assessed separately for each individual risk of the company. Thus, for each individual risk mentioned, the risk report analyses how detailed and differentiated the evaluation, quantification, consequences, and impact of the respective individual risk are dealt with.<sup>12</sup>

In the process of creating categories and criteria, certain sections of DRS 20 were found unsuitable for additional empirical analysis. Accordingly, the certain sections were excluded for three reasons: firstly, some sections were not considered if they concerned only definitions or principles. Secondly, if the certain sections were more of an explanatory nature, or thirdly, if the requirements could not be verified. Requirements cannot always be checked for compliance by companies if internal company information that the external addressee does not have is necessary for their assessment. Table 5 below summarizes the sections excluded for the above reasons.

### 3.4 | Limitations of the developed approach

Our approach is based on the premise that in formulating the regulations, standard setters have correctly judged the level of knowledge of readers and their informational needs. When this premise is applicable, a meaningful measure is used to evaluate the quality of the risk report.

However, our method would be unsuitable for assessing the quality of risk reports, if the standard setters incorrectly evaluate the level of knowledge of the readers.

In such a case, our method would determine the reporting quality as high, although the information contained would only be of little decision-making relevance for the addressee.

Consequently, there would be a discrepancy between the measured quality of the risk report and its perceived quality by the readers.

Against this background, our method presented could also serve as an impetus for standard setters on reflecting upon the decision-relevant information to disclose in a risk

<sup>12</sup> This was again categorized by type of risks, to ensure systematic naming and testing of single risks. Upon request, the authors will supply the categorization.

TABLE 4 Relationship between the DRS 20 and taxonomic levels<sup>13</sup>

Requirement per DRS standard	Definition per DRS 20.11, commentary in literature, or Duden	Matching taxonomic level	Explanation of taxonomic level <sup>14</sup>
Analyze	“List causes and relationships” (DRS 20.11)	Level 4: Analysis	Analyze organizing principles that determine systematic constituents and structures; discover relationships.
Disclose	“Statement of facts or description of matters” (DRS 20.11)	Level 1: Knowledge	Knowledge of facts such as data, events, and classifications.
Assess	“Evaluation of and commentary on matters” (DRS 20.11)	Level 6: Evaluation	Quantitative and qualitative judgments of the extent to which materials and methods meet specific criteria. Connect narratives.
Designate	“Provide Information, inform” (Duden)	Level 1: Knowledge	Knowledge of facts such as data, events, and classifications.
Present	“Statement of facts or description of matters”(DRS 20.11)	Level 1: Knowledge	Knowledge of facts such as data, events, and classifications.
Address	“... disclose risks verbally ...” (MünchKomm-HGB/Lange, 2013, §289 marg. # 104)	Level 1: Knowledge	Knowledge of facts such as data, events, and classifications.
Discussion	“A detailed explanation, commentary and interpretation of a matter that goes beyond a straightforward presentation. It provides information about the conditions, causes, or consequences of matters or measures” (DRS 20.11)	Level 5: Synthesis	Derive the consequences of abstract interrelationships able to reorganize a situation or combine them.
Combine	“Merge individual facts into categories” (Duden)	Level 4: Analysis	Analyze organizing principles that determine systematic constituents and structures; discover relationships.
Summarize	“State or express in a concise form”(Duden)	Level 4: Analysis	Analyze organizing principles that determine systematic constituents and structures; discover relationships.

TABLE 5 Sections in the DRS standard excluded from further analysis

Reason for exclusion	Section in DRS 20
Definitions or principles	135, 146
More of an explanatory nature	136, P138, P141, P143, P145, 147, 153, 154, 158, 161, 163, 164
Requirements cannot be verified	142, 151, 155, 156

report. Therefore, standard setters must evaluate whether they and the readers would come to the same conclusion on classifying information deemed as decision-relevant. In addition, on the one hand the standard setters must correctly assess the degree of decision-making relevance of individual information. On the other hand, the relevant decision-making information must be integrated into the regulations in the form of adequate levels of requirements.

<sup>13</sup> Since DRS 20 is a German financial reporting standard, the linguistic terms were purposely looked up in *Duden*, a German dictionary.

<sup>14</sup> For explanations of taxonomic levels, see Table 1.

If the standard setters correctly consider the level of knowledge of the addressees when formulating the regulations, the corresponding requirement levels must be correctly taken from the standards and assigned to the respective taxonomy levels. In applying our method, it is important to note that the standard setters have defined only some of the key terms.<sup>15</sup> Especially when adopting the meanings of terms from commentaries in the literature and applying these to taxonomic levels, it is necessary to ensure that these coincide with the intentions of standard setters. Otherwise, the decision-relevance of the information in a risk report will be falsely evaluated. Therefore, the assessed quality of the report will be incorrect. It would certainly be helpful, if standard setters would more clearly elucidate their expectations concerning the information to be disclosed, and define the key terms. The reporting entities would then be in a better position to determine which information best meets the decision-relevance requirement of standard setters.

<sup>15</sup> See Chapter 4.1 and Table 4.

## 4 | AREAS OF APPLICATION WITHIN THE SCOPE OF QUALITATIVE ANALYSIS OF RISK REPORTS

### 4.1 | Integrating a taxonomy system into a multi-criteria evaluation procedure

The coding scores from Chapter 3 evaluate the achievement of the objectives for each individual target criterion. However, it is not possible to make consolidated statements of the quality of a risk report (Berens, 1992). The method known as utility value analysis or scoring model has established itself as the method for aggregating information from individual criteria into an overall result.<sup>16</sup> If the quality of the risk report is to be based on a holistic assessment, the measurement methodology developed in the previous chapter can be incorporated into a scoring model (Beattie et al., 2004). Consolidating the evaluation of single criteria into a total score allows ranking of the risk report quality of the analyzed companies (Adam, 2000).

We apply the scoring model to sum the weighted degrees of target achievement of single criteria<sup>17</sup> to come up with a one-dimensional result for the defined overall score (Zangemeister, 1976) as follows:

$$SC_i = \sum_{j=1}^n w_j \cdot ZG_{ji}$$

Where:

- $SC_i$  = Score of risk report  $i$
- $w_j$  = Weighting factor for single criterion  $j$
- $ZG_{ji}$  = Degree of target achievement of single criterion  $j$  for risk report  $i$

The degree of target achievement of single criteria and weighting factors are explained below.

Based on the categories and criteria developed in Table 3, the risk report evaluation gives points to respective single criteria in order to reflect the achievement level. We use two different evaluation scales, one each for binary and quantifiable test criteria. The results are converted to one common scale for comparison. For the analysis, we set the minimum and maximum values at 0% and 100%, respectively, for both scales. The degree of target achievement  $ZG_{ji}$  can be computed with the following formula

<sup>16</sup> For a similar opinion, see Ntim et al. (2013). The dissemination of utility analysis in Germany is based especially on the work of Zangemeister (1976).

<sup>17</sup> In addition to additive approaches, see examples of other multiplicative consolidation principles in Adam (2000).

TABLE 6 Relationships between the taxonomic level and weighting factor

Taxonomic level	Level of expertise	Weighting factor
1	Knowledge	1x
2	Comprehension	2x
3	Application	3x
4	Analysis	4x
5	Synthesis	5x
6	Evaluation	6x

(Hölscher & Schneider, 2014):

$$ZG_{ji} = \frac{K_{ji} - K_{j,\min}}{K_{j,\max} - K_{j,\min}} \cdot (ZE_{\max} - ZE_{\min})$$

Where:

- $K_{ji}$  = Result of risk report  $i$  for criterion  $j$
- $K_{j,\min}$  = Minimal value of criterion  $j$
- $K_{j,\max}$  = Maximum value of criterion  $j$
- $ZE_{\min}$  = Minimum degree of target achievement
- $ZE_{\max}$  = Maximum degree of target achievement

For the two evaluation scales, the following formulas are used to determine the degree of target achievement:

$$\text{Binary criteria : } ZG_{ji} = \frac{K_{ji} - 0}{1 - 0} \cdot (100\% - 0\%) = K_{ji} \cdot 100\%$$

$$\begin{aligned} \text{Quantifiable criteria : } ZG_{ji} &= \frac{K_{ji} - 0}{3 - 0} \cdot (100\% - 0\%) \\ &= \frac{K_{ji}}{3} \cdot 100\% \end{aligned}$$

Hence, for the scoring model, we need a procedure to logically determine the weighting factors. The taxonomic system introduced in Chapter 3 is ideal for this purpose, since it is both logical and consistent.

Accordingly, the higher the taxonomic level, the higher the reporting requirements and therefore significance for the reader. Thereby resulting in a higher weighting factor with which the respective degree of target achievement (level of expertise) is included in the scoring model. Table 6 illustrates the relationship between the taxonomic level, level of expertise, and weighting factor:

Based on the target hierarchy developed in Chapter 3.3, Figure 4 below schematically summarizes how the scores are computed in this model.<sup>18</sup>

The developed evaluation method is suitable for measuring the quality of risk reporting by companies, explicitly

<sup>18</sup> See Table 3 in this paper.

Category	DRS Level				Weighting Level	
	ID	Description	Relevant DRS20 Section	Type of Criteria	Taxonomic Level	Weighting Factor $w_{ID}$
SF	1	Generally recognized framework	K139	B	1	1.92%
	2	Structure of risk management system	K137	Q	1	1.92%
GS	3	Risk management goals	K137	Q	1	1.92%
	4	Risk management strategies	K137	Q	1	1.92%
	5	Risks basically not recorded	K140	B	1	1.92%
	6	Risks basically avoided	K140	B	1	1.92%
RMP	7	Identification step	K144	Q	5	9.62%
	8	Quantification step	K144	Q	5	9.62%
	9	Management and control step	K144	Q	5	9.62%
	10	Internal monitoring of risk management processes	K144	Q	5	9.62%
SR	11	Evaluation of exposure to risk	K149	Q	6	11.54%
	12	Quantifying exposure to risk	K150/152	Q	1	1.92%
	13	Consequences and impacts	K149/157	Q	4	7.69%
RE	14	Changes in single risks	K159	Q	5	9.62%
	15	Risks endangering the status quo	K148	B	1	1.92%
	16	Overall risk exposure	K160	Q	4	7.69%
CT	17	Explicit pre- and post-mitigation risk exposure	K157	B	1	1.92%
	18	Ranking and categorization of risks	K162	Q	4	7.69%
					$\Sigma=52$	

Risks related to	Weighting factor
Finance	20%
Environment	20%
Production	20%
Organization	20%
Market	20%

$$\rightarrow \sum_{ID=1}^{18} w_{ID} \cdot ZG_{ID,i} = Score$$

FIGURE 4 Analysis approach to compute scores

taking the requirement level into account. It is particularly relevant for the evaluation of capital market-oriented companies such as DAX companies.

Finally, the potential limits of the scoring procedure used in this article will be discussed. As with any application, it is important to critically assess the subjectively selected input variables and weighting factors (Adam, 2000; Hölscher & Schneider, 2014). To ensure maximum objectivity and logically compelling weighting of criteria, the system of taxonomic levels utilizes an independent scheme to evaluate the weighting of single criteria based on the DRS 20. If the relationship arising from the significance of the requirement does not follow the assigned taxonomic level and its associated weighting factor, the resulting risk report ranking will vary. By defining the weighting factors based on taxonomic levels, the significance increases linearly from one level to the next. Alternatively, it is also possible to base the scores on another scale of rising significance. Since there are no obvious reasons for such an overweighting of the higher taxonomy levels in the standards, a linear increase in sig-

nificance was used in this paper. In addition, the general points of criticism of scoring apply to the scoring procedure used. In particular, very low degrees of target achievement could be overcompensated by high degrees of target achievement, while a dimensionless score is consolidated.

Furthermore, it should be noted that our method does not compute a higher quality report, if the requirements are overfulfilled. For instance, if an entity explains a situation in a subcategory that falls under the taxonomic level "Describe" the score weighting is one. Since we only count sentences that contain new information, this in-depth analysis of the situation does not lead to a higher quality valuation.

## 4.2 | Integration into other rule-based methods for evaluating text quality

In the above discussion, we detailed the suitability of this approach for qualitative content analysis. In addition, it is also feasible to integrate this method into other methods

of text analysis. The intent of the further explanations is to point out the basic application options as a starting point for further research projects.

Readability evaluates the quality of risk reports regarding the ease of reading verbal material. The more understandable the text, the higher its quality. Typical evaluation units are sentences (e.g., a parameter for “number of words per sentence”) and words (e.g., a parameter for “number of characters per word”). The classic readability measures do not differentiate in terms of the context of an evaluation unit. Instead, they usually compute a simple average for listed key indicators, subject to equal weighting of individual components of indicators. The meaning of a word or sentence is not considered for the information needed by a reader (Das, 2013). Our developed method of requirement levels provides the foundation in order to refine and expand indicators of readability.

In the sense of a weighted mean value, the significance of an evaluation unit could be taken into account when determining the components of the readability measure. This extension implies that text passages with a higher requirement level have a greater influence on the overall readability of a risk report. Using an example, we can show how to integrate Bloom’s taxonomy system into risk reporting evaluation using a readability scale: instead of computing individual components like the average length of a sentence or word, we can do this separately using the categories and criteria listed in Table 3. This initial step makes it possible to specifically consider the readability of DRS 20 categories. Analogous to the scoring procedure of Chapter 4.1, the categories can then be consolidated with the weighting factors in Table 4, to derive an overall readability score. This readability computation accounts for differences in the significance of single elements of the report in terms of information needs of readers, and thereby delivers a differentiated evaluation of the readability of a risk report.

When evaluating the quality of a report using lexicographical methods, it is also conceivable to integrate requirement levels using taxonomic levels. This type of quality evaluation looks at whether the sentiment and intonation conveyed by the text passages correspond with quantitative facts. In a dictionary, we define key terms for each characteristic sentiment. The more frequently predefined key terms in this dictionary appear in a specific passage, the stronger is the characteristic sentiment rating for that text. A possible starting point of the taxonomy system is the measurement of the number of keywords. In contrast to the standard methodology of simply counting the identified characteristic words, the taxonomy system could be used to weight the key terms over the requirement level. Depending on the context in which the word in the dictionary is used, the significance of the sentiment conveyed by

the entire risk report is ranked either higher or lower. For example, if the key term is in a purely descriptive passage (low taxonomic level), it will be of lower significance to the overall sentiment conveyed, than a term found within an evaluation part of the report (higher taxonomic level). This approach implies that in formulating its report, an entity pays more attention to elements in passages of a higher requirement level. Hence, one can conclude that the indirect sentiment conveyed by these elements of the report more strongly reflects the overall sentiment of the report, and thereby that of the entity.

The taxonomy system can also be used for quality assessment purposes via Document Similarity. In its core concept, this methodology is based on checking whether an external company report is based on boilerplate standard elements without specific reference to the company (low quality) or on report texts individually adapted to the current situation on an annual basis (high quality). Similar to the previous methods, it is possible to integrate the requirement level of a reporting element by weighting it for the overall result. For example, if an identical or similar text is regularly used for a purely knowledge-transmitting text passage, this approach is considered to be less important for the overall quality of the risk report than for an evaluative reporting element.

This approach for analyzing risk reports means that, especially in the DRS 20 categories with a low requirement level, the disclosure often concerns purely technical information. If the disclosure and description do not change from one reporting period to another, it does not necessarily reflect a low-quality report. On the contrary, one can certainly argue that standardization and comparability over time reflects a high-quality report. Instead, for high taxonomic levels, the reader should be empowered to self-evaluate and analyze the situation: this calls for adapting the necessary content in the risk report to disclose decision-relevant information reflecting the current situation of the entity. In this sense, reused elements of the report are viewed as a much more severe reduction of the report’s quality. One approach to integrating taxonomic levels is to either completely disregard elements of the report with low taxonomic levels for the evaluation, or assign much lower weighting factors to such passages in comparison with high taxonomic level categories in the report.

Further possible linkages can be found in the area of manual quality measurement especially by means of linguistic analyses. Simply put, the quality analysis is based on the evaluation of various linguistic features regarding their communication efficiency. In a nutshell: an integration of our approach into linguistic analyses can be done by integrating the taxonomy methodology into the so-called texture index. The texture index could thereby be expanded

by a weighting logic according to the taxonomy levels. Through the usage of this procedure it is implied that communication efficiency in reporting elements with a high requirement level highly affects the reporting quality. In addition, possible usage for the developed taxonomy system can also be identified at the level of linguistic characteristics. For example, parameters could be developed that explicitly check the extent to which the risk report examined is capable of conveying the required level of ambition of the taxonomy level.

## 5 | CONCLUSIONS

The focus of this paper was on expanding analysis of financial disclosure by integrating requirement levels. The goal was to advance the state-of-the-art for measuring the quality of reports. Hence, we developed a universally applicable evaluation methodology, based on Bloom's taxonomy system. By integrating the concept of a requirement level, we made it possible to explicitly consider the relevancy of information. We followed this up by integrating the requirement levels into qualitative content analysis, which helped establish the suitability and added benefits of the information generated by the methodology we developed. Showing the usability for risk reports highlighted the categories and criteria for the set of rules in the DRS 20. This also made it evident that it is possible to meaningfully link the requirement levels of the standard setter, which are inherent in the DRS 20, to Bloom's taxonomy system. By summarizing the individual information into an overall result using a scoring system, it was subsequently possible to show that the methodology developed can be used to derive weighting factors. However, the application of requirement levels demands some premises, which must always be considered when interpreting the outcomes.

By applying a scoring approach to consolidate single pieces of information into an overall outcome, we showed that our developed methodology could be used to derive weighting factors. We also demonstrated the universal nature of our methodology through basic applications to other text analysis approaches.

In conclusion, it is evident that through the integration of a taxonomy system our methodology boosts the validity of established text analysis methods, subject to fulfilling the premises discussed critically in this paper.

## CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there is no conflict of interest.

## ORCID

Nils Helms  <https://orcid.org/0000-0002-4240-6872>

## REFERENCES

- Adam, D. (2000). *Investitionscontrolling*. München/Wien: Oldenbourg.
- Barron, O. E., Kile, C. O., & O'Keefe, T. B. (1999). MD&A quality as measured by the SEC and analysts' earnings forecasts. *Contemporary Accounting Research*, 16(1), 75–109.
- Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research - Issues, theory, method, methods and a research framework. *The British Accounting Review*, 46(2), 111–134.
- Beattie, V., McInnes, B., & Fearnley, S. (2004). A method for analysing and evaluating narratives in annual reports - A comprehensive descriptive profile and methodologies for disclosure quality attributes. *Accounting Forum*, 28(3), 205–236.
- Beattie, V., & Thomson, S. J. (2007). Lifting the lid on the use of narrative analysis to investigate intellectual capital disclosures. *Accounting Forum*, 31(2), 129–163.
- Berens, W. (1992). *Beurteilung von heuristiken – neuorientierung und vertiefung am beispiel logistischer probleme*. Wiesbaden: Gabler.
- Beretta, S., & Bozzolan, S. (2004). A framework for the analysis of firm risk communication. *The International Journal of Accounting*, 39(3), 265–288.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1971). *Taxonomy of educational objectives – The classification of educational goals: Handbook I: Cognitive domain*. New York: Longmans.
- Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, 72(3), 323–349.
- Bowman, E. H. (1984). Narrative analysis of annual reports for corporate strategy and risk. *Interfaces*, 14(1), 61–71.
- Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H., & Steele, L. B. (2014). The information narrative of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, 19(1), 396–455.
- Coenenberg, A. G., Haller, A., & Schultze, W. (2016). *Jahresabschluss und Jahresabschlussanalyse: Betriebswirtschaftliche, handelsrechtliche, steuerrechtliche und internationale Grundlagen – HGB, IAS/IFRS, US-GAAP, DRS*. Stuttgart: Schaeffer-Poeschel.
- Core, J. E. (2001). A review of the empirical disclosure literature - Discussion. *Journal of Accounting and Economics*, 31(1–3), 441–456.
- Das, S. R. (2013). Text and context - Language analytics in finance. *Foundations and Trends in Finance*, 8(3), 145–261.
- DRS 20: Deutscher Rechnungslegungs Standard Nr. 20 (DRS 20) Konzernlagebericht, verabschiedet vom Deutschen Rechnungslegungs Standards Committee e.V. am 02. November 2012, Verabschiedung der geänderten Fassung am 24. Oktober 2019, Deutsches Rechnungslegungs Standards Committee e.V. (2019).
- Elshandidy, T., Fraser, I., & Hussainey, K. (2013). Aggregated, voluntary, and mandatory risk disclosure incentives - Evidence from UK FTSE all-share companies. *International Review of Financial Analysis*, 30, 320–333.
- Elshandidy, T., Shrivs, P. J., Bamber, M., & Abraham, S. (2018). Risk reporting - A review of the literature and implications for future research. *Journal of Accounting Literature*, 40, 54–82.
- Filipiuk, B. (2008). *Transparenz in der Risikoberichterstattung – Anforderungen und Umsetzung in der Unternehmenspraxis*. Wiesbaden: Gabler.
- Flick, U. (2016). *Qualitative Sozialforschung – Eine Einführung*. Reinbek: Rowohlt.



- Gerding, A. (2016). *Informationswert des Anhangs zur Jahresrechnung*, Wiesbaden: Gabler.
- Gunning, R. (1971). *The technique of clear writing*. New York: McGraw-Hill.
- Guthrie, J., Petty, R., Ricceri, F., & Yongvanich, K. (2004). Using narrative analysis as a research method to inquire into intellectual capital reporting. *Journal of Intellectual Capital*, 5(2), 282–293.
- Hassan, O. A. G., & Marston, C. (2019). Corporate financial disclosure measurement in the empirical accounting literature: A review article. *The International Journal of Accounting*, 54(2). <https://doi.org/10.1142/S1094406019500069>
- Hoitash, R., & Hoitash, U. (2018). Measuring accounting reporting complexity with XBRL. *The Accounting Review*, 93(1), 259–287.
- Hölscher, R., & Schneider, J. (2014). Multidimensionale Investitionsbewertung - Einsatz von Checklistenverfahren, Nutzwertanalyse und Preis-Leistungsmodell zur Lösung mehrdimensionaler Investitionsentscheidungen - Teil 2: Lösung. *Wirtschaftswissenschaftliches Studium*, 43, 393–397.
- Jones, M. J., & Shoemaker, P. A. (1994). Accounting narratives - A review of empirical studies of narrative and readability. *Journal of Accounting Literature*, 13, 142–184.
- Kaya, D., & Seebeck, A. (2018). Computergestützte Textanalyse von erweiterten Bestätigungsvermerken. *WPg - Wirtschaftsprüfung*, 16, 995–1001.
- Koonce, L., McAnally, M. L., & Mercer M. (2005). How do investors judge the risk of financial items? *The Accounting Review*, 80(1), 221–241.
- Krippendorff, K. (2004a). Reliability in narrative analysis – Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411–433.
- Krippendorff, K. (2004b). Measuring the reliability of qualitative text analysis data. *Quality & Quantity*, 38(6), 787–800.
- Li, F. (2010). Textual analysis of corporate disclosures - A survey of the literature. *Journal of Accounting Literature*, 29, 143–165.
- Lindemann, J. (2006). Kapitalmarktrelevanz der Rechnungslegung - Konzepte, Methodik und Ergebnisse empirischer Forschung. *Zeitschrift für Betriebswirtschaft*, 76(10), 967–1003.
- Linsley, P. M., & Shrivs, P. J. (2006). Risk reporting - A study of risk disclosures in the annual reports of UK companies. *The British Accounting Review*, 38(4), 387–404.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? - Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance - A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Marston, C. L., & Shrivs, P. J. (1991). The use of disclosure indices in accounting research - A review article. *The British Accounting Review*, 23(3), 195–210.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse – Grundlagen und Techniken*. Weinheim/Basel: Beltz.
- Mayring, P., & Fenzl, T. (2014). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 543–556). Wiesbaden: Verlag für Sozialwissenschaften.
- Miihkinen, A. (2012). What drives quality of firm risk disclosure? *The International Journal of Accounting*, 47(4), 437–468.
- Miihkinen, A. (2013). The usefulness of firm risk disclosures under different firm riskiness, investor-interest, and market conditions - New evidence from Finland. *Advances in Accounting*, 29(2), 312–331.
- Nagy, R. A., & Obenberger, R. W. (1994). Factor influencing individual investor behavior. *Financial Analysts Journal*, 50(4), 63–68.
- Ntim, C. G., Lindop, S., & Thomas, D. A. (2013). Corporate governance and risk reporting in South Africa - A study of corporate risk disclosures in the pre- and post-2007/2008 global financial crisis periods. *International Review of Financial Analysis*, 30, 363–383.
- Pérignon, C., & Smith, D. R. (2010). The level and quality of value-at-risk disclosure by commercial banks. *Journal of Banking & Finance*, 34(2), 362–377.
- Pike, R., Meerjanssen, J., & Chadwick, L. (1993). The appraisal of ordinary share by investment analysts in the UK and Germany. *Accounting and Business Research*, 23(92), 489–499.
- Reck, R. (2016). *Kultur und Sprache im Geschäftsbericht - Gestaltung von Textinformationen*. Wiesbaden: Gabler.
- Roseberry, R. L. (1995). A texture index - Measuring texture in discourse. *International Journal of Applied Linguistics*, 5(2), 205–223.
- Ryan, S. G. (2012). Risk reporting quality - Implications of academic research for financial reporting policy. *Accounting and Business Research*, 42(3), 295–324.
- Stein, T. (2011). *Eine ökonomische Analyse der Entwicklung der Lageberichtsqualität. Ein Beitrag zur Diskussion um Regulierung und Deregulierung*. Wiesbaden: Gabler.
- Sydserrff, R., & Weetman, P. (1999). A texture index for evaluating accounting narratives. *Accounting, Auditing & Accountability Journal*, 12(4), 459–488.
- Wagenhofer, A., & Ewert, R. (2015). *Externe Unternehmensrechnung*. Berlin/Heidelberg: Gabler.
- Yekini, L. S., Wisniewski, T. P., & Millo, Y. (2016). Market reaction to the positiveness of annual report narratives. *The British Accounting Review*, 48(4), 415–430.
- Zangemeister, C. (1976). *Nutzwertanalyse in der Systemtechnik – Eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen*. München: Zangemeister & Partner.
- Zhou, S., Jeong, H., & Green, P. A. (2017). How consistent are the best-known readability equations in estimating the readability of design standards? *IEEE Transactions on Professional Communication*, 60(1), 97–111.

**How to cite this article:** Helms, N., Hölscher, R., Nelde, M., & Schneider, J. (2021). The quality of risk reports: Integrating requirement levels of standard setters into text analysis. *Journal of Corporate Accounting & Finance*, 32, 7–23.  
<https://doi.org/10.1002/jcaf.22494>