JODDE Journal of Open, Distance, and Digital Education

Müskens, W. (2026). Assessing the quality of open educational resources: Validation and norming of the IQOER rating scales. *Journal of Open, Distance, and Digital Education, 3*(1), 1-23.
https://doi.org/10.25619/w6s58027

Received: 10 July 2025 | Revised: 30 October 2025 | Accepted: 06 December 2025

Assessing the quality of open educational resources: Validation and norming of the IQOER rating scales

Wolfgang Müskens[1]

[1] Center for Open Education Research (COER), Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

Correspondence:
Wolfgang Müskens | e-mail: wolfgang.mueskens@uni-oldenburg.de

Abstract

This article presents an instrument for assessing the quality of open educational resources (IQOER Rating Scales), which can be used by authors and developers of open educational resources (OER) to evaluate the quality of their own resources. The instrument is based on the comprehensive quality model for higher education OER developed by Zawacki-Richter and Mayrberger (2017) and was created and validated through an extensive multi-stage process. This paper describes the final stage of validation and the norming of the instrument. Norming establishes a reference to the parallel IQOER Classification Scales, which define a minimum and a premium level for each quality dimension. As part of the final validation and norming stage, 50 OER were evaluated by 45 raters using the IQOER Rating Scales, the IQOER Classification Scales, and the MERLOT Peer Review Form. The results show that the IQOER Rating Scales demonstrate strong psychometric properties and substantial correlations with global quality measures, confirming their practical relevance for assessing key aspects of OER quality. The empirically derived thresholds offer indications of whether an OER meets minimum or premium standards, supporting developers and educators in systematically improving their resources. By combining rigorous validation with practical applicability, the IQOER Rating Scales contribute to enhancing the quality and effectiveness of OER in higher education.

Keywords

open educational resources (OER); quality assessment; validation; norming; higher education

1 Introduction

In recent years, the movement of open educational resources (OER) has evolved dynamically in the context of the global digital transformation (Meng et al., 2020; Mishra et al., 2022). It is closely linked to international educational initiatives that embed openness, participation, and equitable access to knowledge as guiding principles of the digital knowledge society (UNESCO, 2019, 2023). Today, OER have become a central component of global strategies to promote inclusion, innovation, and sustainable education (Farrow et al., 2024; Tlili et al., 2022).

The aim of OER is to provide freely available teaching, learning, and research materials that can be used, adapted, and redistributed by teachers and learners. In doing so, they contribute to democratizing access to education and to strengthening participation and collaboration in higher education teaching and research (Hilton, 2020; Iniesto & Bossu, 2023). At the same time, international declarations such as the UNESCO Recommendation on OER (2019) emphasize that openness alone does not guarantee quality. The mere availability of open materials becomes educationally effective only when it is accompanied by pedagogical, technical, and institutional quality (Camilleri et al., 2014; Lopes et al., 2025).

Recent research shows that OER quality represents a multidimensional construct encompassing content-related, didactic, technical, and social aspects (Oliveira et al., 2021). Mishra (2025) and Otto et al. (2021) note that although questions of quality are gaining importance in OER research, they remain theoretically and empirically underdeveloped. This gap is particularly evident in the discrepancy between political demands for quality assurance (UNESCO, 2019, 2023) and the limited number of validated instruments for assessing quality (Lopes et al., 2025; Yuan & Recker, 2015).

Against this backdrop, the present study develops and validates the long version of the Instrument for Quality Assurance of Open Educational Resources (IQOER), an empirically grounded assessment tool that conceptualizes quality as a multidimensional construct consisting of content-related, pedagogical-didactic, and usability-related dimensions. The project's goal is to provide a reliable and practical instrument that enables authors and instructors in higher education to self-assess and further develop the quality of their OER. Furthermore, the study explores the relevance of the pedagogical and didactic design for university instructors' understanding of OER quality.

2 Background and conceptual framework

2.1 OER in higher education

In higher education, OER also demonstrate considerable potential to transform teaching and learning practices. They promote open, collaborative, and adaptive forms of instruction that enhance educators' autonomy and foster active student engagement (Martin & Kimmons, 2020; Tietjen & Asino, 2021). By enabling the adaptation and sharing of materials, OER foster new forms of co-creating learning resources that blur traditional boundaries between teachers and learners (Axe et al., 2020; Clinton-Lisell, 2021).

OER create environments for participatory and inquiry-based learning in which students move from being consumers of knowledge to becoming authors, curators, and producers of content (Adedoyin et al., 2025; Tietjen & Asino, 2021). Such learning environments foster creativity, self-directed learning, and critical media literacy.

Moreover, OER promote institutional openness within universities. They foster stronger collaboration among instructors, increase the visibility of teaching materials, and support the

development of shared quality standards (Farrow et al., 2024; Versantvoort & Schuwer, 2023). In many universities, OER initiatives serve as drivers of digital and pedagogical innovation by establishing channels for exchange and feedback among departments, students, and external communities (Iniesto & Bossu, 2023).

Finally, OER have inspired new perspectives on quality and sustainability in higher education. Quality is increasingly viewed as the outcome of collaborative practices that promote openness, transparency, and reflexivity in teaching and learning (Baas et al., 2019; Lopes et al., 2025; Otto et al., 2021).

2.2 Quality of OER

The quality of OER remains one of the central challenges for their sustainable use in higher education. Earlier research emphasized the need to systematically consider quality criteria in order to promote trust and acceptance among instructors (Camilleri et al., 2014; Yuan & Recker, 2015). In recent years, however, the focus has shifted: quality is no longer understood solely as a static property of materials but as part of a complex interplay of pedagogical, technical, and institutional factors (Otto et al., 2021; Stracke et al., 2019).

From a sociocultural perspective, quality is partly determined by the practices and norms of the communities involved in the creation and reuse of OER. Studies by Algers and Ljung (2015) and Versantvoort and Schuwer (2023) show that collaborative authorship, peer review, and open pedagogy can foster shared understandings of quality. But even in such scenarios of Open Educational Practices (OEP), reliable quality standards form the foundation for the collaborative development of OER.

Numerous empirical studies and meta-analyses in recent years have examined the impact of OER on learning outcomes. Systematic reviews consistently show that students who use OER achieve, on average, comparable or slightly better results than those using traditional learning materials (Colvard et al., 2018; Cho & Permzadian, 2024; Hilton, 2020; Tlili et al., 2023). These effects can largely be explained by increased accessibility, lower costs, and greater pedagogical flexibility. At the same time, the reviews highlight that the OER used in most studies were selected by instructors or subject-matter experts, while students themselves are often unable to reliably assess the quality of OER available in repositories,

Bibliometric analyses further indicate that the number of available OER and related publications has grown exponentially since 2015 (Meng et al., 2020; Mishra et al., 2022). This expansion represents a gain for openness and innovation, but simultaneously complicates the systematic evaluation and selection of high-quality resources. Especially in large repositories, materials vary widely in level, structure, and licensing, making consistent assessments of quality even more challenging (Oliveira et al. 2021).

2.3 The quality model for OER in higher education by Zawacki-Richter and Mayrberger (2017)

Zawacki-Richter and Mayrberger (2017) observed that the quality assessment of OER had, up to that point, been characterized by considerable conceptual and methodological heterogeneity. Existing models tended to focus on individual aspects such as technical accessibility, instructional design, or content accuracy, but rarely considered all dimensions systematically within a unified framework. This fragmentation made it difficult to compare assessment approaches and led to divergent understandings of OER quality in research and practice.

Against this backdrop, the approach proposed by Zawacki-Richter and Mayrberger (2017) sought to systematically analyze and consolidate various international quality models and to integrate

them into a comprehensive quality model for OER in higher education. The authors justified this approach based on three central challenges:

1. the heterogeneity of quality concepts across disciplinary and institutional contexts,
2. the limited empirical foundation of many earlier models, and
3. the lack of integration among technical, content-related, and pedagogical perspectives.

Recent systematic reviews (Lopes et al., 2025; Mishra, 2025; Oliveira et al., 2021) show that even today, no internationally standardized or empirically validated quality models have been established. Instead, a variety of theoretical approaches and sets of criteria continue to coexist, differing by discipline, medium, or target group. Oliveira et al. (2021) emphasize that most existing models are conceptually grounded but seldom empirically validated. Similarly, Lopes et al. (2025) point out that while current quality frameworks cover a broad range of dimensions, they are rarely subjected to psychometric testing and remain only partially implemented in practice.

As part of a project by the Hamburg Open Online University (HOOU), Zawacki-Richter and Mayrberger (2017) conducted an extensive literature review in 2015/16 on approaches to evaluating the quality of OER. They identified a total of eight internationally established instruments, which served as the foundation for their own quality model:

1. Rubric for Selecting Inquiry-Based Activities (Fitzgerald & Byers, 2002).
2. Learning Object Evaluation Instrument (LOEI; Haughey & Muirhead, 2005).
3. Learning Object Review Instrument (LORI; Nesbit et al. 2009).
4. Learning Objects Quality Evaluation Model (eQNet; Kurilovas et al., 2011).
5. OER Rubric (Achieve Inc., 2011).
6. Rubric to Evaluate Learner Generated Content (LGC; Pérez-Mateo et al., 2011).
7. Framework for Assessing Fitness for Purpose in OER (Jung et al., 2016).
8. Multimedia Educational Resource for Learning and Online Teaching (MERLOT Rubric; MERLOT, n.d.).

They compared these models with regard to their quality dimensions, evaluation procedures, scales, and operationalizations. Altogether, these procedures used 161 different quality criteria.

In a synoptic summary, the similarities and differences among the criteria were then analyzed. Of the 161 criteria, 117 addressed pedagogical-didactic and content-related aspects of OER, while 37 concerned technical aspects of the resources.

This synoptic summary initially resulted in an overall model with 29 dimensions. In a subsequent step, incorporating input from OER experts in focus groups, this was condensed into a model with 15 dimensions (Mayrberger et al., 2018; see Figure 1).

The dimensions of the quality model by Zawacki-Richter and Mayrberger include, alongside content-related dimensions, pedagogical-didactic aspects (labelled “Instructional Design”), as well as technical dimensions divided into the areas of “Accessibility” and “Usability”.

In a first attempt at operationalizing this model, Mayrberger et al. (2018) developed corresponding scales for each dimension.

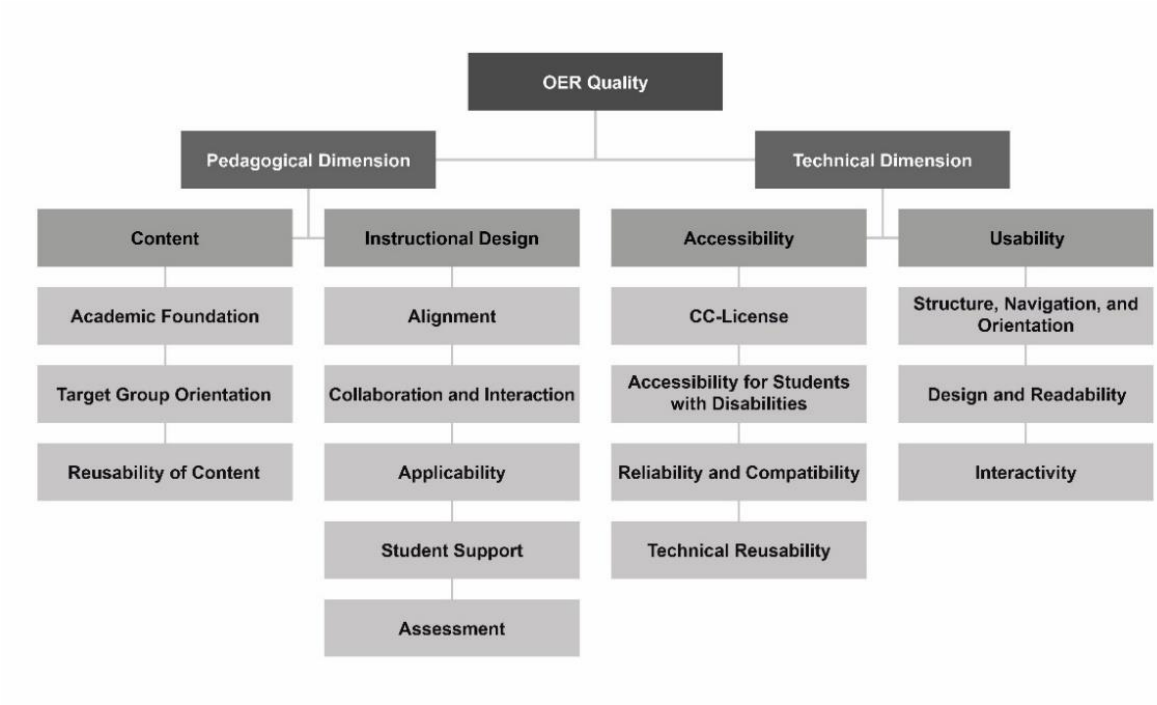

Figure 1: Dimensions of OER quality

Note. Adapted from „Qualitätsentwicklung von OER – Vorschlag zur Erstellung eines Qualitätssicherungs-instrumentes für OER am Beispiel der Hamburg Open Online University. Sonderband zum Fachmagazin Synergie“ by K. Mayrberger et al., 2018. Hamburg: University Hamburg. (https://doi.org/10.25592/978.3.924330.67.5). CC-BY.

3 Development of the IQOER

3.1 Development of the IQOER

The revised scales developed by Mayrberger et al. (2018) form the basis of what is referred to as the Instrument for Quality Assurance of OER (IQOER).

The norming of the IQOER rating scales marks the final stage of an extensive series of development and validation studies conducted in Germany since 2018 (see Figure 2). This process focused on the two main versions of the IQOER instrument: the IQOER Classification Scales and the IQOER Rating Scales.

In an initial step (see right side of the diagram in Fig. 2), the IQOER Classification Scales — also referred to as the short version of the instrument — were developed based on focus groups with educational experts (Mayrberger et al., 2018). These classification scales define, for each dimension of the Zawacki-Richter and Mayrberger (2017) model, both a minimum standard and a premium standard.

It should be emphasized that the distinction between minimum and premium standards within the IQOER framework is not intended as a normative judgment but as a criterion-referenced differentiation that primarily serves a formative and developmental purpose. The minimum standard defines the basic level of quality required for a resource to be pedagogically, technically, and content-related suitable for use in open education. The premium standard, by contrast, denotes characteristics of particularly exemplary or innovative OER that can serve as reference points for further development. It should be noted that the standards defined here refer explicitly to the level of the scales. The instrument is not intended to classify OER into quality categories but to offer formative guidance for improving specific aspects of quality. From 2020 onwards,

and building on methods from test and measurement theory, the IQOER Rating Scales — also referred to as the long version of the instrument — were developed (see left side of the diagram in Fig. 2). These scales cover 11 dimensions of the Zawacki-Richter and Mayrberger (2017) model. Only the four dimensions related to accessibility, which require a technical analysis of the resources, were excluded from this instrument. The dimension "Student Support" was renamed to "Assistance and Support", and the dimension "Applicability" was changed to "Application and Transfer".

In addition, the dimension "Motivation", which was not included in the original model by Zawacki-Richter and Mayrberger (2017) was added to the rating scales (Lübben et al., 2023). The items of the IQOER "Motivation" scale (Table 2) indicate that this dimension does not measure learners' individual motivation to learn but rather the motivational potential of the OER's design. Statements such as "The design of the OER encourages learners to engage with the content" or "The contents are arranged in an interesting way" refer to observable design features that can foster curiosity and engagement. In this sense, motivation is understood not as a psychological state but as a didactic design principle.

The IQOER Rating Scales are based on Likert-type scales. Each of the 12 included dimensions is measured using 5 to 6 items, which are rated on a 5-point scale.

Some scales in the Zawacki-Richter and Mayrberger (2017) model are not suitable for all types of OER. For example, the scales "Assessment" and "Collaboration and Interaction" are only relevant for resources specifically designed with such features. Initial investigations showed that seven of the 12 dimensions could be applied to almost all OER. These seven scales are referred to as the core scales of the IQOER, and they are the focus of the standardization presented here.

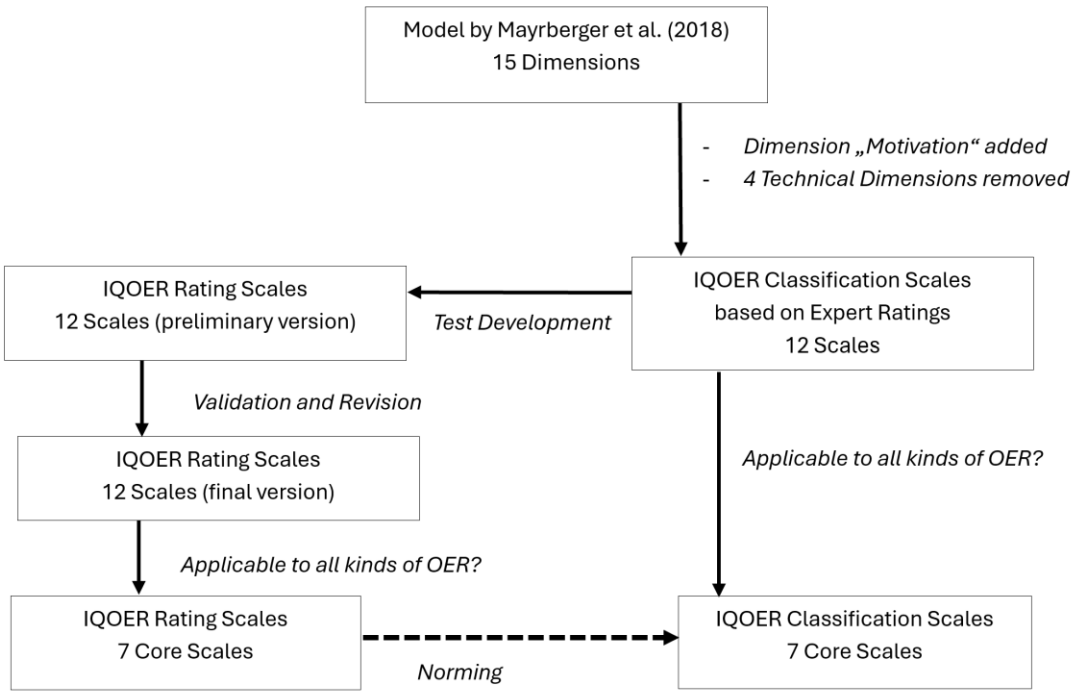

Figure 2: Development Validation of the IQOER

In the first validation study, 50 individuals participated as raters. They assessed a total of 101 OER, which were randomly selected German-language resources for higher education. Each rater was tasked with evaluating two different OER.

The internal consistencies of the IQOER core scales, measured by Cronbach's alpha, ranged from $\alpha = .79$ to $\alpha = .92$, indicating medium to high reliability.

To assess the convergent validity of the scales, correlations were calculated between the mean scores of the item scales and the corresponding classification scales. These correlations ranged from $r = .53$ to $r = .87$ ($p < .001$).

The core scales of the IQOER long version correlated with the overall scale of the MERLOT Peer Review Form, with values between $r = .36$ and $r = .64$ across all assessments. Meanwhile, the classification scales of the IQOER short version correlated between $r = .38$ and $r = .61$ ($p < .001$) with the MERLOT overall scale.

The analysis of the IQOER rating scale items provided detailed insights for optimizing the scales. It suggested that reliability could be further improved by removing or revising specific items. Therefore, the item scales of the long version were revised for the second validation stage, based on the findings of the first. The present paper reports on this second stage of the validation.

3.2 Application possibilities of the IQOER rating scales and norming

Unlike the IQOER Classification Scales, the primary purpose of the IQOER Rating Scales is to enable self-evaluation of OER by their authors and developers. These individuals have in-depth insight into the content design, didactic structure, and technical implementation of their resources. The long version of the IQOER is intended to provide them with an easy-to-use tool for comprehensively assessing all relevant quality aspects of OER development and for critically reviewing their own resources.

Authors and developers may be particularly interested in whether their resources meet the quality standards defined by experts. However, until now, the IQOER Rating Scales have not allowed for direct statements about compliance with these standards. The ratings merely result in a mean scale score for each set of items. To relate these mean scale scores to the minimum and premium standards defined in the IQOER Classification Scales, a criterion-referenced norming of the scales is necessary. In this norming process, the specific mean scale values are determined at which the minimum and premium standards of the corresponding classification scales are just achieved.

This approach allows authors to interpret the results of their self-assessments in relation to the expert-defined quality benchmarks. It provides a practical orientation tool for formative evaluation, helping developers identify areas where their OER meet or fall short of the expected standards, and thereby supporting targeted improvements in the design and implementation of their resources.

3.3 Aims and research questions

First, the fundamental psychometric properties of the revised version of the IQOER Rating Scales are to be determined. This aims to establish whether the revision of the instrument has led to improvements in its reliability and validity.

In addition, a norming of the instrument relative to the minimum and premium standards defined in the IQOER Classification Scales is pursued. This norming is intended to enable criterion-referenced interpretations of the IQOER Rating Scale scores regarding the achievement of the expert-defined minimum and premium standards.

The results of this norming process will form the basis for an interactive online version of the IQOER Rating Scales. This online tool is intended to be freely available to OER developers and authors so they can formatively evaluate whether their materials meet the established standards.

Moreover, with regard to the model developed by Zawacki-Richter and Mayrberger (2017), the question arises as to the validity of this model in capturing the quality of OER. To what extent do the dimensions chosen by the authors comprehensively reflect the quality aspects relevant to higher education OER?

As Yuan and Recker (2015) point out in their seminal review, many existing evaluation instruments and repositories focus primarily on formal and content-related criteria such as accuracy, currency, or technical accessibility. Pedagogical and didactic features—such as alignment with learning objectives, interactivity, or support for self-directed learning—are often not taken into account. Yuan and Recker criticize that quality in these models is overly oriented toward the material level and less toward the learning or application level: "[...] this review indicated that the quality and educational utility of OER depended not only on the content quality (e.g., accuracy, clarity) of resources, but also on the pedagogical values contained in these OER" (Yuan & Recker, 2015, p. 24).

The well-known UNESCO (2019) definition of OER encompasses a very broad range of materials. OER are described as "teaching, learning and research materials in any medium – digital or otherwise – that reside in the public domain or have been released under an open license" (p. 5). This definition thus also includes research reports, policy documents, and other scholarly publications that are not pedagogically or didactically designed. Such diversity makes it necessary to examine empirically—rather than assume theoretically—the relevance of pedagogical dimensions for the perception of OER quality.

Accordingly, the present study examines whether the dimensions included in the IQOER—particularly the pedagogical and didactic ones—empirically contribute to the perception of OER quality. Specifically, it investigates whether an integrative model that encompasses both content-related and didactic dimensions better captures the heterogeneity of OER in higher education than models limited to technical or content-related aspects of quality.

4 Instruments

4.1 The IQOER rating scales

For the norming process, only the seven core scales of the IQOER were considered. These scales are "Scientific Foundation", "Content Reusability", "Application and Transfer", "Assistance and Support", "Motivation", "Structure, Navigation, and Orientation", and "Design and Readability". According to the model by Zawacki-Richter and Mayrberger (2017), the scales "Scientific Foundation" and "Content Reusability" belong to the "Content" category; the scales "Application and Transfer", "Assistance and Support", and "Motivation" are part of the pedagogical-didactic category "Instructional Design"; and the scales "Structure, Navigation, and Orientation" and "Design and Readability" fall under the technical category "Usability".

The IQOER was used in a German-language version (Müskens et al., 2021).

4.2 IQOER classification scales

The IQOER Classification Scales consist of a single item for each dimension. Figure 3 illustrates the classification item for the scale "Scientific Foundation" as an example. The scale has five levels; however, only levels 1, 3, and 5 are explicitly described. Levels 2 and 4 are intended to serve as intermediate stages between these anchor points. Level 3 represents the minimum

standard as defined by OER experts, while level 5 indicates a desirable premium standard. Level 1 describes a clear failure to meet the minimum standard.

The classification scales were developed in 2018 through focus groups with education experts in the field of OER, as part of the Hamburg Open Online University project (Mayrberger et al., 2018).

4.3 IQOER rating scales

The long version of the IQOER used here consists of 5 per dimension. The items are aggregated to mean scales for each dimension. The long version of the instrument comprises a total of 35 items (Table 1).

The long version was revised in terms of form and content based on the results of the first stage of validation. The aim of the revision was to optimize the test regarding objectivity, reliability and validity.

Each of the 63 items of the long version of the IQOER was to be rated by the raters using a 5-point rating scale ("does not apply at all", "rather does not apply", "applies somewhat", "largely applies", "fully applies"). In addition, for each item the alternative "rating not possible" was available, which the raters were asked to select if it was not possible to categorize an OER with regard to the item.

	The content is scientifically correct, up-to-date, and relevant. They are presented in a precise and balanced manner. Bibliographical sources are consistently cited in accordance with academic standards. The argumentation is coherent and clearly comprehensible. (5)
	(4)
	The content is scientifically correct and relevant. The argumentation is comprehensible. The origins of models, methods, and approaches are mostly named. (3)
	(2)
	The content has no recognisable reference to research results. The underlying methods or approaches are either not presented at all or are presented without reference to their origin. (1)

Figure 3: Example classification item for the “Scientific Foundation” scale in the IQOER

Note. Adapted and translated into English by the author from an example item in Müskens et al. (2022, March). Instrument zur Qualitätssicherung von OER – IQOER: Entwicklungsversion 17 [Unpublished manuscript, p. 3]. ResearchGate. https://doi.org/10.13140/RG.2.2.16987.03363/1. Licensed under CC BY 4.0.

4.4 MERLOT peer review form

To determine the construct validity of the IQOER Rating Scales regarding assessing OER quality, an additional instrument with the same purpose was used in the survey. This instrument was a German translation of the MERLOT Peer Review Form (MERLOT, n.d.), which consists of a total

of 31 items. The instrument comprises three scales: "Quality of content" (12 items), "Potential effectiveness as a teaching tool" (8 items), and "Ease of use" (7 items). Additionally, each scale includes a single item that provides an overall assessment of the respective scale. A fourth scale (MERLOT aggregated scale) combines all evaluations related to the material under review. These items use a six-point school grading scale. For the German version, the scale was adapted to the standard German school grading system (1 – best grade, 6 – worst grade).

4.5 Global quality rating

The evaluation questionnaire also included a single item for the global assessment of an OER's quality. This item read: "How would you rate the overall quality of the OER you have worked on (i.e., in terms of content as well as design and implementation)?" This item also used the German six-point school grading scale (1 – best grade, 6 – worst grade).

5 Method

5.1 OER to be evaluated

A total of 50 different OER were evaluated. These included 24 resources with content in the social sciences and humanities, 12 OER with STEM content (science, technology, engineering, and mathematics), and 14 OER that combined elements from both the social sciences/humanities and STEM fields.

Only OER published under a Creative Commons licence (CC BY-SA or CC BY) and aimed at higher education students (ISCED 6/7) were included. Almost all materials originated from the repositories of German federal states or public universities:

- Twillo of the state of Lower Saxony (https://www.twillo.de/), which was still called "OERNds" at the time of the study (September 2020);
- HOOU (https://www.hoou.de/), the repository of the Hamburg Open Online University;
- OpenLearnWare (https://www.openlearnware.de/) of TU Darmstadt;
- ZOERR (https://www.oerbw.de/), the repository of the universities of the state of Baden-Württemberg.

The resources were available in multiple formats: 54% of the OER consisted of video clips, 14% were PDF documents or other text-based materials, 18% combined several formats, and 14% comprised other types such as simulations, presentations, and other digital media formats. Downloading was required only for PDF documents.

5.2 Raters

A total of 45 raters participated in the second stage of the validation. Of these, 44 raters each evaluated two different resources, while one rater evaluated only one resource. In total, 89 OER were assessed. Among these, 37 different resources were assessed by two raters, 12 resources by one rater, and one resource by three raters. Figure 4 shows the composition of this sample.

The raters were students and lecturers at German universities.

Of the raters, 64% identified as male and 36% as female; none selected the "diverse" option. The average age of the raters was 36.6 years (SD = 11.6).

Regarding academic background, 71% of the raters stated that they were studying or working in a humanities subject, 18% in a social science, economics, or law subject, and 11% in a STEM field.

Raters from a STEM field mostly received OER with STEM content for evaluation, while raters from the other disciplines primarily assessed resources with social science or humanities content.

In terms of educational qualifications, 75% of the raters reported already holding a master's degree, 22% held only a bachelor's degree, and one rater did not yet have a university degree.

Moreover, 47% of the raters had already been involved in the development of OER themselves, and 82% reported having prior experience using OER.

The raters did not receive detailed instructions on how to engage with the resources. They were instructed to review shorter OER in full and, for longer ones (with more than two chapters or files), to complete at least the first two sections. It was expected that the raters would adopt the perspective of a self-directed learner engaging with a largely unfamiliar topic.

5.3 Survey procedure

The survey took place from January 2021 to April 2021. Raters from STEM fields were assigned corresponding OER with STEM content, while raters from the social sciences and humanities were given OER relevant to those disciplines. OER that could be classified under both areas were randomly distributed among all raters.

Each rater received their assigned OER along with access to an online platform containing the evaluation instruments. The raters were instructed to work through the OER as completely as possible; for extensive resources, they were asked to process at least the first two chapters or sections. After completing their review of an OER, raters were required to evaluate it using the online instruments: the IQOER Classification Scales, the IQOER Rating Scales, the MERLOT Peer Review Form, and the global quality rating item. No time limit was set for completing the evaluations.

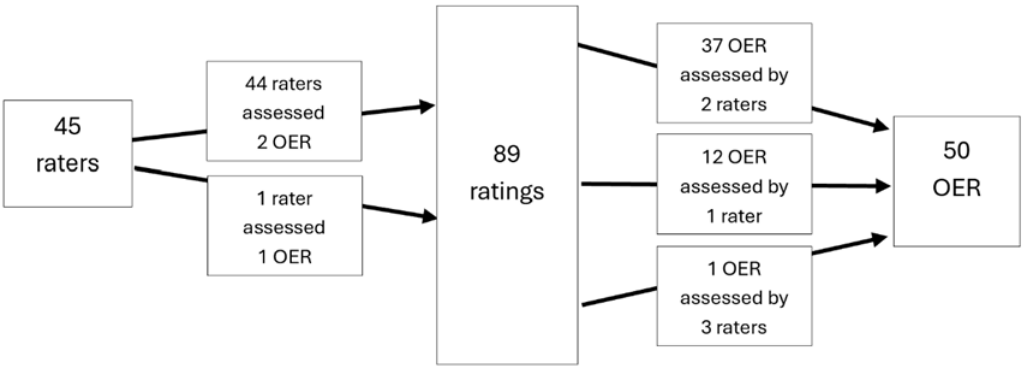

Figure 4: Flowchart of the Allocation of Ratings and Composition of the OER Sample

Note. This figure illustrates the distribution of 89 ratings across 45 raters and 50 OER. It shows how many resources were assessed by one, two, or three raters. Created by the author.

6 Results

6.1 Validation of the IQOER rating scales

6.1.1 Item statistics

Table 1 presents the descriptive statistics of the 35 rating items of the IQOER. In addition to item discrimination (corrected item-total correlation), the correlation between each item and its corresponding IQOER classification scale was also calculated. It can be observed that these two

coefficients are generally very similar. This can be interpreted as evidence for the convergence between the mean scale scores of the rating items (total scale) and the respective classification scales.

Table 1: Items of the IQOER Rating Scales (English translation of German Items)

No.	Item	Scale	N	M	Std.dev.	dis	r(ic)
01	In the OER, bibliographic references are consistently given in accordance with the subject-specific standards.	SCF	74	3,00	1,58	,56**	,70***
02	The origin of models, methods, and approaches is clearly specified in the OER.	SCF	85	3,28	1,39	,56**	,69***
03	The content is up-to-date, accurate, and relevant.	SCF	76	4,24	,88	,55**	,50***
04	The reasoning in the materials is coherent and comprehensible.	SCF	88	4,18	1,09	,52**	,56***
05	The presentation of the content is precise and balanced.	SCF	87	3,93	1,09	,65**	,62***
06	The resource contains or consists of learning objects that can be used without modification in courses or other contexts (e.g. in a degree program).	CRU	85	3,74	1,25	,35	,81***
07	The use of the material does not necessarily require the use of other materials.	CRU	86	3,73	1,27	,50	,46***
08	The OER is closely related to other materials, which knowledge is a prerequisite for understanding the content. (recoded)	CRU	83	3,45	1,38	,11	,14
09	The content/the contents of the material are clearly defined.	CRU	82	3,94	,99	,55	,50***
10	The learning material is self-contained in its contents.	CRU	87	4,05	1,21	,59	,44***
11	The material offers suggestions for applying what has been learned to practice to new problem situations or to practice.	AAT	88	3,17	1,27	,66	,81***
12	The material includes case studies from practice.	AAT	89	3,63	1,43	,63	,42***

No.	Item	Scale	N	M	Std.dev.	dis	r(ic)
13	Within the material, learners are asked to transfer content or methods to new problem situations or practice.	AAT	88	2,65	1,24	,53	,62***
14	The material includes theories or methods that can be applied in practice.	AAT	86	3,59	1,27	,69	,50***
15	The material contains references to (professional) practice.	AAT	81	3,52	1,31	,68	,38***
16	The material offers assistance on subject-specific terms (e.g., glossary, definitions).	AAS	86	2,43	1,43	,54	,76***
17	The resource includes tables, lists, or graphs that summarize or illustrate the topics discussed.	AAS	86	2,97	1,44	,51	,47***
18	The OER includes a summary of the content presented.	AAS	85	2,72	1,50	,57	,41***
19	The OER includes questions that learners can use to check their understanding of the content.	AAS	85	1,94	1,41	,59	,51***
20	The material includes special assistances that support learning.	AAS	85	2,19	1,31	,59	,67***
21	The design of the OER is unique.	MOT	89	2,73	1,29	,86	,85***
22	The design of the OER encourages learners to engage with the content.	MOT	87	3,26	1,28	,91	,86***
23	The contents are arranged in an interesting way.	MOT	88	3,27	1,27	,87	,80***
24	The OER arouses interest in the subject by its design.	MOT	88	3,07	1,40	,87	,82***
25	In some of the elements of OER, at least the first steps of a didactic design are recognizable.	MOT	89	3,40	1,16	,64	,65***
26	The learners always know where they are in the material.	SNO	86	3,86	1,26	,78	,79***
27	The navigation structure is clear.	SNO	87	3,85	1,29	,84	,85***
28	It is possible to access all previous content at any time.	SNO	84	4,29	1,12	,73	,68***

No.	Item	Scale	N	M	Std.dev.	dis	r(ic)
29	The structure is simple and clear.	SNO	88	3,98	1,18	,81	,77***
30	Learners can interrupt the learning sequence at any time and continue later from the same point.	SNO	88	4,55	,96	,61	,52***
31	The whole text is easily readable throughout.	DAR	71	4,24	1,06	,85	,86***
32	The use of the OER is not constrained by the design.	DAR	83	4,28	1,04	,74	,70***
33	All parts of the OER have an integrated, coordinated design.	DAR	79	4,03	1,30	,75	,78***
34	The presentation suits the content.	DAR	80	3,96	1,15	,72	,76***
35	All illustrations of the OER are easily recognizable.	DAR	76	4,17	1,12	,84	,82***

Note. Dimensions: SCF – Scientific Foundation, TGO – Target Group Orientation, CRU – Content Reusability, ALN – Alignment, COI – Collaboration and Interaction, AAT – Application and Transfer, AAS – Assistance and Support, ASM – Assessment, MOT – Motivation, SNO – Structure, Navigation, and Orientation, DAR – Design and Readability, INT – Interactivity. Items are rated on a scale from 1 – does not apply at all to 5 – fully applies (with item 8 reverse-coded). M = mean; dis = item discrimination (corrected item-total correlation); r(ic) = correlation of the item with the corresponding classification scale. * p < .05; ** p < .01; *** p < .001. An English translation is provided for each item.

Item 8 was the only item formulated negatively; that is, the statement of the item corresponded to a low level on the "Content Reusability" scale. Therefore, the ratings for this item were reverse coded for presentation in Table 1. However, the item discrimination and the correlation with the corresponding classification scale remained low (r = .11 and r = .14, respectively).

6.1.2 Mean scales

For the seven IQOER core dimensions, mean scores were calculated based on the item rating scales. Cronbach's alpha values were determined for these scales as a measure of internal consistency (reliability; see Table 2).

Cronbach's alpha values for seven of the eight core scales were at an acceptable level of $\alpha > .75$. Only the "Content Reusability" scale showed lower consistency, with $\alpha = .65$.

The mean scores demonstrated strong correlations with their respective classification scales across all dimensions, and all correlations were highly statistically significant ($r \geqslant .70$; $p \leqslant .001$). These values can be interpreted as evidence of convergent validity, as both versions of the instrument are designed to measure the same constructs.

Furthermore, the correlations between the mean scores of the rating items and the two global measures of OER quality were analyzed. These global measures were the overall score from the MERLOT Peer Review Form and the separate item assessing the overall quality of the OER. The correlations between the mean scores and these global measures indicate their criterion validity, as they reflect the relationship with the perceived quality of the resource that the instrument aims to measure. Indeed, highly significant correlations with the global measures were found for all core scales of the IQOER.

6.2 Criterion-referenced norming of the rating scales

In the following, we examine the relationship between the classification scales and the corresponding mean rating scales of the IQOER in order to standardize the latter. Due to the sample sizes, only the core scales of the instrument were considered.

For this purpose, ordered logistic regressions of the rating scales on the corresponding mean scales were calculated. The PLUM (Polytomous Universal Model) procedure in SPSS 28 was used, based on the cumulative logit model proposed by McCullagh (1980).

The thresholds for the minimum and premium standards are of particular interest for the practical application of the mean scales; these are the values of the mean scales above which the minimum or premium standard is achieved. Knowing these thresholds enables the estimation of whether the respective standard has been met for the OER in question.

Table 2: Means of the IQOER Rating Scale Items

Scale	abbr.	N	M	Std.dev.	Alpha	r(rs)	r(gs)	r(mt)
Scientific Foundation	SCF	88	3.75	.93	.78	.82***	.58***	.51***
Content Reusability	CRU	86	3.77	.84	.65	.70***	.56***	.47***
Application and Transfer	AAT	88	3.32	1.02	.84	.70***	.54***	.41***
Assistance and Support	AAS	86	2.44	1.04	.78	.76***	.56***	.46***
Motivation	MOT	89	3.14	1.13	.94	.89***	.77***	.59***
Structure, Navigation, and Orientation	SNO	88	4.10	.98	.90	.86***	.28**	.35***
Design and Readability	DAR	79	4.14	.98	.92	.89***	.47***	.42***

Note. Mean scales of item ratings; M = mean; Std. Dev. = standard deviation; Alpha = Cronbach's alpha; r(rs) = Pearson correlation of the mean scores with the corresponding classification scale; r(gs) = Pearson correlation of the mean scores with the recoded global quality rating (scale inverted); r(mt) = Pearson correlation of the mean scores with the MERLOT aggregated scale (MERLOT aggregated scale inverted). $p < .05$; $p < .01$; *$p < .001$; n.c. = not calculated.

Therefore, we aimed to identify the mean scale value X where

$$p(Y = j) = p(Y = j + 1) \qquad \text{(Equation 1)}$$

with j = 1...m levels of the classification scale Y.

Transforming Equation 1 yields the following expression for the target Y:

$$p(Y \leq j) = \frac{p(Y \leq (j+1)) + p(Y \leq (j-1))}{2} \qquad \text{(Equation 2)}$$

The model equation of the cumulative logit model of the PLUM procedure for the cumulative probabilities p(Y ⩽ j) is given by:

$$p(Y \leq j) = \frac{e^{\propto j - \beta x}}{1 + e^{\propto j - \beta x}} \qquad \text{(Equation 3)}$$

PLUM provides estimators for the parameters α_j and for the level-independent location parameter β for j = 1…m-1.

By substituting the right-hand side of Equation 3 into Equation 2, the first and last thresholds of the mean scale can be calculated relatively easily.

For a three-level scale, the result for the first level with p(Y = 1) equals p(Y = 2) is

$$X_1 = \frac{ln\frac{(e^{\propto_1 + \propto_2})}{e^{\propto_2} - 2e^{\propto_1}}}{\beta} \qquad \text{(Equation 4)}$$

For the second level with p(Y = 2) = p (Y = 3), the result is

$$X_2 = \frac{ln(-2e^{\alpha_1} + e^{\alpha_2})}{\beta} \qquad \text{(Equation 5)}$$

To calculate the thresholds for the minimum and premium standards, the original five-level classification scale was transformed into a three-level scale by aggregating levels 1 and 2 (below minimum standard) and levels 3 and 4 (minimum standard). Table 3 shows the thresholds for the minimum and premium standards calculated using Equations 5 and 6.

Table 3: Thresholds for the Minimum and Premium Standards of the IQOER Mean Scales

Scale	**abbr.**	**N**	**Threshold for minimum standard (SE) (classification scale = 3)**	**Threshold for premium standard (SE) (classification scale =5)**
Scientific Foundation	SCF	76	2.99 (0,10)	4.27 (0.08)
Content Reusability	CRU	83	2.84 (0.19)	4.10 (0.14)
Application and Transfer	AAT	85	3.25 (0.14)	4.84 (0.20)
Assistance and Support	AAS	85	2.34 (0.17)	4.65 (0.34)
Motivation	MOT	89	2.60 (0.12)	4.38 (0.12)
Structure, Navigation and Orientation	SNO	85	3.20 (0.15)	4.45 (0.11)
Design and Readability	DAR	78	3.27 (0.12)	4.42 (0.10)

Notes: Thresholds calculated using ordered logistic regression (PLUM model) based on the three-level recoded classification scales. Thresholds indicate the mean scale values at which the minimum or premium standards defined by experts are met.

7 Discussion

The statistical analyses presented above provide evidence for the reliability and validity of the IQOER Rating Scales.

Cronbach's α coefficients for the individual scales ranged from .65 to .92. Cronbach's α is a measure of internal consistency—that is, the degree to which items within a scale correlate with one another and thus appear to measure the same underlying construct. From a practical standpoint, high internal consistency indicates that the items within each IQOER dimension form a coherent set. The raters tended to respond to these items in a similar manner, suggesting that the dimensions capture clearly defined and conceptually homogeneous aspects of OER quality.

The correlations between the IQOER Rating Scales and the shorter IQOER Classification Scales (r = .70 to .89) demonstrate that both formats are conceptually aligned. This means that the two instruments assess the same quality dimensions, at different levels of detail.

The positive and statistically significant correlations between the IQOER Scales and the MERLOT Peer Review Form (r = .35 to .59) provide evidence of convergent validity—that is, the extent to which different instruments designed to measure similar constructs yield consistent results.

Despite the relatively small sample size, the thresholds show only minor standard errors, indicating that the levels of the Classification Scales corresponding to the IQOER Rating Scale results can be determined with high precision.

A particularly noteworthy finding concerns the three scales in the Instructional Design domain—Motivation, Application and Transfer, and Assistance and Support.

The Motivation scale showed the strongest association with the global quality rating (r = .77), the highest correlation among all dimensions examined. This finding suggests that an appealing, learner-oriented design and a clear didactic structure were perceived by raters as especially relevant indicators of quality. Such characteristics correspond to the motivational design principles described by Keller (2010) in the ARCS model and align with empirical findings indicating that interest and engagement are key indicators of quality in digital learning resources (Evelyn & Kromer, 2018; Tseng et al., 2024).

The Application and Transfer scale likewise showed strong relationships with the global quality measures. OER that facilitate the transfer of knowledge to new contexts or establish practical relevance were consistently rated more highly. This underscores the importance of action-oriented and practice-based learning arrangements that enable students to actively apply what they have learned—a core principle of constructivist learning theory (Jonassen, 1999).

Finally, the Assistance and Support scale captures supportive elements within OER, such as glossaries, summaries, or self-assessment questions. This dimension also correlated highly significantly with the global quality indicators, showing that the provision of learning aids is perceived as an essential quality feature—particularly for independent and adaptive learning.

Taken together, these findings demonstrate that the perceived quality of OER depends to a large extent on didactic elements that foster learning motivation, practical application, and learner support. The results thus provide empirical evidence that pedagogical and didactic dimensions—not merely technical or content-related aspects—constitute integral components of a comprehensive understanding of OER quality.

8 Conclusion

Overall, the results indicate that the IQOER Rating Scales exhibit high internal consistency and convergent validity and provide meaningful reference points for interpreting and assessing OER quality.

A key limitation of the study is the small and unevenly composed sample of raters. With only 45 participants and a total of 89 OER evaluations, the findings should be interpreted with caution, as they may not fully represent the diversity of perspectives within higher education. Furthermore, the majority of raters came from social sciences and humanities backgrounds, while raters with STEM expertise were underrepresented. This imbalance could have biased the evaluations toward certain quality aspects prioritized in these disciplines. Future research should aim to validate the IQOER Rating Scales with a larger and more diverse sample that includes a balanced representation of disciplinary backgrounds to ensure the generalizability of the results across different academic contexts.

With the criterion-referenced norming of the scales used here, an easy-to-use interactive version is available to users, allowing them to see, as they respond to the items, whether the OER they have created meets the minimum or premium standards defined by the IQOER. Based on the calculated thresholds for the minimum and premium standards, an online version of the IQOER Rating Scales with automated evaluation was developed. This tool is available under a Creative Commons license and can be freely used by OER developers (https://umfragen.uni-oldenburg.de/index.php?r=survey/index&sid=413357&lang=en).

Although the development of the IQOER took place within the German-speaking higher education context, the instrument can also help address the international need for reliable and empirically grounded methods for assessing OER quality. For example, Tlili et al. (2022) show that in Africa, OER have not yet achieved adequate dissemination, and educators often doubt the quality of available resources. Freely available instruments for assessing the quality of OER may help establish quality assurance processes even in countries with less developed OER infrastructures.

Our findings highlight the crucial role of pedagogical and didactic dimensions in assessing the perceived quality of OER in higher education. By providing empirically established thresholds for minimum and premium standards, the IQOER Rating Scales offer a valuable tool for developers and educators to systematically evaluate and improve their resources. The strong correlations with global quality measures confirm that aspects such as motivation, structure, and support are central to creating effective and high-quality OER.

Like most assessment-based instruments, the IQOER also relies on human judgment and therefore inevitably involves a certain degree of subjectivity. The pedagogical backgrounds and contextual interpretations of the evaluators may influence how specific criteria are applied. Although this represents a recognized limitation in quality assessment, it simultaneously points to opportunities for further development.

The structured design of the IQOER—with clearly defined scales and descriptors—provides a possible foundation for future AI-supported evaluation tools. Models trained on annotated OER datasets could approximate human ratings and generate automated preliminary assessments of quality, for instance, by identifying missing didactic elements or weak alignment with learning objectives.

Beyond summative evaluation, such tools could also serve a formative function by supporting educators during the development of OER. When integrated into authoring environments, AI

systems based on the IQOER dimensions could provide real-time feedback that helps authors plan design steps, check coherence, and iteratively improve their materials.

Future research should focus on examining the use of the IQOER in real-world development and implementation contexts to empirically validate its effectiveness as an instrument for promoting quality in OEP. While the present study provides evidence of the psychometric properties of the scales, it remains an open question how the instrument can be applied in authentic teaching and learning settings, and what effects its use may have on quality, collaboration, and sustainability in OER development.

In addition, future research projects should explore how the IQOER can be integrated as a formative tool into digital development environments—such as repositories, authoring tools, or peer-review platforms. Achieving sustainable OER quality requires not only reliable instruments but also a systematic linkage of quality assurance with open practices, professional training, and institutional incentive structures (Ramírez-Montoya, 2020).

Overall, the IQOER should not be regarded in future research merely as an evaluation instrument but as a tool for reflection and development that contributes to the professionalization and sustainability of OEP.

References

Achieve Inc. (2011). *Rubrics for evaluating open educational resource (OER) objects.* https://www.achieve.org/files/AchieveOERRubrics.pdf

Adedoyin, O. B., Altinay, F., Altinay, Z., Shadiev, R., Baştaş, M., & Dagli, G. (2025). Open educational resources: Qualitative evaluation of faculty interest in motivating students to create and use. Innovations *in Education and Teaching International, 62*(1), 216–232. https://doi.org/10.1080/14703297.2023.2271493

Algers, A., & Ljung, M. (2015). Peer reviewing of OER in a contested domain – An activity theoretical analysis. *Journal of Interactive Online Learning, 13*(4), 21–38. https://www.ncolr.org/issues/jiol/v13/n4/2.html

Axe, J., Childs, E., DeVries, I., & Webster, K. (2020). Student experiences of open educational practices: A systematic literature review. *Journal of e-Learning and Knowledge Society, 16*(4), 67–75. https://doi.org/10.20368/1971-8829/1135340

Baas, M., Admiraal, W., & van den Berg, E. (2019). Teachers' adoption of open educational resources in higher education. *Journal of Interactive Media in Education, 2019*(1), Article 9. https://doi.org/10.5334/jime.510

Bozkurt, A. (2024). GenAI et al.: Cocreation, Authorship, Ownership, Academic Ethics and Integrity in a Time of Generative AI. *Open Praxis, 16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654.

Camilleri, A. F., Ehlers, U.-D., & Pawlowski, J. (2014). *State of the art review of quality issues related to open educational resources (OER)* (EUR 26347 EN). Publications Office of the European Union. https://doi.org/10.2791/80171

Cho, K. W., & Permzadian, V. (2024). The impact of open educational resources on student achievement: A meta-analysis. *International Journal of Educational Research, 126*, 102365. https://doi.org/10.1016/j.ijer.2024.102365

Clinton-Lisell, V. (2021). *Open pedagogy: A systematic review of empirical findings. Journal of Learning for Development, 8*(2), 255–268. https://doi.org/10.56059/jl4d.v8i2.511

Colvard, N. B., Watson, C. E., & Park, H. (2018). The impact of open educational resources on various student success metrics. *International Journal of Teaching and Learning in Higher Education, 30*(2), 262–276. https://www.isetl.org/ijtlhe/pdf/IJTLHE3386.pdf

Evelyn, S., & Kromer, J. (2018). OER evaluation as a means of teaching information literacy in individual and small group settings. *The Reference Librarian, 59*(1), 1–9. https://doi.org/10.1080/02763877.2017.1402730

Farrow, R., Iniesto, F., Pitt, B., Weller, M., & Bossu, C. (2024). Innovation with open educational resources: An integrative review of drivers, barriers and enablers. *Journal of Open, Distance, and Digital Education, 1*(2), 1–41. https://doi.org/10.25619/y3yte265

Fitzgerald, M. A., & Byers, A. (2002). A rubric for selecting inquiry-based activities. *Science Scope, 26*(1), 22–25. https://www.researchgate.net/publication/267156453

Haughey, M., & Muirhead, B. (2005). Evaluating learning objects for schools. *Australasian Journal of Educational Technology, 21*(2), 226–241. https://files.eric.ed.gov/fulltext/EJ850358.pdf

Hilton, J. (2020). Open educational resources, student efficacy, and user perceptions: A synthesis of research published between 2015 and 2018. *Educational Technology Research and Development, 68*, 853–876. https://doi.org/10.1007/s11423-019-09700-4

Iniesto, F., & Bossu, C. (2023). Equity, diversity, and inclusion in open education: A systematic literature review. *Distance Education, 44*(4), 694–711. https://doi.org/10.1080/01587919.2023.2267472

Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory (Vol. 2, pp. 215–239).* Lawrence Erlbaum. https://www.taylorfrancis.com/chapters/edit/10.4324/9781410603784-12

Jung, I., Sasaki, T., & Latchem, C. (2016). A framework for assessing fitness for purpose in open educational resources. *International Review of Research in Open and Distributed Learning, 17*(4), 66–84. https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-016-0002-5

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach.* Springer. https://doi.org/10.1007/978-1-4419-1250-3

Kurilovas, E., Bireniene, V., & Serikoviene, S. (2011). Methodology for evaluating quality and reusability of learning objects. *Electronic Journal of E-Learning, 9*(1), 39–51. https://files.eric.ed.gov/fulltext/EJ930254.pdf

Lopes, A. P., Soares, F., Softic, S., & Andriichenko, Y. (2025). The art of quality assessment for open educational resources (OER). In M. E. Auer & T. Rüütmann (Eds.), *ICL 2024: The International Conference on Interactive Collaborative Learning (pp. 259–270).* Springer. https://doi.org/10.1007/978-3-031-83523-0_24

Lübben, S., Müskens, W., & Zawacki-Richter, O. (2023). Quality of OER: Test theoretical development and validation of an assessment tool. In D. Otto, G. Scharnberg, M. Kerres, & O. Zawacki-Richter (Eds.), Distributed learning ecosystems: Concepts, resources, and repositories (pp. 139-160). Springer VS. https://doi.org/10.1007/978-3-658-38703-7

Mayrberger, K., Zawacki-Richter, O., & Müskens, W. (2018). Qualitätsentwicklung von OER – Vorschlag zur Erstellung eines Qualitätssicherungsinstrumentes für OER am Beispiel der

Hamburg Open Online University. Hamburg: Universität Hamburg. https://doi.org/10.25592/978.3.924330.67.5

McCullagh, P. (1980). Regression models for ordinal data. Journal of the Royal Statistical Society: Series B (Methodological), 42(2), 109–142. https://www.jstor.org/stable/2984952

Meng, X., Cui, C., & Wang, X. (2020). Looking back before we move forward: A systematic review of research on open educational resources. In *Proceedings of the 20th International Conference on Educational Innovation through Technology (EITT) (pp. 20–28).* IEEE. https://doi.org/10.1109/EITT50754.2020.00022

MERLOT. (n.d.). *MERLOT peer review information.* Multimedia Educational Resource for Learning and Online Teaching (MERLOT). https://info.merlot.org/merlothelp/assets/docs/Peer_Review_form_CC_0519.pdf

Martin, T., & Kimmons, R. (2020). Faculty members' lived experiences with choosing open educational resources. *Open Praxis, 12*(1), 131–144. https://doi.org/10.5944/openpraxis.12.1.987

Mishra, M., Dash, M. K., Sudarsan, D., Santos, C. A. G., Mishra, S. K., Kar, D., Bhat, I. A., Panda, B. K., Sethy, M., & da Silva, R. M. (2022). Assessment of trend and current pattern of open educational resources: A bibliometric analysis. *The Journal of Academic Librarianship, 48* (3), 102520. https://doi.org/10.1016/j.acalib.2022.102520

Mishra, S. (2025). A review of reviews on open educational resources. *Open Praxis,* 17(2), 305–325. https://doi.org/10.55982/openpraxis.17.2.851

Müskens, W., Zawacki-Richter, O., & Dolch, C. (2022, March). Instrument zur Qualitätssicherung von OER – IQOER: Entwicklungsversion 17 [Unpublished manuscript]. ResearchGate. https://doi.org/10.13140/RG.2.2.16987.03363/1

Nesbit, J. C., Belfer, K., & Vargo, J. (2009). A convergent participation model for evaluation of learning objects. *Canadian Journal of Learning and Technology, 31*(3), 1–13. https://doi.org/10.21432/T25C8C

Oliveira, M. M. de, Paschoal, L. N., & Barbosa, E. F. (2021). Quality models and quality attributes for open educational resources: A systematic mapping. In *IEEE Frontiers in Education Conference (FIE) (pp. 1–10)*. IEEE. https://doi.org/10.1109/FIE49875.2021.9637309

Otto, D., Schroeder, N., Diekmann, D., & Sander, P. (2021). Trends and gaps in empirical research on open educational resources (OER): A systematic mapping of the literature from 2015 to 2019. *Contemporary Educational Technology, 13*(4), ep325. https://doi.org/10.30935/cedtech/11145

Pérez-Mateo, M., Maina, M. F., Guitert, M., & Romero, M. (2011). Learner generated content: Quality criteria in online collaborative learning. *The International Review of Research in Open and Distributed Learning, 12*(5), 125–148. https://openaccess.uoc.edu/items/a3b8df75-34a2-4845-abc4-8dc479201dea?locale=en

Ramírez-Montoya, M. S. (2020). Challenges for open education with educational innovation: A systematic literature review. *Sustainability, 12*(17), 7053. https://doi.org/10.3390/su12177053

Stracke, C. M., Downes, S., Conole, G., Burgos, D., & Nascimbeni, F. (2019). Are MOOCs open educational resources? A literature review on history, definitions and typologies of OER and MOOCs. *Open Praxis, 11*(4), 331–341. https://doi.org/10.5944/openpraxis.11.4.1010

Tietjen, P., & Asino, T. I. (2021). What is open pedagogy? Identifying commonalities. *The International Review of Research in Open and Distributed Learning, 22*(2), 186–204. https://doi.org/10.19173/irrodl.v22i2.5161

Tlili, A., Altinay, F., Huang, R., Altinay, Z., Olivier, J., Mishra, S., Jemni, M., & Burgos, D. (2022). Are we there yet? A systematic literature review of open educational resources in Africa: A combined content and bibliometric analysis. *PLOS ONE, 17*(1), e0262615. https://doi.org/10.1371/journal.pone.0262615

Tlili, A., Garzón, J., Salha, S., Huang, R. Xu, L. Burgos, D., Denden, M. Farrell, O., Farrow, R. Bozkurt, A., Amiel, T. McGreal, R. Lopez-Serrano, A. & Wiley, D. (2023). Are open educational resources (OER) and practices (OEP) effective in improving learning achievement? A meta-analysis and research synthesis. *International Journal of Educational Technology in Higher Education 20*(54). https://doi.org/10.1186/s41239-023-00424-3

Tseng, H., Sciuchetti, M., & Tang, Y. (2024). Exploratory study of students' creative mindset, perceived values of open educational resources, and flow. *Open Praxis, 16*(4), 595–609. https://doi.org/10.55982/openpraxis.16.4.747

UNESCO. (2019). *Recommendation on open educational resources (OER).* United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000370936

UNESCO. (2023). *Implementation of standard-setting instruments, part VI, 2019 recommendation on open educational resources (OER) – Consolidated Report on the Implementation of the Recommendation.* https://unesdoc.unesco.org/ark:/48223/pf0000386595

Versantvoort, M., & Schuwer, R. (2023). Towards sustainable OER practices: The case of bachelor nursing in the Netherlands. *Open Praxis, 15*(2), 113–123. https://doi.org/10.55982/openpraxis.15.2.542

Yuan, L., & Recker, M. (2015). Not all rubrics are equal: A review of rubrics for evaluating the quality of open educational resources. *International Review of Research in Open and Distributed Learning, 16*(5), 16–38. https://doi.org/10.19173/irrodl.v16i5.2389

Zawacki-Richter, O., & Mayrberger, K. (2017). *Qualität von OER – Internationale Bestandsaufnahme von Instrumenten zur Qualitätssicherung von OER*. HOOU-Projektbericht. Hamburg Open Online University. https://epub.sub.uni-hamburg.de/epub/volltexte/2018/76394/pdf/qualitaet_von_oer_2017_1.pdf

Author's Contribution (CRediT)

WM: Investigation, methodology, visualization, writing – original draft, writing – review & editing. The author has read and agreed to the published version of the manuscript

Data Availability

The instruments mentioned in this article were published in German under a Creative Commons licence (see Müskens et al., 2022). The English version of the IQOER Rating Scales is fully included in this publication. The original dataset is available upon request from the author.

Ethics and Consent

All participants in this study provided informed and explicit consent for the collection, storage, and processing of their personal data. Personal data were stored separately from the participants'

evaluations and were handled in accordance with Article 13 of the German Data Protection Regulation (Datenschutz-Grundverordnung, DSGVO), which implements the European General Data Protection Regulation (GDPR). Participants were informed of their rights under this regulation. A procedural protocol was submitted for inclusion in the register of processing activities at the University of Oldenburg. No sensitive personal data were collected beyond the demographic information reported in this article.

Competing Interests

The authors declare no competing interest.

Acknowledgement of Use of Generative AI Tools

Based on Academic Integrity and Transparency in AI-assisted Research and Specification Framework (Bozkurt, 2024): In the creation of this scholarly product, the authors benefited from GAI related to following areas: Language Translation and Localization. For the translation and localization of content, ChatGPT v4.0 was employed. Human translators subsequently reviewed and adjusted the translations to ensure accuracy, cultural appropriateness, and contextual relevance. The final text was thoroughly reviewed and approved by the authors to ensure it accurately reflects the intended research outcomes and ethical standards. The authors also assessed and addressed potential biases inherent in the AI-generated content. The final version of the paper is the sole responsibility of the human authors.

Acknowledgement

The author would like to thank all participants who contributed to the validation of the IQOER Rating Scales by evaluating Open Educational Resources. This study was funded by the German Federal Ministry of Education and Research (BMBF) as part of the project „Verbundprojekt: Digitale Bildungsarchitekturen – Offene Lernressourcen in verteilten Systemen“ (Collaborative Project: Digital Educational Architectures – Open Learning Resources in Distributed Systems).