

Article

# Development of an ESD Indicator for Teacher Training and the National Monitoring for ESD Implementation in Germany

Eva-Maria Waltner <sup>1,\*</sup> , Werner Rieß <sup>1</sup>  and Antje Brock <sup>2,\*</sup>

<sup>1</sup> Department of Biology and Pedagogy of Biology at the University of Education, Kunzenweg 21, 79117 Freiburg, Germany; riess@ph-freiburg.de

<sup>2</sup> Department of Education and Psychology, Freie Universität Berlin, Fabeckstraße 37, 14195 Berlin, Germany

\* Correspondence: eva-maria.waltner@ph-freiburg.de (E.-M.W.); brock@institutfutur.de (A.B.); Tel.: +49-761-682-903 (E.-M.W.)

Received: 15 May 2018; Accepted: 16 July 2018; Published: 18 July 2018



**Abstract:** Education for Sustainable Development (ESD) is a core element of UNESCO's Sustainable Development Goal (SDGs) Target 4.7, which seeks to ensure that all learners acquire the knowledge and skills needed to promote sustainable development through education for sustainable development. The German Federal Ministry for Education and Research (BMBF) followed suit in 2015 and launched a high scale national monitoring of the current state of ESD implementation. In this context, suitable ESD indicators should be analyzed to inform policy and research agendas. The present project is part of the national monitoring within Germany's Global Action Program (GAP) actions. The research team at the University of Education in Freiburg conducted a study to evaluate the accessibility of data and the measurability of ESD-relevant teacher training (TT). During the two-step procedure for data collection on ESD-relevant TTs in Germany, an extensive list of ESD related search terms first captured 66,935 TTs with possible ESD relevance in the evaluation period. Second, the collected data was analyzed using Mayring's qualitative content analysis. The 66,935 TTs were thereby reduced to 3818 TTs with different degrees of ESD relevance. The results of the evaluation study show that suitable ESD indicators, the FESD (formula for the ESD-indicator for TTs) (basic), FESD (basic, rated) and FESD (pro), could be developed and calculated for 15 of 16 federal states in Germany. The gathered insights show a path towards ESD monitoring in TT to clarify the needs and achievements of ESD implementation in the field of continuing education of teachers. However, the presented indicators only show a possible path for ESD indicator development. A comprehensive set of ESD indicators should also focus on the micro or output (e.g., ESD competencies) level. These insights for the future seem worth striving for not only in Germany or on the national level but also internationally to foster ESD, Target 4.7 of the SDGs and the SDGs in general.

**Keywords:** education for sustainable development; ESD; in-service teacher training; indicator; implementation; national monitoring; education for sustainability; future education; global action program (GAP) on ESD; sustainable development goals; SDGs

## 1. Introduction

ESD is placed at the center of the 2030 Sustainable Development Agenda and has been widely recognized as a key enabler of sustainable development and an integral element of quality education [1] (p. 4). This agenda is an equally important focus of European and transnational endeavors connected to the SDGs and ESD. These national endeavors, processes, and analyses are not only 'globally' relevant,

they contribute equally to a general understanding of the challenges, needs, and success criteria of the ESD implementation process in general (e.g., [2]). Accordingly, we present the project located in Germany, ESD\_indicator\_teacher training, placed in the broader concept of the national monitoring actions of ESD, hoping that the project contributes to a broad transnational understanding of ESD implementation efforts, measurement possibilities, and remaining challenges. In 2015, the German Federal Ministry for Education and Research (BMBF), established a multi-level implementation structure to realize the GAP on ESD. It includes a steering body (the national platform), six expert forums (early childhood education, school, vocational education and training, higher education, informal and non-formal learning/youth and local authorities), a youth forum, and 10 partner networks consisting of practitioners in the different thematic and educational fields of ESD (such as early childhood education and higher education but also biodiversity or media). The main bodies of the structure—national platform, expert forums as well as partner networks—are designed to facilitate the information flow between top-down and bottom-up hierarchies and within the vertical structures of the complex system of actors involved in ESD governance in Germany (for ESD governance, see [3,4]) and to connect (inter)national and sometimes abstract political aims with concrete educational practices. No central or national German educational policy exists. Given the importance of the federal states concerning education in Germany, representatives of the single federal states are included in almost all the expert forums. The implementation process is informed by an international advisor at the German UNESCO Commission and a scientific advisor at the *Institut Futur*, located at the Freie Universität Berlin. The department of the scientific advisor is also responsible for the wide-scale national monitoring of the current state of ESD implementation with the aim of informing the stakeholders involved in the GAP structure, the scientific community, and the public about the status quo of ESD in Germany and effective measures for upscaling it. This implementation structure of the GAP addresses what Læssøe and Mochizuki [5] have identified as missing elements in the actual governance efforts related to ESD, i.e., a lack of concrete national action plans, curricula frameworks and guidelines, fragmentation, a lack of coordinating bodies, a lack of systematic monitoring and evaluation of ESD policy efforts, as well as what actually happens in practice (see [5], p. 38). Consequently, a consensus has formed among stakeholders for the need of ESD operationalization and a major desideratum to develop quantitative indicators (see [6], p. 85). The aim to develop indicators for ESD has been explicitly voiced since the beginning of the UN Decade for ESD and first attempts for its operationalization have been suggested (see, e.g., [7–9]). However, many tasks on how to generate a set of ESD indicators that are specific, measurable, attainable, realistic, and timely [10] (p. 35) remain incomplete. Moreover, preceding analyses on ESD indicators have mostly drawn on either qualitative interviews or state government reports [8,9,11]. Additionally, these studies mainly had a macro focus or relied on self-report surveys with dichotomous scale (yes/no) formats (e.g., see [12–15]).

This article seeks to contribute to the emerging research agenda on how to measure ESD implementation. Agreeing on standardized ways of operationalizing ESD is crucial, given that the operationalization can be understood as being “influenced by international as well as unique socio-political and cultural contexts which make them [...] more or less floating concepts rather than consistent and fixed concepts” [5] (p.32). This statement mirrors the level of operationalization that is seen as an important characteristic of ESD, that is, standardization should address ‘glocality’. When operationalizing ESD, the definition must incorporate a large-scale or international comparability as well as aspects that reflect national priorities or other local specifics are important (see [9], p. 100). On the content-level, ESD is linked to an idea of universal concepts and human rights applicable to everyone, but at the same time concrete links to the local circumstances (and lifeworlds) should also be taken into account (see [16], p. 28). Although discussions about the openness or vagueness of the concept are ongoing, the increasing research and monitoring efforts of ESD (indicators) provide a chance to not only enhance the visibility of the concept but to bring about a new level of concreteness by means of indicators. In that context and on an international level, the observation is that “the lack of common visions for SD and ESD hampers the definition and the selection of common patterns for

ESDI” (ESD Indicators, the authors) ([9], p. 102). Clearly, a certain level of openness is essential for ESD implementation, given the participatory demand of the concept and its applicability to cultures and local contexts (see, e.g., [16]). Thus, the increasing efforts to monitor and evaluate ESD must carefully address preserving the necessary contingency of the concept, while concurrently effectively contributing to upscale it by operationalizing the core ideas of where and what ESD should be.

Against this background, the University of Education in Freiburg conducted an evaluation study to analyze accessibility of data and the measurability of ESD in the specific field of in-service TTs accredited by the states. The databases for TT therefore equally include extracurricular TTs from the non-formal sector or TTs cooperation between the states and NGOs (if accredited by the states). The 16 federal states are each responsible and, to a large extent, sovereign in terms of educational policy, which is why the TTs are also listed in non-centralized databases of each federal state. Through collecting primary data, the project, ESD\_indicator\_teacher training, aimed to answer the following two research questions. RQ 1: How accessible is the relevant data on TTs in the 16 German federal states for continuous monitoring? RQ 2: Is it possible to develop a replicable and applicable ESD indicator for teacher training (as part of a larger, longitudinal set of ESD indicators)?

To situate these questions in the broader framework of the ESD implementation process in Germany, Section 2 provides a brief status of the national ESD monitoring process, and then the theoretical implications for (ESD) indicators (and TT in ESD) are discussed. Section 3 describes the method, the development of several potential indicators for ESD-relevance in TTs, and the analysis. Section 4 contains the findings and the proposed indicators: FESD (basic), FESD (basic, rated) and FESD (pro). Section 5 concludes the study. Section 6 provides an outlook and the results of the indicator development and their implications are discussed.

## 2. Formal and Conceptual Context of the Study: National Monitoring of ESD Implementation

### 2.1. Broader Framework of National Monitoring for ESD

National as well as international trends of quantifying and measuring can be observed in the context of education [17,18], critical [19], for example, the visibility and influence of the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS).

The same trend applies to the context of sustainability-related research, for example, the indicator development of the Sustainable Development Goals (SDGs) [20] and the German Sustainable Development Strategy [21] (pp. 146–147). This trend also expands to the field of ESD, e.g., [22,23]. These tendencies can be framed and explained in the context of the growing importance of evidence-based policy making, which deepens cooperation between research and policy making and the role of research outcomes for political decision making and practice [24–27]. The described trends contribute to the growing demand of monitoring efforts in education, which can be understood as an indicator-based, systematic observation of different aspects of an educational system ([28], p. 163). This focus on measurability underscores the crucial role of the development of adequate indicators, especially for complex phenomena such as educational processes, particularly for ESD or sustainability in general [13,29–33].

The implementation of the UNESCO GAP on ESD in Germany is monitored with the aim of providing robust information about the extent and quality of the ESD implementation. The complexity of ESD monitoring in six different educational areas mirroring the expert forums (early childhood education, school, vocational education and training, higher education, informal and non-formal learning, and local authorities) provides ample opportunity for scientific advice on effective national and regional strategies for further upscaling ESD in Germany. Additionally, the monitoring enables the crosscutting approach to become a vital part of the structures underlying educational systems.

The general monitoring process stretches over a period of three years and comprises four research phases. During the first phase, the main aim was to provide an overview over the use of different

ESD-related concepts (for more details, see [34]) within different groups of central documents for five out of the total of six educational areas (excluding informal learning because of a lack of comparable documents in this field of education). The encompassing monitoring process reflects not only the need for a differentiated view, according to the educational areas, but also to balance and complement the strengths and drawbacks of the various methodological approaches of grasping ESD.

The research in the different monitoring phases was based on national [7,8] and international indicator propositions, for example, on the suggestions of UNECE (see, e.g., [11]), which include information about whether, for example, the proposition refers to ESD in national policy documents, in curricula of formal education, or in teacher training [35–38]. The indicators for that broader area of national monitoring have been partly adapted to fit the specifics of the six educational fields (for the set of indicators see: [34] (p. 4–5)). This monitoring has been set up as an encompassing process not only to fulfill the specifics of the different educational areas but also to go beyond a merely descriptive approach and to also include explanatory aims to better grasp the diffusion process of ESD. This research at the science–policy interface, therefore, was conducted to provide an overview of the status quo of ESD in Germany and to provide recommendations of how to accelerate the uptake of ESD at the structural level of educational systems. The results of the first monitoring phases show that a great need still exists for advancing the indicators on ESD to address the different educational areas and concurrently base them on suitable datasets (e.g., that show a certain dynamic and are meaningful for ESD).

The national ESD monitoring in Germany is the broader context in which the TTs indicators described in this article have been developed. The study on TTs was commissioned by the scientific advisor of the GAP as a part of the national monitoring of ESD in Germany. The underlying theoretical and methodological orientations of the evaluation study for the ESD\_indicator\_teacher training will be described in the following section.

## 2.2. Theoretical Implications for the Indicator Development

For the present project, we used the indicator definition given by the authoring group for educational reporting in Germany. This definition nonetheless also comprises the purpose of indicators in a global perspective. In their report on the development of indicators for educational reporting, they defined indicators as quantitative tools (or proxy variables) providing a simple and comprehensible status report on the quality or the state of art of a more complex, usually multi-dimensional system [39] (p. 15). This definition, however, fails to capture another important aspect of indicators—their use as a tool or policy instrument to show trends and development tendencies (e.g., [40], p. 234). Thus, indicators “are considered to provide condensed information that can be transformed into knowledge relevant for decision making” [41], which is why they are equally used in educational monitoring to assure quality education or other relevant policy fields, e.g., [9,14,42–49].

Important preconditions of indicators are that they need to be seen as comprehensive, credible, and fair in the eyes of their recipients (see [50] pp. 23–24). In developing the method of this study, we placed special emphasis on the target of developing an indicator that is “based on adequate samples, [and has] appropriate levels of reliability, good validity, and above all, positive reactivity” ([51], p. 3). Other important aspects are the criteria of applicability and acceptance. To ensure that these criteria are addressed, many stakeholders, including all spokespersons for ESD and TTs from the 16 federal ministries of educations, have been included in the process of developing the national ESD indicators. Given that the discourse on how to operationalize and define ESD is still an ongoing debate in research, all stakeholders need to be committed to a high level of transparency from the outset in the development the indicators.

Compared to the ongoing discussions on different ESD definitions, a consensus exists on the essential role of teachers in the ESD implementation (see e.g., [52], p. 172) and that teacher professional development can be an effective way to enhance and ensure the quality of schooling [53–55].

Additionally, Lipsky's classical oeuvre on implementation research drew attention to the distinctive role of teachers (called street ministers of education) as policy makers on the ground (see [56], p. 12). This conceptualization of policy makers is equally important in the context of indicators relevant for policy making, discussed later in Sections 5 and 6. In line with this statement is the assumption of the essential role of the teacher as change agent in ESD development. However, ESD research investigating the state of the art and effects of ESD, especially in the first phase of teacher education is limited (e.g., see [57–59]). Few studies have conducted research on in-service teacher education courses related to ESD, conceptualized TT for ESD, or developed assessment tools, see e.g., [15,60]. ESD research needs to follow the example of other research disciplines in education, especially the research field of vocational education and training (i.e., continuing education of teachers). Considering these aspects, Lipowsky's [61] and Huber's [62] findings on the importance and success criteria of TT served as a general theoretical background for the presented study.

In the context of ESD, the role of the teacher is equally stressed by the UNESCO Roadmap that states that “[e]ducators and trainers [as] powerful agents of change for delivering the educational response to sustainable development. However, for them to help usher in the transition to a sustainable society, they must first acquire the necessary knowledge, skills, attitudes and values” [63] (p. 20). As described above, research on ESD (implementation) still lacks data and quantitative methods for operationalization. In the future, it would be desirable to develop indicators or measurement instruments that assess the more complex construct of ESD, for example, competencies or the above-mentioned dimensions of knowledge, skills, attitudes, and values.

The research on indicators of the ESD implementation process offers many possibilities for mapping the progress of ESD with TT as one important field. In this analysis, quantitative data was gathered to examine multiple aspects of ESD-relevant TT offered at the federal state level.

### 3. Method and Data for ESD-Relevant TTs

#### 3.1. Implications of Method for Indicator Development

The German Institute for International Educational Research (DIPF) describes two possibilities for indicator development. One approach is a deductive top-down procedure in which indicators are derived from scientific modelling. The other approach is an inductive bottom-up procedure that includes the selection and aggregation of relevant factors extracted from existing data (see [44], p. 13). The chosen indicator development approach for the described project combines the two aspects. We employed statistical modelling to analyze ESD indicators for TT (FESD), and we analyzed the existing datasets from the federal states' databases that list the TTs. Section 3.2 describes the first step of the data retrieval. Subsequently, the coding process of the raw data will be presented in Section 3.3.

#### 3.2. Data Collection

A two-step procedure was chosen for data collection on ESD-relevant TT. In reference to other ESD-monitoring projects, we worked with the given key words (see [34], p. 7) and further extended the list to incorporate the complex implications of the ESD concept. Note that this first step was applied only to ensure that no TT containing ESD-relevant content would be overlooked. The output of the key word retrieval was scanned in a second step to ascertain the degree that the TT courses were ESD-relevant (see Section 3.3). The list of keywords was sent to ESD spokespersons and those responsible for TT in the 16 federal states. We received positive feedback on the keyword list from all 32 stakeholders of the ministries and the TT centers.

Consequently, the databases of the 16 federal states of Germany were scanned with the help of the comprehensive list of 37 search terms connected to multidimensional aspects of ESD (“*esd\**”OR“*sustain\**”OR“*sustain\**” (“*sustain\**” appears twice because this list constitutes a translation from the German keyword list in which the English “*sustain\**” word was used as well). OR“*develop\**”OR“*sdg\**”OR“*world\**”OR“*environment\**” (The German keyword “*welt\**” = *world*,

captures “Umwelt” = *environment* as well. This is the reason why the English translation of the keyword list contains one additional key word (“*environment\**”). OR”global\*”OR”consum\*”OR”fair\*”OR”justice\*” OR”peace\*”OR”system\*”OR”social\*”OR”intercultural\*”OR”intergeneration\*”OR”particip\*”OR”gestaltungs\*” OR”efficien\*”OR”sufficien\*”OR”consisten\*”OR”perma\*”OR”natur\*”OR”species\*”OR”biodivers\*”OR”eco\*” OR”climate\*”OR”greenhouse\*”OR”mobility\*”OR”econom\*”OR”energ\*”OR”resources\*”OR”material circ\*” OR”seasonal\*”OR”nutrition\*”OR”futur\*”OR”transformation\*”OR”change\*”)(own translation, original accessible under Supplementary Materials, Table S1).

These retrieved TTs were further analyzed in the second content analytical step, which was a major part of the analysis. If the search string results exceeded 1500 TT courses. 1500 TT courses were chosen randomly from the list extracted by the key word retrieval and coded following the usual procedure. This procedure was repeated for four federal states (Bavaria, Berlin, Lower Saxony and North Rhine-Westphalia). Step two of the data analysis is described in Section 3.3. It defines the qualitative approach that was applied to investigate the degrees of ESD relevance in the TTs. The databases of the federal states contained all in-service TTs accredited by the states. As stated above, the list includes extracurricular TTs from the non-formal sector or TT cooperation between the states and NGOs. Altogether, 111,589 TTs were listed in the databases of the federal states during the evaluation period that included the school year 2015/16 and half of the school year 2016/17. This evaluation period was chosen, because TTs were still being added to the databases for the second half of the school year 2016/17 by the time the study began. Therefore, the data collection was limited to 1.5 years during the time in which all TT events were listed in the databases. The extensive list of search terms captured 66,935 TTs with ESD relevance to be further analyzed in the second data analysis described below.

### 3.3. Data Analysis—Coding ESD-Relevant TT Using a Weighted Ordinal Coding System

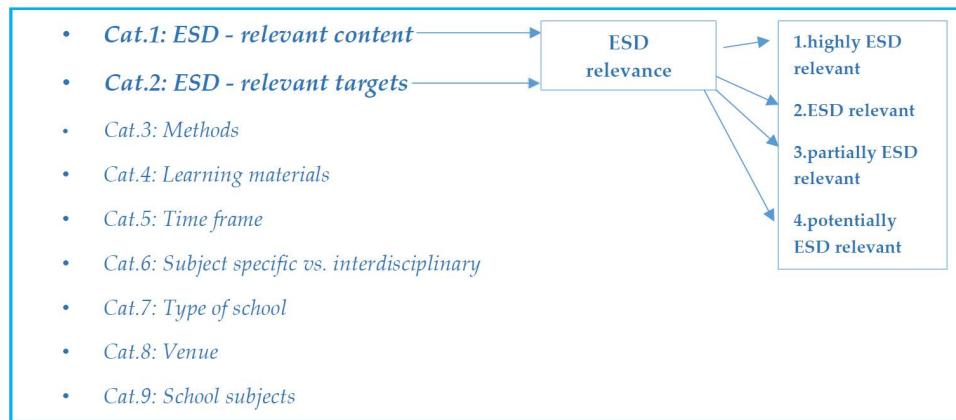
The collected data consisting of more than half of the TT courses that were offered during the evaluation phase (66,935 out of 111,589) was analyzed using Mayring’s qualitative content analysis [64]. The name of this method might be misleading as it involves systematic quantitative steps of analysis. There are several advantages in using this methodological approach, because it allows a rule-bound procedure, a theory-guided character and it integrates quantitative steps of analysis (see [64], pp. 39–42).

We formulated strict content-analytical rules for the scanning process of ESD relevance, measured by defined categories. Figure 1 shows a translated extract of the coding agenda. Only the highest value of Category 1 is displayed. (original accessible under Supplementary Materials, Document S2).

Variable	Value	Definition	Anchor samples	Encoding rules
Category 1: ESD-relevant content/topics	Cat: high ESD relevance	<p>The contents mentioned in the teacher training course (TT) are topics of the sustainability discourse or are considered from a sustainability point of view and these form the focal point of the TT.</p> <p>à R1: reference to:</p> <ul style="list-style-type: none"> <li>- Sustainable Development / Agenda 21</li> <li>- Sustainability awareness</li> <li>- Sustainable development in school</li> <li>- central sustainability strategies (e.g., efficiency, sufficiency, permanence, consistency)</li> <li>...</li> </ul> <p>à R2: reference to:</p> <ul style="list-style-type: none"> <li>- Education for Sustainable Development (ESD)</li> <li>- Global learning</li> <li>...</li> </ul> <p>à R3: examples of multidimensional topics:</p> <ul style="list-style-type: none"> <li>- Reasons and consequences of biodiversity loss</li> <li>- Fairtrade</li> <li>- Sustainable mobility</li> <li>- Freshwater shortages and their economic / social consequences</li> <li>- Reasons and consequences of climate change</li> <li>- Resources and basic human needs</li> <li>- Economic globalization and global development disparities / poverty</li> <li>- Cultural diversity and social inequality</li> <li>- Contributions of natural resources management for environmental protection</li> <li>- Ecological friendly production and consumption patterns</li> <li>...</li> </ul>	<p>R1: The main idea NE is explicitly mentioned in the analysis unit (the entire tender text for a teacher training with title, description, other information), or it forms the main focus of the analysis unit.</p> <p>OR</p> <p>R2: The topic is explicitly assigned to the ESD, so it should be considered from a sustainability point of view; or</p> <p>OR</p> <p>R3: The multidimensional (ecological, economic, socio-cultural, intergenerational dimension) importance of the subject of the training clearly emerges from the unit of analysis / inter-dimensional relationships are made clear, for example, social-cultural and ecological aspects or economic and socio-cultural aspects are addressed together;</p> <p>OR</p> <p>R4: A theme of the ecological, social or economic dimension and the associated additional normative weighting: 1) Protection of the environment (not just the environment) Promote intergenerational justice ...; 2) Fairer society (not just society); 3) Sustainable economy (not just economics).</p>	

**Figure 1.** Translated extract of the coding agenda for the second step of the TT analysis.

The coding agenda with nine categories (see Figure 2) was validated and tested by four external ESD experts. It contained the name of the variable, their values (on an ordinal grading scale), the definition, anchor samples, and the coding rules.



**Figure 2.** Categories of the coding procedure for analyzing the TT course descriptions with nine categories validated and tested by external ESD experts.

The category system is the central point in quali-/quantitative content analysis. The coding agenda consisted of nine categories and 20 coding rules. The gradations were designated using several categories of the above mentioned comprehensive coding system obtained from scanning the descriptions of teacher formation courses, retrieved by the above given list of keywords. As stated above, wide acceptance and understanding are the basis of a successful indicator development. For this reason, transparency and the consultation of other ESD experts were considered as essential criteria in the development process of the ESD category system. To assure validity of the coding agenda, it was tested, in an expert hearing, by four external ESD experts from different universities and different disciplines and backgrounds. They provided feedback and recommendations for the coding agenda. Furthermore, the stakeholders from the ministries were equally consulted. To ensure validity of the coding agenda in addition to the states and academic stakeholders, we integrated ESD topics listed by the teachers from a former study and used them as anchor samples [65]. In a second validation step, three young practitioners (teacher students with practical school experience) tested and evaluated the coding agenda (in addition to the ESD academics and ministerial experts.)

Afterwards, the final coding procedure was given to different coders who applied the content analytic coding rules on 235 TT courses. The intercoder reliabilities were good, ranging from 0.71 to 0.93 (Cohen's Kappa  $\tau$  correlation coefficient)—“For content analysis it is inter-coder reliability which is of particular significance. Several content analysts work on the same material independently from one another and their findings are compared.” (see [64], p. 42). That means that different people using our coding plan for the same TT course they came to the same or similar results when designating the categories. In the data analysis phase of the evaluation project, the data gathered by the described procedure was analyzed and interpreted using MAXQDA software. To determine ESD relevance, we did not rely on a binary decision. Instead, we used the above-mentioned gradation system ranging on an ordinal scale from highly to potentially ESD-relevant (see Figure 2). The coding procedure consisted of two main categories, Category 1: ESD-relevant content, and Category 2: ESD-relevant targets. If the texts of the TTs were not sufficient to qualify for the two categories, coders coded them with missing values. This procedure addressed coding errors caused by coders' uncertainty.

Further categories (Category 3–9, see Figure 2) were coded when the unit of analysis was at least partially ESD relevant according to Category 1 or 2 and when the information given in the TT course description was sufficient.

Figure 2 displays the coding process conversion into the rating scale. Given that the accessibilities of information on categories 3–9 varied greatly across the different states, we built our rating scale based only on the first two categories. These two categories parallel with the structure of other international ESD-indicator grids, for example, the grid of indicators and sub-indicators from the UNECE adapted by Capelo et al. [9], (p. 107). We equally used “key themes of SD” as our first category. Instead of a binary ‘yes/no’ self-reporting format, we used strict coding rules to assign the category values that also contained a graduation system from potentially ESD-relevant to highly relevant. Based on the coding (using an ordinal scale) of the first two categories, we classified the TT courses as highly ESD relevant if they, according to the strict coding rules, were coded as containing (1) a high ESD-relevant content (see Figure 1, extract of the coding agenda) and (2) they additionally mentioned an ESD-relevant goal (see Category 2). Instead of focusing on learning outcomes (see [9], p. 107) as it is expressed in the UNECE documents, we analyzed whether the TT courses stated at least any ESD-relevant goals and coded them according to the strict content analytic rules for Category 2 ESD-relevant targets. This procedure was used, because for the present evaluation study we could not yet utilize other appropriate assessment tools that capture ESD competencies or the expected learning outcomes. Additional measurement instruments would be needed to integrate further outcome indicators that capture, for example, SD-related competencies or the dimensions of knowledge, skills, attitudes, values, and behavior.

## 4. Results

### 4.1. Total Numbers and Proportions for Germany—The Indicator FESD (Basic)

Primary data on ESD-relevant TT could be gathered for 15 of 16 federal states in Germany. Only 0.6% of the 66,935 TTs (see Section 3.2) were coded highly ESD relevant. All federal states totaled 3818 TTs with different degrees of ESD relevance. For the international audience, we mainly focused on the results for Germany. However, the ministries and ESD spokespersons from the 15 federal states received the exact numbers and indicators for each of their specific states in the form of a report. TTs received the designation, highly ESD relevant, if they contained at least ESD-relevant content and ESD-relevant targets. Table 1 shows the findings for the 15 states analyzed in Germany. The numbers are not yet multiplied by the rating coefficient, as explained below.

**Table 1.** Total numbers for Germany (15 of 16 federal states)<sup>1</sup> in the evaluation period—school year 2015/16 and half of the school year 2016/17.

TT <sub>total</sub> = total number of TTs accredited by the state(s)	111,589			
N = number of teachers who worked in Germany during the evaluation period	574,893 <sup>2</sup>			
	highly ESD-relevant TT (rated by y4 = 1)	ESD-relevant TT (rated by y3 = 0.75)	partially ESD-relevant TT (rated by y2 = 0.5)	potentially ESD-relevant TT rated by y1 = 0.25)
TT <sub>ESD</sub> = ESD-relevant TT accredited by the state(s)	385	987	938	1508
				3818

<sup>1</sup> Without data for TT in Baden-Wuerttemberg, numbers are not yet weighted; <sup>2</sup> Total numbers of teachers from the Federal Statistical Office (*Destatis*) [66] without Baden-Wuerttemberg.

Note that potentially ESD-relevant TT is not a placeholder for coding uncertainty. Potentially ESD-relevant means that the identified TT dealt with topics that could be conceivable starting points for teaching ESD or SD topics, but the course description had no EDS-relevant goals nor explicit text stating ESD or SD as course content. Examples for potentially ESD-relevant TT courses: (1) A TT course that included waste management but no further discussion about the environmental protection dimension; (2) A TT course about the political or economic system (without discussing any other sustainability dimensions); (3) A TT course only discussing integration (without the cross-social or global context).

If the texts of the TTs were not sufficient to decide whether a TT course contained ESD-relevant topics or goals (i.e., the coding Categories 1 and 2), the courses were coded with missing values.

Based on the gathered data, we developed several indicator possibilities, the *FESD (basic)*, the *FESD (basic, rated)*, and the *FESD (pro)* (1)–(3). *FESD (basic)* (1) shows the basic indicator for ESD-relevant TT

$$\begin{aligned} FESD \text{ (basic)} &= \frac{TT(ESD)}{TT(\text{total})} \times 100 \\ TT_{ESD} &= \text{ESD-relevant TTs accredited by the state(s)} \\ TT_{\text{total}} &= \text{total number of TT accredited by the state(s)} \end{aligned} \quad (1)$$

*FESD (basic)* is the proportion of ESD-relevant TTs to the total number of TTs offered by the states, calculated for 15 of the 16 federal states. Afterwards, the indicator was also calculated for the entire Federal Republic of Germany. The *FESD (basic)* for each federal state varied between 1.08 and 9.01 and was 3.42 for Germany. Eight of the federal states obtained indicators below this ratio and seven were above. These findings clearly show a high variance in ESD-relevant TTs offers among the 15 federal states.

#### 4.2. Rated ESD Relevance—*FESD (Basic, Rated)*

Table 1 also shows the applied ratings  $y4$ ,  $y3$ ,  $y2$ , and  $y1$  for highly to potentially ESD-relevant TTs. The TTs coded highly to potentially ESD-relevant were rated using quartile weights based on the ESD relevance (see Table 1) as shown in the Formula (2) below. We used this ratio procedure to stress those TTs that explicitly mentioned ESD-relevant content or goals (multiplied by  $y4 = 1$ ) in comparison to those that only contained a possible ESD-relevant topic, that is, without calling it an ESD topic or explicitly stressing one of the sustainability dimensions. Additionally, as explained earlier in Section 3.3, these courses did not contain any ESD-goals. The potentially ESD-relevant TTs were multiplied by  $y1 = 0.25$ .

If the texts of the TTs were not sufficient to decide whether a TT course contained ESD-relevant topics or goals (i.e., coding Categories 1 and 2) the courses were coded with missing values.

$$FESD \text{ (basic, rated)} = \sum_{i=1}^4 (TT(ESD) \times yi) \times \frac{100}{TT(\text{total})} \quad (2)$$

Calculations of the *FESD (basic, rated)* weighted values are similar to the *FESD (basic)* values. The results for the federal states varied between 0.63 and 4.20. Six federal states achieved indicator results that were above the ratio for Germany, *FESD (basic, rated, Germany)* = 1.77. Nine of the federal states were below this ratio.

The clear advantage of the *FESD (basic, rated)* lies in the gradation ratings instead of a simple binary ESD-relevant/non-ESD-relevant assessments. The underlying assumption of the rating system is that a highly ESD-relevant TT course will promote the development of teachers' ESD competencies to a greater extent than a TT that only marginally addresses ESD topics. However, further research is needed on the effectiveness of ESD TTs and on factors that contribute to successful developmental outcomes to substantiate this assumption. Further implications and interpretation of these results are discussed in Sections 4.4, 5 and 6.

#### 4.3. *FESD(pro)—Qualitative Rating System for ESD Relevance with a Quantitative Aspect*

The Formula (3) above displays the third indicator possibility, the *FESD (pro)*. It combines a qualitative rating system for ESD relevance with a quantitative aspect. The preceding *FESD (basic, rated)* indicator includes the qualitative aspect of ordinal rating ratios applied according to the results of the coding process. The quantitative aspect consists of the integration of the ratio for the numbers of teachers that theoretically could attend a highly ESD-relevant TT course during the evaluation phase. Therefore, the *FESD (pro)* calculates the degree and frequency of professionalization possibilities, that is, the highly ESD-relevant TT courses in relation to the total number of teachers in the country.

The indicator answers the question: What is the theoretical possibility teachers had in the observation period to effectively develop their ESD competencies (by attending a highly ESD-relevant TT)? *FESD (pro)* varied between 0.01 and 0.19 among the 15 states. The *FESD (pro, min)* of 0.01, for example, indicates that every 100th teacher could attend a highly ESD-relevant TT. In contrast, the *FESD (pro)* of 0.19 signifies that every fifth teacher could attend a highly ESD-relevant TT. The national mean for Germany was 0.07, that is, every 14th teacher could attend a highly ESD-relevant TT during the evaluation phase. According to calculations based on the weighted total numbers for the *FESD (pro)*, six federal states were below the German indicator result of 0.07. The nine remaining states reached the 0.07 or above this indicator ratio.

$$\begin{aligned} FESD (\text{pro}) &= \sum_{i=1}^4 (TT(\text{ESD}) \times yi) \times \frac{20}{N(\text{total})} \\ N_{\text{total}} &= \text{Total Number of teachers that worked in each federal state} \\ 20 &= \text{Average Number of participants (teachers) per TT} \end{aligned} \quad (3)$$

#### 4.4. Summary of the Results

By the end of this project, the states' coding results were applied in different formulas to calculate the indicators for ESD relevance in TTs. The results of the evaluation study showed that suitable ESD indicators, the *FESD (basic)*, *FESD (basic, rated)*, and *FESD (pro)*, could be developed and assessed. The calculated indicators constitute a major evaluation that covers the entire German TT system (except one federal state), the proportion of ESD-relevant TTs for 15 of the 16 federal states, and distinct dimensions and characteristics of ESD relevance in TTs.

The results in Table 2 show the high variance among the 15 federal states. However, note that the purpose of these indicator results is not to determine strength or weakness by comparing indicators between the states (intercomparisons) but to use these indicators to make intracomparisons of longitudinal results.

**Table 2.** Indicator results for the FESD (basic), FESD (basic, rated), and FESD (pro) in Germany <sup>1</sup>.

	FESD (Basic)	FESD (Basic, Rated)	FESD (Pro)
Federal states' min.	1.08	0.63	0.01
Federal states' max.	9.01	4.20	0.19
National indicator ratio (for Germany)	3.42	1.77	0.07

<sup>1</sup> Only the min. and max. indicator and the national mean for Germany are displayed. We decided not to display the ratios by state rankings, because each state should set their own FESD-indicator targets independent of other states.

## 5. Discussion

This article presented the indicator development process for the FESDs and gave several possible indicators for capturing ESD relevance in TT. We discussed the underlining guidelines for indicators and their development. Thereupon, we presented the methods and the data as well as the three indicator possibilities—*FESD (basic)*, *FESD (basic, rated)*, and *FESD (pro)*—for practical application. The main contribution and strength of the study is that the results provide the possibility to assess a transparent indicator ratio that provides condensed information on ESD-relevant TTs, based on a dynamic dataset (i.e., the state databases of the TT course listings). The presented findings of indicator development constitute the most recent research on the ESD relevance in TTs.

Against this background, the FESD can be understood as a robust proxy indicator that can quickly enable stakeholders dealing with ESD or TT to evaluate how the ESD implementation evolves over time. Recalling the research questions of the evaluation study, RQ 1: How accessible is the relevant data on TTs in the 16 German federal states for continuous monitoring? RQ 2: Is it possible to develop a replicable and applicable ESD indicator for teacher training (as part of a larger, longitudinal set of ESD indicators)? We can state the following: (1) The relevant data was accessible in 15 of 16 federal states

in Germany in the time of the evaluation phase. (2) These databases could be used for continuous monitoring. (3) A replicable and applicable ESD indicator (the FESDs variants) could be developed.

Nevertheless, some shortcomings of the indicator development and application occurred. As stated above, no central database exists listing all the TT courses for the federal states in Germany. Consequently, unavoidable differences were encountered in the characteristics of databases from the federal states (e.g., differences in the length of the description and the amount of information on the training). These differences should be considered when comparing the results at the federal state level. Another consequence of the differences among the TT course descriptions was that Categories 3–9 (see Figure 2) could not be coded in many cases because of insufficient course descriptions. Hence, only Categories 1–2 were used as a basis of information for calculating the FESD indicator. Structural changes of the different federal state databases (to facilitate access to relevant information) would improve future research on TTs. For example, course attendance records of teachers would be a very important component for all the federal states.

Some shortcomings are project specific and thus mirror general limitations of single indicators. The proposed indicators—*FESD (basic)*, *FESD (basic, rated)*, and *FESD (pro)*—constitute only partial indicators that focus on the field of ESD in TT. Ideally, however, they should function as embedded in a more comprehensive set of ESD indicators. Only such a set of ESD indicators would be able to capture the interactions, multiple facets, and developmental levels to provide a more holistic view of the state of art of ESD in Germany (see e.g., [40] p. 234). Another limitation of the proposed indicators relates to a general problem associated with data collection in field studies that lack experimental control, especially in the social sciences. The use of indicators (or other research methods) to capture a multi-dimensional and highly complex system always entails a considerable reduction of this complexity of the examined system (see [39], p. 15). Additionally, further ESD-specific research is needed on the effectiveness of ESD TT and on factors that contribute to successful developmental outcomes of TTs. For a general overview, see e.g., [53,67–69].

The study also assessed detailed characteristics of the TTs (e.g., methods, duration, learning material, subject specific vs. interdisciplinary). Although the evaluation study took place in Germany, the approaches for the indicator development at the meso and micro level is nevertheless transferable to transnational application. As stated above, one of the main characteristics of the indicators is that they provide an easily understood index to report on a more complex domain that is relevant for decision making. Accordingly, the FESDs offer insights on the status of ESD implementation in TTs. Hence, the main aim of these indicators is that they provide relevant and reliable measures to inform decision makers about the status of ESD development in their state (s. [65], p. 8), also for long-term perspectives. Each administrative region should therefore decide where they want to set their targets or benchmarks. However, this aim is only one aspect of how ESD indicators can be used. In fact, they provide information relevant for different levels, such as institutional and political, and various stakeholders on several levels (e.g., curriculum developers, teacher educators, teachers, school leaders, and researchers; [9], p. 95).

For the political level, this study has revealed a suggestion for applicable and replicable indicators that provide condensed, relevant, and reliable data that can be used for decision making [41] (p. 510). In general, federal agencies “view indicators as essential for monitoring the status of the nation’s educational system and for tracking how it changes over time. [...] policy makers also seek indicator data that can inform new improvement efforts” ([70], p. 181). These two statements apply to the *FESD* indicators and ESD indicators in a broader context. As already stated above, subsequent to the *FESD* indicator results, each federal state should decide on their own targets, which are determined and influenced by many factors.

Moreover, in line with Lipsky, teachers, as policy makers on the ground can also refer to these indicator results when asking for more or better TT courses that incorporate ESD. Additionally, teachers should be given more ownership in the development of indicators in the future (see [71], p. 71). However, in developing a comprehensive set of ESD indicators for portraying the ESD implementation,

this article constitutes only a starting point of suggesting a possible path for ESD monitoring in the field of TT. Further implications for research and practice and an outlook will be given in the concluding section of this article.

## 6. Conclusions & Outlook

The ESD indicator on TT, as part of a larger set of ESD indicators, can serve to promote “comparison of performance, monitoring and data collection” [72] and foster benchmarking processes and political target setting. It can therefore be seen as an important driver in promoting the structural implementation of ESD, while addressing the following questions: “How do stakeholders know that an increased understanding of ESD and indicators develops? How can stakeholders work productively to build ownership and commitment? and Are stakeholders improving their ability to assess ESD progress?” ([10], p. 32). In sum, the monitoring phases of ESD implementation are well complemented by the indicators on TT. They address the need for dynamic information by which changes can be observed in relatively short time periods and can be compared on different scales. One of the central features of the national monitoring is to grasp the level of ESD implementation for each of the educational fields and thereby contribute to a broader picture of ESD. Without the development of powerful indicators for single educational fields, this overview would remain too blurred at a macro or meso level of analysis of ESD implementation [73]. The shift to the micro level at which ESD effects are measured on outcomes, such as SD-related knowledge, skills, attitudes, values, and behavior, is still needed. However, the complementary advantage of the broader scope of the national monitoring and the TTs indicator also lies in the analysis of documents with relatively long update cycles of up to 10 years, as in the case of school curricula versus the dynamic change in the database for TTs. Another complementary element is the analysis across educational fields, which focused on the conceptual level and therefore will not be able to capture the breadth on the thematic level of ESD, whereas the TT indicator is the reverse such that the whole conceptual range and educational areas of ESD might not be covered. It is therefore insightful to view the project findings in the context of the teacher-relevant outcomes examined in the document analysis of the ESD national monitoring see [34]. The findings of the teacher education documents underscore the importance for further integration into TT, and they suggest that a lack of ESD-relevant knowledge is not just the concern of an older generation of teachers but also teachers who have recently finished their education. The findings also stress the relevance of TT as a valuable continuation and complement of teacher education. The concentration of references to ESD on single subjects in teacher education highlights the need for TT on ESD that addresses teachers in general and all subjects if it is to be effectively implemented as a cross-cutting educational concept.

Future research on ESD indicators should also focus on not only the input and processes of indicators related to ESD, but also on its outputs (see also the definition for monitoring education systems [28], p. 163). Although the indicators for TT proposed in this article, depending on the perspective, have elements of input, process, and output in the developed indicators (for differentiations of indicators, see [14]), a complement within a set of indicators that targets the level of teacher and student ESD competencies [74–80] appears worth striving for. As stated above, this shift to the micro level at which ESD effects are measured on outcomes—such as SD-related knowledge, skills, attitudes, values, and behavior—is still needed.

**Supplementary Materials:** The following documents are available online at <http://www.mdpi.com/2071-1050/10/7/2508/s1>, Table S1: List of keywords for ESD data retrieval in original language (German); Document S2: Coding procedure in original language (German).

**Author Contributions:** W.R. was leading the research group at the University of Education in Freiburg. E.-M.W. was part of the research group. She was responsible for the corresponding contact, data collection, and reporting. Antje Brock contributed by placing the subproject into the broader framework of the national monitoring and categorization. E.-M.W., W.R., and A.B. wrote the paper.

**Acknowledgments:** The authors gratefully acknowledge the Federal Ministry for Education and Research (BMBF) for funding and the Institut Futur, located at the Freie Universität Berlin for supporting this project (2016-Futur-01).

The article processing charge was funded by the Baden-Württemberg Ministry of Science, Research and Culture and the University of Education, Freiburg in the funding program Open Access Publishing. A big thank you to the associates from the University of Education: Dietmar Glaubitz and Jutta Nikel for their feedback and guidance during the project and to the assistants, Sarah Jaster, Lorena Mutter, and Frederic Petersen for coding. Last but not least, we would like to thank Inka Bormann, Matthias Barth, Hansjörg Seybold and Stephan Schuler and all the other stakeholders for their contributions to the coding agenda by giving valuable feedback in the expert hearings.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

BMBF	German Federal Ministry for Education and Research
ESD	Education for Sustainable Development
FESD	Formula for the ESD-indicator for TTs
GAP	Global Action Program
PISA	Program for International Student Assessment
TIMSS	Trends in International Mathematics and Science Study
TT	In-service teacher training accredited by the states
UNECE	United Nations Economic Commission for Europe
UNESCO	United Nations Educational, Scientific and Cultural Organization
SDGs	Sustainable Development Goals

## References

1. Leicht, A.; Heiss, J.; Byun, W.J. *Issues and Trends in Education for Sustainable Development*; UNESCO: Paris, France, 2018; Available online: <http://unesdoc.unesco.org/images/0026/002614/261445e.pdf> (accessed on 8 March 2018).
2. UNECE (United Nations Economic Commission for Europe). *Learning from Each Other, the UNECE Strategy for Education for Sustainable Development*; UNECE: Geneva, Switzerland, 2009; Available online: [http://www.unece.org/fileadmin/DAM/env/esd/01\\_Typo3site/LearningFromEachOther.pdf](http://www.unece.org/fileadmin/DAM/env/esd/01_Typo3site/LearningFromEachOther.pdf) (accessed on 6 January 2017).
3. Bormann, I.; Heinrich, M.; Hamborg, S.; Lambrecht, M.; Nikel, J.; Haker, C.; Brüsemeister, T. Governance von Transferprozessen im Mehrebenensystem. Gegenstandsbezogene und methodologische Überlegungen. In *Governance-Regime des Transfers von Bildung für Nachhaltige Entwicklung: Qualitative Rekonstruktionen*; Bormann, I., Hamborg, S., Heinrich, M., Eds.; Springer VS: Wiesbaden, Germany, 2016; pp. 7–41.
4. Kolleck, N.; Jörgens, H.; Well, M. Levels of Governance in Policy Innovation Cycles in Community Education: The Cases of Education for Sustainable Development and Climate Change Education. *Sustainability* **2017**, *9*, 1966. [[CrossRef](#)]
5. Læssøe, J.; Mochizuki, Y. Recent Trends in National Policy on Education for Sustainable Development and Climate Change Education. *J. Educ. Sustain.* **2015**, *9*, 27–43. [[CrossRef](#)]
6. Adomßent, M.; Bormann, I.; Burandt, S.; Fischbach, R.; Michelsen, G. Indikatoren für Bildung für nachhaltige Entwicklung. In *Bildung für Nachhaltige Entwicklung: Beiträge der Bildungsforschung*; Bundesministerium für Bildung und Forschung (BMBF), Ed.; Bundesministerium für Bildung und Forschung (BMBF) Referat Bildungsforschung: Bonn, Germany, 2012; pp. 71–90.
7. Rode, H. *Different Indicators for Different Contexts? Developing Indicators for ESD in Germany*; University of Bath: Bath, UK, 2006; Available online: <http://www.bath.ac.uk/cree/resources/esrcesd/rode.pdf> (accessed on 11 July 2018).
8. Siemer, S.; Elmer, S.; Rammel, C. Pilotstudie zu Indikatoren Einer Bildung für Nachhaltige Entwicklung. Available online: [http://afnk.de/wp-content/uploads/2012/10/Exkurse\\_2\\_Pilotstudie\\_Indikatoren.pdf](http://afnk.de/wp-content/uploads/2012/10/Exkurse_2_Pilotstudie_Indikatoren.pdf) (accessed on 5 January 2017).
9. Tilbury, D.; Janousek, S. *Development of a National Approach to Monitoring, Assessment and Reporting on the Decade of Education for Sustainable Development, Stage 1: Identification of National Indicators. Summarising Documented Experiences on the Development of ESD Indicators and Networking with Expert Groups on ESD Indicators*; Australian Research Institute of Education for Sustainability and Australian Government Department of the Environment and Water Resources: Sydney, Australia, 2006. Available online: [http://aries.mq.edu.au/projects/esdIndicators/files/ESDIndicators\\_Feb07.pdf](http://aries.mq.edu.au/projects/esdIndicators/files/ESDIndicators_Feb07.pdf) (accessed on 4 May 2018).

10. Capelo, A.; Santos, C.; Pedrosa, M.A. Education for sustainable development in East Timor. *Asian Educ. Dev. Stud.* **2014**, *3*, 98–117. [[CrossRef](#)]
11. Blum, N.; Nazir, J.; Breiting, S.; Goh, K.C.; Pedretti, E. Balancing the tensions and meeting the conceptual challenges of education for sustainable development and climate change. *Environ. Educ. Res.* **2013**, *19*, 206–217. [[CrossRef](#)]
12. Capelo, A.; Conceição Santos, M.; Pedrosa, M.A. Chapter 5: Education for Sustainable Development Indicators, Competences and Science Education. In *Contributions to the UN Decade of Education for Sustainable Development*; Goncalves, F.J., Pereira, R., Leal Filho, W., Eds.; Lang Peter GmbH Internationaler Verlag der Wissenschaften: Frankfurt, Germany, 2012.
13. Haan, G.D. Gestaltungskompetenz als Kompetenzkonzept für Bildung für nachhaltige Entwicklung. In *Kompetenzen der Bildung für Nachhaltige Entwicklung: Operationalisierung, Messung, Rahmenbedingungen, Befunde*; Bormann, I., Haan, G.D., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2008; pp. 23–42.
14. Lingard, B. The impact of research on education policy in an era of evidence-based policy. *Crit. Stud. Educ.* **2013**, *54*, 113–131. [[CrossRef](#)]
15. Organisation for Economic Co-operation and Development (OECD). Evidence in Education, Linking Research and Policy 2007. Available online: <http://www.oecd.org/education/ceri/47435459.pdf> (accessed on 27 March 2018).
16. Biesta, G. Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* **2009**, *21*, 33–46. [[CrossRef](#)]
17. United Nations. Final List of Proposed Sustainable Development Goal Indicators, Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2016/2/Rev.1). Available online: <https://sustainabledevelopment.un.org/content/documents/11803Official-List-of-Proposed-SDG-Indicators.pdf> (accessed on 9 April 2018).
18. The Federal Government of Germany. Deutsche Nachhaltigkeitsstrategie, Neuauflage 2016. Available online: [https://www.bundesregierung.de/Content/DE/\\_Anlagen/Nachhaltigkeit-wiederhergestellt/2017-01-11-nachhaltigkeitsstrategie.pdf?\\_\\_blob=publicationFile&v=20](https://www.bundesregierung.de/Content/DE/_Anlagen/Nachhaltigkeit-wiederhergestellt/2017-01-11-nachhaltigkeitsstrategie.pdf?__blob=publicationFile&v=20) (accessed on 27 March 2018).
19. McKenzie, M.; Bieler, A.; McNeil, R. Education policy mobility: Reimagining sustainability in neoliberal times. *Environ. Educ. Res.* **2015**, *21*, 319–337. [[CrossRef](#)]
20. Pauw, J.; Gericke, N.; Olsson, D.; Berglund, T. The Effectiveness of Education for Sustainable Development. *Sustainability* **2015**, *7*, 15693–15717. [[CrossRef](#)]
21. Oliver, K.; Lorenc, T.; Innvar, S. New directions in evidence-based policy research: A critical analysis of the literature. *Health Res. Policy Syst.* **2014**, *12*, 34. [[CrossRef](#)] [[PubMed](#)]
22. Smith, P.C.; Nutley, S.M.; Davies, H.T.O. *What Works? Evidence Based Policy and Practice in Public Services*, 1st ed.; The Policy Press: Bristol, UK, 2000.
23. Nutley, S.M.; Davies, H.T.O.; Walter, I. Evidence Based Policy and Practice: Cross Sector Lessons from the United Kingdom. *Soc. Policy J. N. Z.* **2003**, *2*, 29–48. Available online: <https://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/journals-and-magazines/social-policy-journal/spj20/20-pages29-48.pdf> (accessed on 9 April 2018).
24. OECD (Organization for Economic Co-operation and Development). *Evidence in Education: Linking Research and Policy*; OECD Publishing: Paris, France, 2007.
25. Ioannidou, A. Educational Monitoring and Reporting as Governance Instruments for Evidence-Based Education Policy. In *International Educational Governance*, 1st ed.; Amos, K.S., Ed.; Emerald Group Publishing: Bingley, UK, 2010; pp. 155–172.
26. McCool, S.F.; Stankey, G.H. Indicators of sustainability: Challenges and opportunities at the interface of science and policy. *Environ. Manag.* **2004**, *33*, 294–305. [[CrossRef](#)] [[PubMed](#)]
27. Arima, A.; Konaré, A.O.; Lindberg, C.; Rockefeller, S. Draft International Implementation Scheme, United Nations Decade of Education for Sustainable Development 2005–2014. Available online: <http://www.env-edu.gr/Documents/files/Basika%20Keimena/DESD.pdf> (accessed on 4 June 2017).
28. Bormann, I. Criteria and indicators as negotiated knowledge and the challenge of transfer. *Educ. Res. Policy Pract.* **2007**, *6*, 1–14. [[CrossRef](#)]
29. Burford, G.; Tamás, P.; Harder, M. Can We Improve Indicator Design for Complex Sustainable Development Goals? A Comparison of a Values-Based and Conventional Approach. *Sustainability* **2016**, *8*, 861. [[CrossRef](#)]

30. Meadows, D. Indicators and Information Systems for Sustainable Development, a Report to the Balaton Group. Available online: <https://pdfs.semanticscholar.org/3372/06350e14a75581b88550fadfd0b39d144d87.pdf> (accessed on 12 June 2017).
31. Singer-Brodowski, M.; Brock, A.; Etzkorn, N.; Otte, I. Monitoring of education for sustainable development in Germany—Insights from early childhood education, school and higher education. *Environ. Educ. Res.* **2018**, *1*–16. [CrossRef]
32. Di Giulio, A.; Schweizer Ruesch, C.; Adomßent, M.; Blaser, M.; Bormann, I.; Burandt, S.; Fischbach, R.; Kaufmann-Hayoz, R.; Kirkser, T.; Künzli David, C.; et al. Bildung auf dem Weg zur Nachhaltigkeit: Vorschlag eines Indikatoren-Sets zur Beurteilung von Bildung für Nachhaltige Entwicklung. Available online: [http://www.ikaoe.unibe.ch/publikationen/PDF-Schriftenreihen/Schriftenreihe%2012%20%282011%29/BNE-Indikatoren\\_2011\\_AOe\\_Nr12.pdf](http://www.ikaoe.unibe.ch/publikationen/PDF-Schriftenreihen/Schriftenreihe%2012%20%282011%29/BNE-Indikatoren_2011_AOe_Nr12.pdf) (accessed on 27 March 2018).
33. Michelsen, G.; Adomßent, M.; Bormann, I.; Burandt, S.; Fischbach, R. Indikatoren der Bildung für Nachhaltige Entwicklung. Ein Werkstattbericht. Available online: [http://www.bne-portal.de/sites/default/files/Indikatoren\\_2520der\\_2520BNE.File\\_0.pdf](http://www.bne-portal.de/sites/default/files/Indikatoren_2520der_2520BNE.File_0.pdf) (accessed on 27 March 2018).
34. UNECE (United Nations Economic Commission for Europe). *Indicators for Education for Sustainable Development, Addendum—Draft Format for Reporting on Implementation of the UNECE Strategy for Education for Sustainable Development*; UNECE: Geneva, Switzerland, 2006.
35. UNECE-Expert Group on Indicators for ESD. Extract of Issues Relevant to Competences in ESD from the Reporting Format. Indicators/Sub-Indicators under Issues for Reporting 2 and 3. Available online: [http://www.unece.org/fileadmin/DAM/env/esd/inf.meeting.docs/EGonInd/8mtg/ExtractRFCompetencesEG\\_ESD\\_8\\_5.pdf](http://www.unece.org/fileadmin/DAM/env/esd/inf.meeting.docs/EGonInd/8mtg/ExtractRFCompetencesEG_ESD_8_5.pdf) (accessed on 3 June 2017).
36. UNECE (United Nations Economic Commission for Europe). *2014 National Implementation Reporting, Phase III: Format for Reporting on the Implementation of the UNECE Strategy for Education for Sustainable Development*; UNECE: Geneva, Switzerland, 2014; Available online: <http://www.unece.org/2014esdreporting.html> (accessed on 11 July 2018).
37. UNECE (United Nations Economic Commission for Europe). *Indicators for Education for Sustainable Development, Progress Report on the Work of the Expert Group*. ECE/CEP/AC.13/2006/5; UNECE: Geneva, Switzerland, 2006; Available online: <https://www.unece.org/fileadmin/DAM/env/documents/2006/ece/cep/ac.13/ece.cep.ac.13.2006.5.e.pdf> (accessed on 9 April 2018).
38. Baethge, M.; Brunke, J.; Döbert, H.; Fest, M.; Freitag, H.-W.; Fitzsch, B.; Fuchs-Rechlin, K.; Christian, K.; Kühne, S. Indikatorenentwicklung für den Nationalen Bildungsbericht “Bildung in Deutschland”, Grundlagen, Ergebnisse, Perspektiven. Available online: [https://www.bmbf.de/pub/Bildungsforschung\\_Band\\_33.pdf](https://www.bmbf.de/pub/Bildungsforschung_Band_33.pdf) (accessed on 12 January 2017).
39. Nikel, J.; Müller, S. Indikatoren einer Bildung für nachhaltige Entwicklung. In *Kompetenzen der Bildung für Nachhaltige Entwicklung: Operationalisierung, Messung, Rahmenbedingungen, Befunde*; Bormann, I., Haan, G.D., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2008; pp. 233–251.
40. Bormann, I.; Michelsen, G. The Collaborative Production of Meaningful Measure(ment)s: Preliminary insights into a work in progress. *Eur. Educ. Res. J.* **2010**, *9*, 510–518. [CrossRef]
41. Berger-Schmitt, R.; Noll, H.-H. Conceptual Framework and Structure of a European System of Social Indicators. Available online: [http://www.gesis.org/fileadmin/upload/dienstleistung/daten/soz\\_indikatoren/eusi/paper9.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/soz_indikatoren/eusi/paper9.pdf) (accessed on 12 January 2017).
42. Brock, A. Indikatorenset zur Verankerung von BNE in den verschiedenen Bildungsbereichen. In *Wegmarken zur Transformation: Nationales Monitoring von Bildung für Nachhaltige Entwicklung in Deutschland*; Brock, A., Haan, G.D., Etzkorn, N., Singer-Brodowski, M., Eds.; Verlag Barbara Budrich: Berlin, Germany; Toronto, ON, Canada, 2018.
43. DIPF (Deutsches Institut für Internationale Pädagogische Forschung). Das weiterentwickelte Indikatorenkonzept der Bildungsberichterstattung. Available online: <http://www.bildungsbericht.de/de/forschungsdesign/pdf-grundlagen/indikatorenkonzept.pdf> (accessed on 26 September 2016).
44. Gehrlein, U. *Nachhaltigkeitsindikatoren zur Steuerung Kommunaler Entwicklung*; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2004.
45. Huckle, J. *A UK Indicator of Education for Sustainable Development, Report on Consultative Workshops*; Sustainable Development Commission: London, UK, 2006; Available online: <https://research-repository.st-andrews.ac.uk/>

- [ac.uk/bitstream/handle/10023/2263/sdc-2006-education-for-sd-indicators.pdf?sequence=2](http://ac.uk/bitstream/handle/10023/2263/sdc-2006-education-for-sd-indicators.pdf?sequence=2) (accessed on 12 June 2017).
- 46. Huckle, J. Consulting the UK ESD community on an ESD indicator to recommend to Government: An insight into the micro-politics of ESD. *Environ. Educ. Res.* **2009**, *15*, 1–15. [CrossRef]
  - 47. Barrett, A.M.; Sørensen, T.B. *Indicators for All? Monitoring Quality and Equity for a Broad and Bold Post-2015 Global Education Agenda*; Open Society Foundations: New York, NY, USA, 2015; Available online: <https://www.opensocietyfoundations.org/sites/default/files/barrett-indicators-for-all-20150520.pdf> (accessed on 27 June 2018).
  - 48. EU-Commission of the European Communities. *Progress Towards the Common Objectives in Education and Training, Indicators and Benchmarks*; EU-Commission of the European Communities: Brussels, Belgium, 2004. Available online: [http://www.nefmi.gov.hu/letolt/eu/progress\\_report\\_indicators.pdf](http://www.nefmi.gov.hu/letolt/eu/progress_report_indicators.pdf) (accessed on 28 June 2018).
  - 49. Van Ackeren, I.; Hovestadt, G. Indikatorisierung der Empfehlungen des Forum Bildung. Available online: <http://d-nb.info/971373620/34> (accessed on 11 January 2017).
  - 50. Fritz-Gibbon, C.T.; Tymms, P. Technical and Ethical Issues in Indicator Systems: Doing Things Right and Doing Wrong Things. Available online: <http://epaa.asu.edu/ojs/article/viewFile/285/411> (accessed on 12 January 2017).
  - 51. Seybold, H. Bedingungen des Engagements von Lehrern für Bildung für nachhaltige Entwicklung. In *Bildung für Nachhaltige Entwicklung: Aktuelle Forschungsfelder Und-Ansätze*; Rieß, W., Apel, H., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2006; pp. 171–183.
  - 52. Yoon, K.S.; Duncan, T.; Lee, S.W.-Y.; Scarloss, B.; Shapley, K.L. Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement. Available online: [https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL\\_2007033.pdf](https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf) (accessed on 21 March 2018).
  - 53. Stoll, L.; Harris, A.; Handscomb, G. Great Professional Development Which Leads to Great Pedagogy, Research and Development Network National Themes: Theme Two. Available online: <https://www.appa.asn.au/wp-content/uploads/2015/08/stoll-article2.pdf> (accessed on 21 March 2018).
  - 54. Lipowsky, F. Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. *Z. Pädagog.* **2006**, *51*, 47–70. Available online: [http://www.pedocs.de/volltexte/2013/7370/pdf/Lipowsky\\_Auf\\_den\\_Lehrer\\_kommt\\_es\\_an.pdf](http://www.pedocs.de/volltexte/2013/7370/pdf/Lipowsky_Auf_den_Lehrer_kommt_es_an.pdf) (accessed on 29 November 2016).
  - 55. Lipsky, M. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*; Russell Sage Foundation: New York, NY, USA, 1980.
  - 56. Andersson, K.; Jagers, S.; Lindskog, A.; Martinsson, J. Learning for the Future? Effects of Education for Sustainable Development (ESD) on Teacher Education Students. *Sustainability* **2013**, *5*, 5135–5152. [CrossRef]
  - 57. Rieckmann, M.; Holz, V. Zum Status Quo der Lehrerbildung und-weiterbildung für nachhaltige Entwicklung in Deutschland. *Pädagog. Blick-Z. Wiss. Prax. Pädagog. Berufen* **2017**, *25*, 4–18.
  - 58. Hallitzky, M. Forschendes und selbstreflexives Lernen im Umgang mit Komplexität. In *Kompetenzen der Bildung für nachhaltige Entwicklung: Operationalisierung, Messung, Rahmenbedingungen, Befunde*; Bormann, I., Haan, G.D., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2008; pp. 159–178.
  - 59. Franz, J.; Frieters, N. Kompetenzmodelle in Fortbildungen—Pragmatische Wege. In *Kompetenzen der Bildung für Nachhaltige Entwicklung: Operationalisierung, Messung, Rahmenbedingungen, Befunde*; Bormann, I., Haan, G.D., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2008; pp. 75–87.
  - 60. Lipowsky, F.; Razek, D. Lehrerinnen und Lehrer als Lernen-Wann gilt der Rollentausch?—Merkmale und Wirkung wirksamer Lehrerfortbildungen. In *Reform der Lehrerbildung in Deutschland, Österreich und der Schweiz: Teil 1: Analysen, Perspektiven und Forschung*; Bosse, D., Criblez, L., Hascher, T., Eds.; Prolog-Verlag: Immenhausen, Germany, 2012; pp. 235–254.
  - 61. Huber, S.G. Wirksamkeit von Fort-und Weiterbildung. In *Lehrprofessionalität: Bedingungen, Genese, Wirkungen und ihre Messung*; Zlatkin-Troitschanskaia, O., Ed.; Beltz: Weinheim, Germany, 2009; pp. 451–463.
  - 62. UNESCO. UNESCO Roadmap for Implementing the Global Action Programme on Education for Sustainable Development. Available online: <http://unesdoc.unesco.org/images/0023/002305/230514e.pdf> (accessed on 5 January, 2017).
  - 63. Mayring, P. Qualitative Content Analysis. Available online: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173> (accessed on 6 January 2017).

64. Rieß, W.; Mischo, C.; Reinholtz, A.; Richter, K.; Dobler, C.; Seybold, H. Evaluationsbericht Bildung für nachhaltige Entwicklung an weiterführenden Schulen in Baden-Württemberg. Maßnahme Lfd. 15 im Aktionsplan Baden-Württemberg. Available online: [https://www.researchgate.net/profile/Werner\\_Riess/publication/278619237\\_Evaluationsbericht\\_Bildung\\_fuer\\_nachhaltige\\_Entwicklung\\_BNE\\_an\\_weiterfuhrenden\\_Schulen\\_in\\_Baden-Wurttemberg\\_Massnahme\\_Lfd\\_15\\_im\\_Aktionsplan\\_Baden-Wurttemberg/links/5583dab808ae89172b85fdbd/Evaluationsbericht-Bildung-fuer-nachhaltige-Entwicklung-BNE-an-weiterfuehrenden-Schulen-in-Baden-Wuerttemberg-Massnahme-Lfd-15-im-Aktionsplan-Baden-Wuerttemberg.pdf](https://www.researchgate.net/profile/Werner_Riess/publication/278619237_Evaluationsbericht_Bildung_fuer_nachhaltige_Entwicklung_BNE_an_weiterfuhrenden_Schulen_in_Baden-Wurttemberg_Massnahme_Lfd_15_im_Aktionsplan_Baden-Wurttemberg/links/5583dab808ae89172b85fdbd/Evaluationsbericht-Bildung-fuer-nachhaltige-Entwicklung-BNE-an-weiterfuehrenden-Schulen-in-Baden-Wuerttemberg-Massnahme-Lfd-15-im-Aktionsplan-Baden-Wuerttemberg.pdf) (accessed on 26 June 2018).
65. Statistisches Bundesamt. Bildung und Kultur: Allgemeinbildende Schulen, Schuljahr 2015/2016. Available online: [https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100167004.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100167004.pdf?__blob=publicationFile) (accessed on 31 May 2017).
66. Razejak, D.; Küsting, J.; Lipowsky, F.; Fischer, E.; Dezhgahi, U.; Reichardt, A. Facetten der Lehrerfortbildungsmotivation-eine faktorenanalytische Betrachtung. *J. Educ. Res. Online (JERO)* **2014**, *6*, 139–159.
67. Lipowsky, F.; Rzejak, D. Lehrerinnen und Lehrer als Lerner—Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. *Reform Lehrerbildung* **2012**, *3*. [CrossRef]
68. Beck, K.; Zlatkin-Troitschanskaia, O. *Lehrerprofessionalität: Was wir Wissen und Was wir Wissen Müssten*; Verl. Empirische Pädagogik: Landau in der Pfalz, Germany, 2010.
69. Oekes, J. What Educational Indicators? The Case for Assessing the School Context. *Educ. Eval. Policy Anal.* **1989**, *11*, 181–199. [CrossRef]
70. Mogensen, F.; Schnack, K. The action competence approach and the ‘new’ discourses of education for sustainable development, competence and quality criteria. *Environ. Educ. Res.* **2010**, *16*, 59–74. [CrossRef]
71. Reid, A.; Nikel, J.; Scott, W. *Indicators for Education for Sustainable Development: A Report on Perspectives, Challenges and Progress*; Anglo-German Foundation: Berlin, Germany, 2006; Available online: [http://www.agf.org.uk/cms/upload/pdfs/CR/2006\\_CR1515\\_e\\_education\\_for\\_sustainable\\_development.pdf](http://www.agf.org.uk/cms/upload/pdfs/CR/2006_CR1515_e_education_for_sustainable_development.pdf) (accessed on 6 January 2017).
72. Tilbury, D.; Janousek, S.; Elias, D.; Bacha, J. Asia-Pacific Guidelines for the Development of National ESD Indicators. Available online: <http://unesdoc.unesco.org/images/0015/001552/155283e.pdf> (accessed on 6 January 2017).
73. Rode, H.; Michelsen, G. Levels of indicator development for education for sustainable development. *Environ. Educ. Res.* **2008**, *14*, 19–33. [CrossRef]
74. Cebrián Bernat, G.; Junyent Pubill, M. Competencias profesionales en Educación para la Sostenibilidad: Un estudio exploratorio de la visión de futuros maestros. *Ensciencias* **2014**, *32*. [CrossRef]
75. Cebrián, G.; Junyent, M. Competencies in Education for Sustainable Development: Exploring the Student Teachers’ Views. *Sustainability* **2015**, *7*, 2768–2786. [CrossRef]
76. Sleurs, W. Competencies for ESD (Education for Sustainable Development) Teachers. A Framework to Integrate ESD in the Curriculum of Teacher Training Institutes. Available online: [https://www.unece.org/fileadmin/DAM/env/esd/inf.meeting\\_docs/EGonInd/8mtg/CSCT%20Handbook\\_Extract.pdf](https://www.unece.org/fileadmin/DAM/env/esd/inf.meeting_docs/EGonInd/8mtg/CSCT%20Handbook_Extract.pdf) (accessed on 27 March 2018).
77. Wiek, A.; Withycombe, L.; Redman, C.L. Key competencies in sustainability: A reference framework for academic program development. *Sustain. Sci.* **2011**, *6*, 203–218. [CrossRef]
78. Shephard, K. Higher education for sustainability: Seeking affective learning outcomes. *Int. J. Sustain. High. Educ.* **2008**, *9*, 87–98. [CrossRef]
79. Hellberg-Rode, G.; Schrüfer, G. Which specific professional action competencies do teachers need in order to implement education for sustainable development in schools? Findings of an exploratory study. *Seiten/Biol. Lehr. Lern.—Z. Didakt. Biol.* **2016**, *1*–29. [CrossRef]
80. UNECE (United Nations Economic Commission for Europe). *Lernen für die Zukunft, Kompetenzen für Bildung für Nachhaltige Entwicklung*; UNECE: Geneva, Switzerland, 2012; Available online: [http://www.education21.ch/sites/default/files/uploads/Lernen%20f%C3%BCr%20die%20Zukunft\\_dt\\_3.pdf](http://www.education21.ch/sites/default/files/uploads/Lernen%20f%C3%BCr%20die%20Zukunft_dt_3.pdf) (accessed on 23 September 2016).

