Development and Validation of the Education for Sustainable Development School Organisation Questionnaire

Dries Verhelst[a]*, Jan Vanhoof[a], and Peter Van Petegem[a]

[a]*Department of Training and Education Sciences, University of Antwerp, Antwerp, Belgium*

*corresponding author (email: dries.verhelst@uantwerpen.be, phone: +32 265 92 52, adress: Universiteit Antwerpen, Departement opleidings- en onderwijswetenschappen, Gratiekapelstraat 10, 2000 Antwerp, Belgium)

Empirically based tools to map education for sustainable development within school organisations are not readily available, which is both a cause and a consequence of the scarce empirical and quantitative research on school organisations and education for sustainable development. In present study, the Education for Sustainable Development School Organisation Questionnaire (ESD-SOQ) was developed and validated. This instrument measures the organisational characteristics of an ESD-effective school. During the development and validation of the ESD-SOQ different steps were taken. First, data from prior research was consulted to draft the initial scales. These were further refined via cognitive interviews (n = 3) and consulting critical peers (n = 7). Next a pilot study (n = 108) was set out to retain the best items per scale and thus reduce the size and load of the questionnaire. Finally, a large-scale validation study (n = 764) ensured the reliability, the content validity, the cognitive validity and the construct validity of the final instrument. Based on the findings of our development and validation study, ESD-SOQ proves to be a valid and reliable instrument for the investigation school organisational characteristics related to education for sustainable development.

Education for Sustainable Development, School Organisations, Questionnaire Development, Questionnaire Validation

1. Problem statement

1.1. Need for empirical and quantitative evidence on ESD schools

The implementation of education for sustainable development (ESD) has been increasing over recent decades fostered by initiatives such as the United Nations Decade of Education for Sustainable Development (DESD, UNESCO, 2014). ESD is an educational approach empowering learners "*with knowledge, skills, values and attitudes to take informed decisions and make responsible actions for environmental integrity, economic viability and a just society*" (UNESCO, 2020, p. 8). Despite these increased efforts, the DESD concluded with a call for the further monitoring and evaluation of ESD (UNESCO, 2014; Wals, 2009). This challenge remains today, with empirical research on ESD still appearing to lag behind the pressing needs. Without ignoring the valuable steps that have been taken towards a more empirical approach in the ESD field, most of the existing research efforts and instruments focus on the student level (e.g. Olsson et al., 2020), the teachers and classroom level (e.g. Varela-Losada et al., 2020) or on higher education, thus overlooking the organisation of compulsory education (e.g. Holm et al., 2015; Rampasso et al., 2019). The fact that measurement instruments to map ESD within school organisations are not readily available is both a cause and a consequence of this lack of empirical and quantitative research on ESD within school organisations.

While school effectiveness research has recognised the importance of the school organisation in facilitating the outcomes and impact of education (Creemers & Kyriakides, 2010; Teddlie & Reynolds, 2006), the gap in the knowledge base of ESD is concerning. Moreover, while existing studies provide valuable insights into the functioning of school organisations implementing or working with ESD, these studies often take a theoretical or qualitative approach and lack an empirical or quantitative perspective (e.g.: Bennell, 2015; Iliško & Badyanova, 2014; Laurie et al., 2016; Scott, 2013). Nonetheless, this is not to say that no empirical studies have been done. For example, Mogren et al. (2019) adapted a conceptual framework for school improvement into a measurement instrument that allowed them to quantitatively investigate ESD implementation in relation to the whole school approach. As a number of scholars have also highlighted the need for more evidence on the impact and outcomes of ESD (Boeve-de Pauw et al., 2015; Bormann & Nikel, 2017; Singer-Brodowski et al., 2019; Waltner et al., 2018), this study set out to develop and validate a questionnaire on school organisational characteristics in relation to ESD.

A frequently expressed concern about effectiveness in the context of ESD relates to bias in effectiveness research; that is, it only measures that which is readily measurable and neglects precisely what lies behind the outcomes (Biesta, 2009). In line with Nikel and Lowe (2010), we argue that

effectiveness does not exclude a pluralistic perspective on quality education. Acknowledging that it is important to be 'aware of outcomes and impacts – both positive and negative – that are not pre-specified' (Nikel & Lowe, 2010, p. 596), a focus on school effectiveness can provide valuable insights into school organisational characteristics that influence both the effectiveness and impact of education. As to what this effectiveness perspective should focus on within the school organisation, Laurie et al. (2016, p. 440) list the adaptation of ESD management practices as one of the requirements for ESD to contribute to quality and effective education. While the need for empirical research on the school organisation level is evident, the tools to conduct this research in the area of ESD are still lacking. To investigate how school organisations can facilitate effective or quality ESD, a validated and comprehensive measurement tool is desirable. Such a measurement instrument can aid in bridging the gap on school effectiveness in the field of ESD and facilitate an empirical perspective on the school organisation's influence on ESD outcomes.

Recognising these needs in the field of ESD, the goal of this study was to develop and validate a questionnaire on ESD school effectiveness, aiming to operationalise a previously developed framework for the ESD-effective school (Verhelst et al., 2020). This framework consists of eight organisational characteristics and is based on a critical review combining existing ESD literature on school organisations with other sources on school effectiveness and educational administration. The theoretical validity of both the framework and the organisational characteristics was substantiated through a qualitative inquiry with school leaders and teachers at ESD active schools (Verhelst et al., 2021). The previous conceptual and qualitative research served as a baseline for the development of the different scales and items on the ESD School Organisational Questionnaire (ESD-SOQ). Following the development of the instrument, we conducted a validation of the ESD-SOQ to check the construct validity and the reliability of the different items and scales (Cohen et al., 2011, pp. 188-189). To address the central goal of this study – the development of a tool enabling the measurement of the constructs within the framework for the ESD-effective school in a reliable and valid way – we posed the following research questions:

- How can the characteristics of the ESD-effective school organisation be operationalised?
- To what extent are the items in the ESD-SOQ valid for measuring the characteristics of the framework for the ESD-effective school organisation?
- How reliable are the different scales in the ESD-SOQ?

The target population for the ESD-SOQ includes all staff members within a school organisation. These may be teachers and school management, as well as policy officers, support staff and others. These

people, who all make and shape the organisation together, are expected to be able to give a good indication of what the school organisation looks like.

1.2. Conceptual framework: the ESD-effective school

In the following, we provide a description of the conceptual framework for an ESD-effective school. The framework for an ESD-effective school holds a specific focus on the organisational level of the school (Verhelst et al., 2020). The organisational level includes all processes and entities connected to organisational functioning of the school that transcend the classroom, individual student, and individual teacher level. Arguably, an ESD-effective school organisation will be able to achieve its ESD-related goals via the facilitation of the processes at the different levels of the school. Different from the school-wide capacity perspective (Sleegers et al., 2014; Thoonen et al., 2012), the present framework is situated at the organisational level of the school whereas the school-wide capacity perspective includes both the school and teacher level. For an in-depth report on this framework, we refer a prior conceptual article (Verhelst et al., 2020). The framework for an ESD-effective school consists of eight characteristics: sustainable leadership, the school resources, pluralistic communication, democratic decision-making, collective efficacy, shared vision, adaptability and supportive relations. As Mogren and Gericke (2019) found, specific aspects (factors) of a school organisation can facilitate ESD within the school via an transmissive (providing structure and anchors) of transformative perspective (focussing on development and processes). The different organisational characteristics will thus facilitate ESD via for example offering anchors via the professional structures (e.g.: workgroups) or having a development focussed perspective via a strong believe in the collective efficacy. Arguably, an ESD-effective school organisation will be able to achieve its ESD-related goals via the facilitation of the processes at the different levels of the school. These processes can include an interdisciplinary approach, project implementation, specific didactical approaches, such as an interdisciplinary or holistic perspective on sustainability issues, or any other process within the school that relates to ESD (Sinakou et al., 2019).

In order to facilitate and enable those ESD-processes an ESD-effective school organisation will need an enabling organisational context. The resources the school has at its disposal and sustainable leadership shape this *sub-contextual level* of the organisation, referring to the internal organisational context which in itself can be situated within the broader educational context. Leadership within an ESD-effective school organisation is described as *sustainable leadership*, indicating leadership that is sustainable over time and that incorporates a holistic and integrated perspective on the past, present and future, locally and beyond (Bottery, 2012; Hargreaves & Fink, 2006). Sustainable leadership will contribute to a long-term holistic strategy, as it adapts to the specific time and context while

considering holistic, pluralistic and action-oriented perspectives. Sustainable leadership is not understood as a one-person show: an ESD-effective school organisation might have several *sustainable leaders*. Moreover, sustainable leadership is ought to have a clear perspective on the main focus of the school, namely teaching and learning (Starratt, 2007), and this in the context of ESD.

Central to the sub-contextual level of an ESD-effective school organisation is the reciprocal relationship between sustainable leadership and the resources the school has at its disposal (Leo & Wickenberg, 2013). These *school resources* fall into three different categories: *time management*, *professional structures* and *physical structures*. Within an ESD-effective school, the available time will be efficiently managed in order to facilitate pluralistic, interdisciplinary and integrated teaching and learning. Moreover, the specific configuration of different professional positions will enable pluralism and holism within the school team (Mogren & Gericke, 2017a). Finally, in an ESD-effective school organisation, the *physical structures* refer to the infrastructure and financial resources of the school organisation (Kuzich et al., 2015; Schelly et al., 2012). These are allocated, used and managed so that the school is able to reach its educational goals while at the same time demonstrating the organisational values related to ESD.
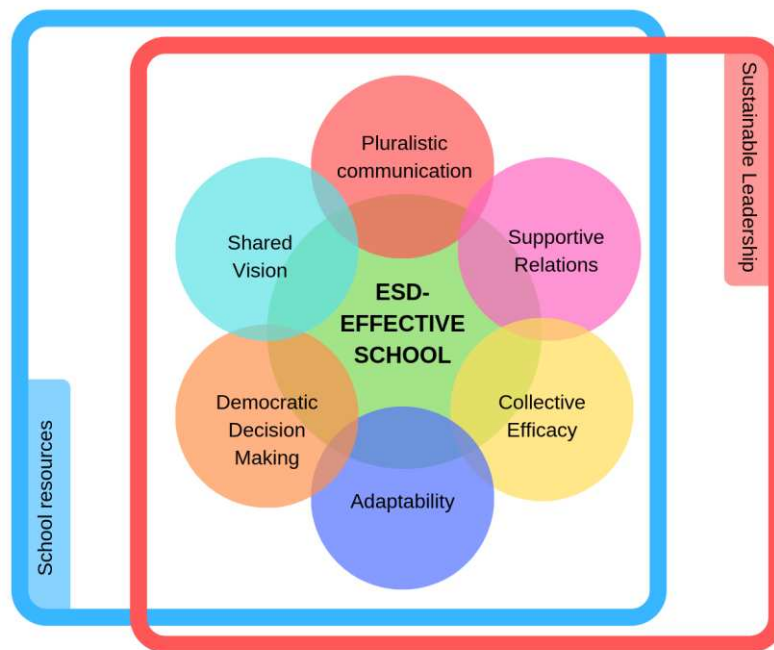

Figure 1: Conceptual framework of an ESD-effective school

In an ESD-effective school organisation, the presence of sustainable leadership and adequate resource management is expected to lead to a situation that facilitates those organisational traits necessary for the school to attain the educational goals related to ESD. The conceptual framework for the ESD-effective school organisation includes six central, interrelated organisational characteristics contributing to ESD effectiveness: *pluralistic communication, democratic decision-making,*

adaptability, supportive relations, collective efficacy and *shared vision. Pluralism* is the recognition of, and dialogue between, different viewpoints and ideas (Boeve-de Pauw et al., 2015; Lijmbach et al., 2002; Rudsberg & Öhman, 2010). As it is central to ESD, it is evident that members of an ESD-effective school organisation will be able to communicate in a pluralistic fashion. Within an ESD-effective school organisation, *pluralistic communication* contributes to a climate where there is a willingness to learn from, and a critical reflection on, the experiences, viewpoints and arguments of others. Pluralistic communication enriches decision-making processes, with different opinions, viewpoints and critical self-reflection contributing to a *democratic decision-making process*. Such a mode of decision-making involves all of the relevant stakeholders in making decisions, leading to broadly supported decisions in an ESD-effective school organisation.

Next, the adaptation strategy of an ESD-effective school organisation is characterised by being able to make well-supported decisions regarding both internal and external demands. Moreover, an effective school knows how to improve itself when responding to internal and external demands (MacBeath & Mortimore, 2001). Thus, *adaptability* is the characteristic of the school linked to its ability to change or not, depending on the situation. By having a holistic perspective when responding to opportunities or demands, the school can ensure that existing aspects are not overlooked in favour of future change (Hargreaves, 2007).

Another central characteristic is the school's *supportive relations,* which may be manifested within the team, with other schools or external partners and between school leaders of different schools (Leo & Wickenberg, 2013). These supportive relations aid in knowledge-sharing and facilitate the establishment of a collective initiative to pursue ESD, both within and between schools. The support of external partners, such as the local community, parents or families, guidance services and external experts, may also aid schools by providing expertise, resources and other benefits contributing to the attainment of the educational goals of the school (Bennell, 2015; Mogren & Gericke, 2017a).

Lastly, a heightened sense of collective efficacy in an ESD-effective school organisation indicates that the school organisation has a shared believe that they can positively influence the learning of the pupils in the school (Bandura, 1997; Hoy et al., 2013). School members feel able to achieve positive outcomes with their students in the area of ESD and, in turn, they positively contribute to the actual student outcomes and school effectiveness. Finally, a *shared vision* in the pursuit of ESD will ensure that everyone within the ESD-effective school organisation is aware of what the school means by ESD and that the school as a whole is motivated to invest in it. Figure 1 provides a visual representation of the conceptual framework, with its two subcontextual characteristics and the six central characteristics.

2. Validity and reliability in questionnaire development

A measurement instrument should demonstrate its validity in a number of areas. Firstly, the theoretical constructs that are to be measured should be translated into items that cover the full concept they intend to measure. If this is the case, the *content validity* of the items or questionnaire is guaranteed (Carmines & Zeller, 1979). This content validity can be evaluated by consulting people who have expertise in the subject area of the items (Karabenick et al., 2007). Secondly, as respondents read and interpret the different items in a questionnaire independently, they should be able to do this correctly. Thus, the *cognitive validity* of the questionnaire items also needs to be addressed.

Karabenick et al. (2007) described a series of sequential cognitive tasks during which the respondent has to perceive, process and respond to a questionnaire item. Questionnaire items that are cognitively valid will evoke a response similar to the response the item intends to elicit. However, Karabenick et al. (2007) notes that several things can go wrong during this process. For example, if a respondent misinterprets a word because, for example, in his or her school this word has a different meaning, this can affect how the respondent retrieves the necessary information to address the item, thus influencing their response. Thus, in order to assess cognitive validity, the intended respondent population should be investigated through cognitive pre-testing. This cognitive pre-testing offers several advantages for instrument development, as qualitative data about how items are interpreted by the respondents can be used to assess the cognitive validity of the items and adjust or remove them as needed (Karabenick et al., 2007; Willis, 2005).

Moreover, it is also a prerequisite that the items measure the intended construct (Carmines & Zeller, 1979). Construct validity thus also refers to the extent to which an item used to measure a construct 'fairly and comprehensively covers that what it claims to cover', thus indicating how valid it is for measuring a certain construct and is thus consistent with the theory that lies at its base (Cohen et al., 2011, p. 188). Another aspect of validity that is to be established, concerns the issue of discriminant validity. This relates to items being empirically distinguishable and not measuring another theoretical construct of interest (Rönkkö & Cho, 2020).

Even if it is established that a questionnaire is valid, it is still necessary to check whether the scales in the questionnaire are reliable. Reliability can be seen as a condition of validity, and it requires an instrument or scale to be stable, replicable and internally consistent under similar conditions ((Cohen et al., 2011). Reliability can be evaluated in terms of the internal consistency of questionnaire scales (Cohen et al., 2011).

3. Developing a valid and reliable questionnaire: process and methodology

In the following sections, we describe the steps taken in the operationalisation of the conceptual framework and the development and validation of the ESD-SOQ. Figure 2 gives a general overview of this process. The four main steps in this study were: the development of the scales for the questionnaire, pre-testing the questionnaire, piloting the questionnaire and, finally, validating the ESD-SOQ.
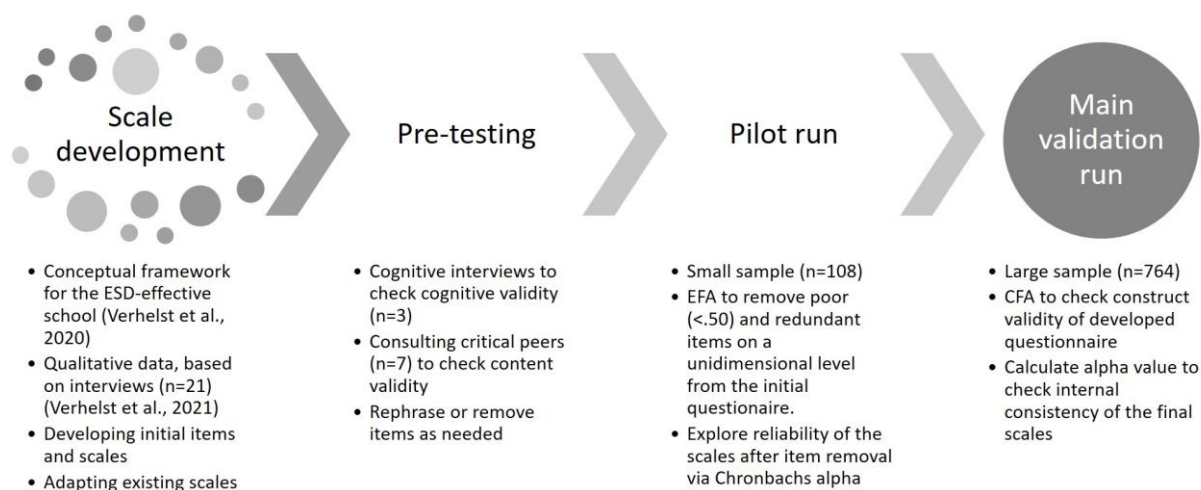

Figure 2: Steps of the development and validation process

Step 1: Scale development

The starting point for the development of the questionnaire was a conceptual framework for the ESD-effective school organisation (Verhelst et al., 2020). The conceptual framework, which serves as the theoretical basis for the novel questionnaire, was the result of a review that synthetized ESD-literature, school management literature and organisational literature. However, the focus was the identification and development of a framework with a focus on ESD-effectiveness. For the development of the different scales and items, the conceptual framework was substantiated by qualitative data provided by school leaders and teachers with experience and expertise in ESD (n = 21) (Verhelst et al., 2021). As described above, school resources include three components (*time management, professional structures* and *physical structures)*. For each component a separate scale was developed. The same

applies for supportive relations, also leading to three scales: *supportive relations within the school team, supportive relations with external partners and other schools* and *supportive relations between school leaders*. The latter was exclusively oriented at school leaders seeing that teachers might not be able to provide valid reports on the supportive relations between school leaders. As a result, twelve scales were developed to measure the eight characteristics of the ESD-effective school. Throughout the process of drafting the first version of the scales, the conceptual framework for the ESD-effective school organisation and the data from the qualitative study were consulted to effectuate the content validity of the items and scales.

Of these twelve scales, two were based on existing instruments. The scale for collective efficacy was based on the collective efficacy subscale from the SAOS questionnaire, with a reported alpha coefficient of .91 (Hoy et al., 2006). The original scale was translated into Dutch and adapted to the context of ESD. The scale for supportive relations within the school team was based on a scale developed by Aelterman et al. (2002) in their research on teacher well-being. This scale was also used by Van Petegem et al. (2010), who reported a Cronbach's alpha value of .84. The other scales were developed from scratch based on data from our conceptual study and the qualitative framework itself.

Furthermore, in order to retain the best items for each of the scales, we purposely developed more items than needed. Typically, a minimum of four items is preferred to provide good coverage of the theoretical construct (Hair et al., 2010). With this in mind, we developed eight to nine items for each of the twelve scales, with the intention to ultimately reduce this number to four to six items for each scale. Every item of the ESD-SOQ has the following answering options: (1) completely disagree; (2) somewhat disagree; (3) neither agree nor disagree; (4) somewhat agree; (5) completely agree. There is also the option to indicate if an item is not applicable.

Step 2: Pre-testing

To address cognitive validity (Karabenick et al., 2007; Willis, 2005), a pre-test consisting of three cognitive interviews was conducted: one with a primary school teacher, one with a secondary school teacher and one with a primary school policy advisor. Based on these interviews, items that respondents perceived as unclear were reformulated. If needed, items were reformulated in order to make them comprehensible for all staff from both primary and secondary schools. Following these cognitive interviews, a panel of critical peers (n = 7) with expertise in ESD, school policy and teacher development programmes related to ESD was consulted. They filled in the questionnaire and simultaneously provided written feedback which was used to check the content validity of the items and scales (Carmines & Zeller, 1979). Moreover, this feedback was used to adapt the questionnaire

look and feel, as well as flow and ambiguities in the formulation. This pre-test led to an initial version of the measuring instrument that was piloted on a small sample of schools in step three.

Step 3: Pilot

The initial version of the ESD-SOQ was piloted on a sample of fifteen schools. This pilot enabled us to remove redundant items if necessary, thus leading to a more parsimonious questionnaire. The pilot sample had an equal representation of both primary and secondary schools. As this pilot still had an exploratory and developmental character, the sample of n = 108 (consisting of school leaders, teachers and other staff of the school) was deemed sufficient. The sample consisted of a small number of school leaders (n = 11), and therefore, it was not possible to conduct analyses on the scale for supportive relations between school leaders. This scale was therefore omitted from further analyses bringing back the number of scales in the ESD-SOQ from twelve to eleven.

The data from this pilot allowed for a unidimensional exploratory factor analysis (EFA) with varimax rotation. Items with a factor loading lower than .50 were removed (Hair et al., 2010). To support this unidimensional approach, we assessed the screeplots and eigenvalues of the separate organisational characteristics. The main rationale for this unidimensional approach comes from the theoretical grounding of the different factors: each scale measures one of the organisational characteristics. Moreover, as we developed a larger number of items with the intent to limit the final scales to the best four to six items, those items with the lowest factor loadings within a scale were also removed. In the case of the scale for 'adaptability', an item with a factor loading of .47 was retained, as this was only slightly below the cut-off value and allowed for the retaining of four items in the scale. Following the EFA, the internal consistency of the provisional scales was checked using Cronbach's alpha, leading to the removal of items that negatively influenced the internal consistency of the scales. Table 1 gives an overview of the different scales, the number of items in the scale, an example item and the corresponding alphas and the factor loadings of the items within the scale. Overall, the reliability of the scales after the pilot proved to be more than sufficient, with alphas ranging from .69 to .94. The pilot led to the reduction of the total number of items in the ESD-SOQ from 85 to 55, with four to six items for each scale. The post- pilot version of the ESD-SOQ, as a product of the process of developing, pre-testing and piloting, allowed us to answer the first research question in this study: How can the characteristics of the ESD-effective school organisation be operationalised? This version of the ESD-SOQ with the eleven scales for the eight characteristics, was then further validated to assess its construct validity.

Table 1: Overview of the scales after the pilot

(Sub)scale	Description	No. of items		Factor loadings of retained items	Example item	Post pilot alpha
		Pre-pilot	Post-pilot			
Sustainable leadership	The presence of sustainable leadership at the school.	8	5	.70 .68 .87 .71 .78	The leadership at this school shows that ESD is here to stay.	.86
Time management	The extent to which planning and time allocation facilitate ESD.	7	6	.69 .70 .71 .74 .68 .88	At this school, we have enough time to work in an investigative fashion with ESD.	.88
Professional structures	The extent to which professional resources facilitate ESD.	7	6	.73 .85 .51 .78 .91 .80	At this school, teachers from different courses and grades collaborate on ESD.	.87
Physical structures	The extent to which the school's physical structures facilitate ESD.	8	4	.55 .70 .59 .50	The infrastructure at this school shows that we focus on ESD.	.69
Pluralistic communication	The extent to which there is a pluralistic communication climate in the school.	9	5	.85 .93 .92 .80 .79	At this school, different opinions on ESD are respected.	.92
	The extent to which (ESD related)	8	6	.73 .64	At this school, different viewpoints	.85

Democratic decision-making	decisions are made in a democratic fashion.			.69 .73 .76 .71	are considered when making decisions about ESD.	
Adaptability	The extent to which a school is able to deal with internal and external demands.	7	4	.86 .87 .47 .52	At this school, we are open to suggestions about ESD that come from within the school.	.80
Supportive relations school team	The extent to which supportive relations within the school team are present and facilitate ESD.	7	5	.75 .90 .91 .75 .75	At this school, teachers are involved in each other's ESD tasks and projects.	.91
Supportive relations external partners	The extent to which supportive relations with external partners and other schools are present and facilitate ESD.	8	6	.77 .83 .92 .97 .89 .73	This school gains inspiration for ESD from collaboration with external partners.	.94
Collective efficacy	The extent to which collective efficacy, focused on ESD, is present at the school.	9	5	.78 .79 .59 .72 .70	At this school, we are convinced that we can provide good ESD education to our pupils.	.87
Shared vision	The presence of a common understanding and motivation towards ESD.	7	4	.86 .79 .70 .52	At this school, it is clear for most of the colleagues what is meant by ESD.	0.82

Step 4: Validation

The last step in the development of the ESD-SOQ involved a large-scale validation study using a purposively selected sample (Teddlie & Yu, 2007). This allowed the assessment of construct validity and the reliability of the final version of the ESD-SOQ. As this study was part of the VALIES research project – a research and development project with a professionalization trajectory for schools that want to implement ESD – the primary and secondary schools participating in this project proved to be an appropriate purposive sample for this study. An online questionnaire was distributed to 52 schools via email. This communication made clear that all members of the school organisations were eligible to complete the questionnaire. Following the first round of email invitations, paper versions of the questionnaire were sent out with a second invitation to fill in the online questionnaire. The final sample size was set at 764 responses. The resulting sample contained an equal representation of experienced (> 10 years of experience) and relatively novice staff (< 10 years of experience), from primary (36%) and secondary (64%) schools. The majority of the respondents identified as female (70%). The larger part of the respondents were teachers (83%), with other respondents having a coordinating function or being the school leader, a policy advisor or having another function within the school.

The validation analysis started off with the calculation of the descriptive statistics. The mean was calculated for each scale, along with the kurtosis and skewness of the scales to check whether the data was normally distributed. This allowed an informed decision on which estimator to use for the confirmatory factor analysis (CFA), which was conducted to check the construct validity of the ESD-SOQ. Factor loadings higher than .50 indicate a good item (Hair et al., 2010). The factor analysis was computed using the lavaan package for latent variable modelling version 6.5 (Rosseel, 2012). The sample size of n = 764 was sufficient for this analysis (Hair et al., 2010). Using the scales of the ESD-SOQ to define the latent variables, an eleven-factor model was defined, with each of the characteristics of the ESD-effective school organisation as factors.

The MLR estimator was used to calculate standard errors that were robust to non-normality and hierarchically clustered data, as this estimator uses more conservative standard errors. Moreover, this estimator is suitable to use with incomplete data (Beaujean, 2014). The following indices were consulted to assess the model fit of the CFA: χ^2 (cut-off: p = 0.05, sensitive to sample size), Standardised Root Mean Residual (SRMR, cut-off: .08), the Tucker Lewis index (TLI, cut-off: .90), the Root Mean Square Error of Approximation (RMSEA, cut-off: 0.05 to 0.10) and the comparative fit index (CFI, cut-off: .90) (Hooper et al., 2008). To improve model fit, modification indices were consulted to add error covariances between items to the model. To assess the discriminant validity, we applied the CI_{cfa} approach as described by Rönkkö and Cho (2020) and we inspected the standardized factor solution.

4. Results of the validation

4.1. Descriptive statistics

The descriptive statistics of the dataset provided a general overview of the data collected by the ESD-SOQ. Overall, the scales for physical structures, pluralistic communication, adaptability, collective efficacy and shared vision differed from normal distribution and tended to be somewhat leptokurtic, confirming the need for an estimator that calculates robust standard errors. Table 2 provides an overview of the different scales and subscales in the ESD-SOQ.

Table 2: Descriptive statistics

	Mean	Std. Deviation	Kurtosis	Skewness
Sustainable leadership	3.419	.731	-.475	.513
Time management	3.114	.831	-.365	.021
Professional structures	3.616	.785	-.591	.569
Physical structures	2.982	.561	-.393	1.272
Pluralistic communication	3.898	.649	-.464	1.182
Democratic decision-making	3.371	.759	-.484	.316
Adaptability	3.760	.626	-.552	1.236
School team	3.261	.575	-.347	.618
External partners	3.177	.784	-.488	.798
Collective efficacy	3.377	.533	-.518	1.023
Shared vision	3.038	.452	.005	2.484

4.2. Confirmatory factor analysis

The eleven-factor CFA model provided a good fit for the observed data. To further improve model fit, error covariances were added between the items of 'Adaptability 1' and 'Adaptability 2' (0.58) and between 'Professional structures 2' and 'Professional structures 4' (-0.32). Although the p-value for χ^2 was not significant, this fit indicator is known to be sensitive to larger sample sizes. The model fit suggested by the other fit indices did show a good fit. The SRMR value of .05 remained well below the cut-off value of .08 and indicated an acceptable fit (Hooper et al., 2008; Hu & Bentler, 1999). In addition, the TLI (.90) and CFI (.91) indicated a fair model fit with respect to the cut-off point of .90 (Hooper et al., 2008). Moreover, the RMSEA value of .045 (with 90% CI [.043, .047]) also support our model. Table 4 gives an overview of the standardised factor loadings for the eleven factors in the CFA model. All factor loadings were sufficiently high, with 41 items having a factor loading higher than .70.

Only 'Physical structures 1' had a factor loading of .46 and did not meet the proposed .50 cut-off value suggested by Hair et al. (2010), but only by a small margin. Next, the correlation matrix (Table 5) shows that most of the factors correlated on a level that was to be expected, indicating that they are related but also sufficiently distinctive to measure different constructs. As none of the correlations were higher than .80, we did not assume any worrying issues with regard to the discriminant validity of the different factors (Rönkkö & Cho, 2020). Secondly, we compared the 95% confidence intervals (CI) against the cut-off values proposed by Rönkkö and Cho (2020) and found that the upper level for the CI for the correlation between democratic decision making and sustainable leadership just exceeds the .80 cut-off value proposed by Rönkkö and Cho (2020). Nevertheless, even if the correlations between the factors are high, this does not automatically mean that there is a discriminant validity problem as these correlations can also be expected based on the theory (Rönkkö & Cho, 2020). Table 3 gives an overview of the proposed classification for the upper level (UL) CI.

Table 3: Classification and cut-offs by Rönkko and Cho (2020)	
Classification	CI_{cfa}
Severe problem	$1 \leq UL$
Moderate problem	$.9 \leq UL < 1$
Marginal problem	$.8 \leq UL < .9$
No problem	$UL < .8$

The results of the CFA thus indicate that the ESD-SOQ and the different scales are valid to measure the constructs represented by the different factors. In answering the second research question, our results show that the different items provide a valid measurement of the characteristics of the framework for an ESD-effective school organisation.

4.3. Reliability

The reliability of the ESD-SOQ was evaluated by examining the internal consistency of the different scales. Cronbach's alpha values of the scales indicated an acceptable to high internal consistency. The only scale with an alpha value below .70 was that for physical structures in the school, with a value of .69. The values suggested high internal consistency for each of the ESD-SOQ scales. Based on these results, it can be stated that the scales are a reliable measurement of the characteristics of an ESD-effective school organisation.

Table 6 provides sample items for each of the validated scales of the ESD-SOQ. A complete version of the questionnaire can be found in appendix 1 (original Dutch version) and appendix 2 (English translation).

Table 4: Standardised factor loadings[a]

Factor	Item	Factor loading	Factor	Item	Factor loading
Sustainable leadership (α = .89)	Item 1	.78	Pluralistic communication (α = .91)	Item 1	.75
	Item 2	.76		Item 2	.83
	Item 3	.80		Item 3	.81
	Item 4	.75		Item 4	.86
	Item 5	.80		Item 5	.87
Time management (α = .90)	Item 1	.68	Democratic decision-making (α = .92)	Item 1	.80
	Item 2	.63		Item 2	.68
	Item 3	.85		Item 3	.86
	Item 4	.87		Item 4	.88
	Item 5	.77		Item 5	.87
	Item 6	.89		Item 6	.76
Professional structures (α = .86)	Item 1	.67	Supportive relations school team (α = .86)	Item 1	.72
	Item 2	.81		Item 2	.80
	Item 3	.74		Item 3	.89
	Item 4	.74		Item 4	.83
	Item 5	.72		Item 5	.56
	Item 6	.57	Supportive relations external partners (α = .94)	Item 1	.80
Physical structures (α = .71)	Item 1	.46		Item 2	.86
	Item 2	.59		Item 3	.91
	Item 3	.79		Item 4	.86
	Item 4	.64		Item 5	.93
Adaptability (α = .86)	Item 1	.69		Item 6	.81
	Item 2	.70	Collective efficacy (α = .83)	Item 1	.57
	Item 3	.78		Item 2	.69
	Item 4	.84		Item 3	.68
Shared vision (α = .78)	Item 1	.86		Item 4	.76
	Item 2	.83		Item 5	.84
	Item 3	.51			
	Item 4	.60			

[a] All factor loadings were statistically significant: $p < 0.005$

Table 5: Correlations between the factors (lower triangle) and upper limits of 95% CI (upper triangle in italic)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Sustainable leadership (1)	1.00	*0.686*	*0.737*	*0.651*	*0.735*	*0.580*	*0.562*	*0.658*	*0.773*	*0.737*	*0.810*
Time management (2)	.615	1.00	*0.669*	*0.674*	*0.642*	*0.601*	*0.422*	*0.573*	*0.631*	*0.619*	*0.614*
Professional structures (3)	.631	.576	1.00	*0.621*	*0.765*	*0.592*	*0.476*	*0.644*	*0.702*	*0.705*	*0.702*
Physical structures (4)	.529	.582	.475	1.00	*0.551*	*0.535*	*0.354*	*0.544*	*0.563*	*0.555*	*0.541*
Supportive relations school team (5)	.638	.554	.692	.441	1.00	*0.569*	*0.557*	*0.729*	*0.725*	*0.745*	*0.722*
Supportive relations external partners (6)	.475	.501	.516	.423	.476	1.00	*0.410*	*0.558*	*0.531*	*0.550*	*0.550*
Pluralistic communication (7)	.483	.319	.392	.228	.466	.319	1.00	*0.515*	*0.679*	*0.528*	*0.628*
Collective efficacy (8)	.565	.491	.550	.409	.664	.461	.434	1.00	*0.725*	*0.747*	*0.652*
Adaptability (9)	.693	.539	.600	.436	.625	.445	.616	.639	1.00	*0.666*	*0.751*
Shared vision (10)	.636	.527	.627	.435	.666	.451	.431	.651	.566	1.00	*0.664*
Democratic decision-making (11)	.736	.526	.620	.431	.640	.460	.558	.586	.684	.587	1.00

Table 6: Sample items for each of the final scales in the ESD-SOQ

Scale	Sample item (translated from Dutch)
Sustainable leadership	The school leadership makes efforts towards anchoring ESD-initiatives on the long term.
Time management	At this school, lessons are scheduled in such a way that they facilitate cross-curricular ESD-activities.
Professional structures	At this school, we devote a lot of attention to who works with who in order to facilitate ESD.
Physical structures	At this school, the infrastructure shows that we are working on ESD.
Supportive relations school team	At this school, there is a pleasant atmosphere among colleagues when working on ESD.

Supportive relations external partners	At this school is stronger when it comes to ESD thanks to the cooperation with external partners.
Pluralistic communication	At this school, we are allowed to be critical about ESD.
Collective efficacy	At this school, we are convinced that we can provide good ESD to our students.
Adaptability	At this school, we are open for suggestions related to ESD that come from inside our school (e.g., teachers, students, staff).
Shared vision	At this school, we work on ESD because we think it is important as a school.
Democratic decision-making	At this school, teachers are involved when making decisions about ESD.

6. Discussion and conclusions

6.1. Contributions to research and practice

With the need for more empirical and quantitative research on school organisations in the area of ESD, this study developed and validated the ESD School Organisational Questionnaire (ESD-SOQ). A framework for the ESD-effective school (Verhelst et al., 2020) provided the starting point for the development of the scales and items in this questionnaire. In the following, we discuss the results of the development and validation process in relation to the three research questions of this study.

The first research question, concerning the operationalisation of the conceptual framework in a questionnaire, was answered by developing a survey instrument consisting of eleven scales that encompassed the breadth and depth of the conceptual framework. By conducting several cognitive interviews (Willis, 2005), consulting critical peers and piloting the questionnaire, we effectuated the cognitive and content validity of the questionnaire (Carmines & Zeller, 1979; Karabenick et al., 2007).

To answer the second research question, concerning the extent to which the items were valid for measuring the characteristics of an ESD-effective school organisation, the construct validity of the ESD-SOQ was taken into account. The confirmatory factor analysis using the MLR estimator with robust standard errors to compensate for non-normality and nested data showed that the items of the eleven-factor model were indeed valid for measuring the characteristics of the conceptual framework for an ESD-effective school organisation. With the exception of 'Physical structures 1', which was marginally below the advised cut-off value, all items exceeded the minimum factor loading of .50

proposed by Hair et al. (2010). Fit indices indicate that the model is adequate in terms of construct validity of the questionnaire (SRMR: .05; TLI: .90; CFI: .91; RMSEA: .041) (Hooper et al., 2008; Marsh et al., 2004). With regards to the discriminant validity, no moderate or severe issues were found. The marginal problem of the 0.810 upper CI for sustainable leadership and democratic decision-making does stand out but does not point at serious issues (Rönkkö & Cho, 2020). We expected the relatively high correlations between the different factors, especially for sustainable leadership and democratic decision-making. The characteristics they aim to measure are themselves interrelated, as was discussed in the description of the framework (Verhelst et al., 2020)

With regard to the third research question, we found that all of the scales are more than sufficient internal consistency, with alpha values ranging from .69 to .94, and thus deemed reliable. As none of the scales had a value higher than .95, we argue that no redundant items were left in the scales, as very high alpha values are an indication of this (Taber, 2018). Comparing the alpha values of our scales for supportive relations within school teams and collective efficacy to the original scales, shows that our scale for supportive relations within school teams has a similar alpha value of .82 (alpha: 0,84; Van Petegem et al., 2010), while the alpha value for collective efficacy dropped to .78 (alpha: 0,91; Hoy et al., 2006). Although our alpha for collective efficacy is still acceptable, a potential explanation can be found in the fact that we transformed the scale so that it would appropriate for the context of ESD, while the original scale concerned collective efficacy within school in general.

The ESD-SOQ provides researchers and educational stakeholders with a means to survey different organisational characteristics related to ESD in school organisations. This instrument offers plenty of opportunity for research on school effectiveness and ESD, and it is also a promising starting point as a self-evaluation tool for schools who want to gain insight into their own working procedures and policy. Via this questionnaire, a school team can obtain insights in characteristics of their organisation related to ESD. This can be a valid starting point for (self-)reflection and organisational development as opening up this information can enable schools to set their own goals for their organisational growth. Knowing that similar measurement instruments are scarce, especially in the school organisational context, we believe that the ESD-SOQ will allow further exploration of the effects of the school organisation on ESD effectiveness, outcomes and impact. By operationalising the framework of the ESD-effective school organisation, we are now able to obtain a broad and holistic perspective on school organisational characteristics. Linking this novel instrument to existing instruments allows for a broad array of possible future research. Seeing that more and more instruments are being developed with a focus on action competence as a potential outcome of ESD (Olsson et al., 2020; Sass et al., 2021), future research could explore the empirical connections between ESD at the student level and the school

organisational characteristics. As the organisational characteristics measured by the ESD-SOQ are argued to facilitate ESD-processes, empirically establishing this connection is made possible by the instruments of the different validation studies. Moreover, there are several instruments on attitudes related to ESD or sustainable development (Gericke et al., 2019; Varela-Losada et al., 2020). As the school staff makes up the school organisation, their attitudes and believes may prove an important aspect of what shapes the different characteristics of an ESD-effective school organisation. With regards to the study of Mogren et al. (2019), which also focused on the school organization, we found that there was a different yet potentially complementary, perspective between their study and our own. While Mogren et al. (2019) had a slightly different perspective, giving prominence to a different conceptual framework and focusing on ESD implementation, they also underlined the need for a holistic and whole school approach (WSA). Nevertheless, we believe that there is a complementarity between the instrument and work by Mogren et al. (2019) and this present study. With a focus on WSA, Mogren and colleagues highlighted the importance of a holistic vision, routines and structures, professional knowledge creation, and practical pedagogical work. An interesting approach would be to investigate how those four constructs, related to a WSA in ESD, are facilitated by the organisational characteristics in present study, which stem from a more organizational-oriented perspective. For example, understanding how the school organization can contribute and facilitate pedagogical practice, would allow us to gain a more in depth-understanding of how a school influences student outcomes. Additionally, as the quality criteria defined by Mogren (2017a, 2017b), who themselves build on the quality criteria perspective by Breiting et al. (2005) provided an inspiration for several of the characteristics for the conceptual framework of an ESD-effective school (Verhelst et al., 2020), a similar endeavour concerning the quality criteria linked to ESD-implementation could expand our understanding of what makes a school ESD-effective. Several of the quality criteria could be understood via the different organizational characteristics. For example, creating, implementing and understanding a holistic idea as a quality criteria would be facilitated via the presence of a shared vision and vice versa. In addition, sustainable leadership, as an example, could be facilitating for each of the different quality criteria. However, the focus of the quality criteria goes to the implementation of ESD, sets them apart from the organizational characteristics of an ESD-effective school. The latter are to be understood as organizational characteristics in which a school can grow and develop and that are argued to facilitate processes within the school that are related to ESD, be it its implementation or other processes such as revising and the continuous development of established ESD practices. The role of sustainable leadership in the creation of an holistic idea, as identified by (Mogren & Gericke, 2017a) would be one of the tracks worth exploring.

6.2. Limitations of the study and the ESD-SOQ

While there are perspectives for further research and for ESD-practice in this paper, it has to be said that limitations for both the study and the questionnaire itself have to be acknowledged. One limitation to our study that should be mentioned, is that although our sample met all of the criteria required to validate the questionnaire, it solely consists of Flemish schools. This limits our claims to the generalisability of the questionnaire and the results to other contexts. It is therefore desirable to translate, adapt and validate the ESD-SOQ so that it can be used in different contexts. Moreover, the translation of the ESD-SOQ into different languages and its validation in different contexts will offer opportunities for additional and comparative research on ESD in a variety of different regions and contexts. To facilitate this, the original Dutch questionnaire and an English translation are included as supplementary materials. Secondly, while one would conceptually be able to describe a school that has no clear understanding of ESD, based on the characteristics of an ESD-effective school, this would not be measurable with the present version of the ESD-SOQ. This limits the use of the ESD-SOQ to schools that have at least some notion of ESD. Notwithstanding these limitations, the questionnaire allows for empirically substantiated insight into school organisational traits related to ESD. On the methodological side, some limitations and concerns are ought to be mentioned. While the CFA found a good model fit, meaning that the items and scales do indeed measure the intended constructs, there is some reason for caution. As organisational research is often confronted with the issue of multicollinearity, one should be cautious for the high correlations among the latent factors. On the other hand, these correlations and connectedness among the organisational characteristics is to be expected, seeing that they are conceptually interrelated. This interrelatedness is also highlighted by school leaders and teachers of ESD-active schools (Verhelst et al., 2021). Lastly, a prominent limitation of this study is that the original target group of the ESD-SOQ was not very well reflected in the sample obtained for the validation. This had its repercussions for the ESD-SOQ itself. While the instrument itself was originally designed to be used by an array of stakeholders within the school team, the sample had an overrepresentation of teachers. Thus, at this moment we are not able to say how this instrument will work for other actors within the school organisation. Nevertheless, as proven by this validation study, the ESD-SOQ is valid and reliable to be used with a similar, e.g. teacher dominated, sample.

A delimitation to this study is that the questionnaire did not include or recognise other stakeholders' perspective on the school organisation. While this is not a limitation in itself, as our focus was on developing and validating a questionnaire for school characteristics without specifying what source to consult, students and parents are important stakeholders within the school organisation. Future research should investigate how this questionnaire might be adjusted, enabling students and parents

to report their perceptions of the school organisation. The addition of their perspectives on the organisational characteristics will offer researchers an even broader insight into the ESD functioning of school organisations. Another delimitation is related to the fact that this instrument is oriented at a broad and exploratory cross-section of an ESD school organisation that included both a battery of organizational characteristics and a diversity of school staff. While this allows for a good overview of the school organisation in relation to ESD, follow-up research can invest in a more in-depth and specific perspective that focusses on one specific characteristic or the perspective of a specific categories of the school staff separately (e.g.: school leaders, teachers and so on).

6.3.   Main conclusion

This study and the ESD-SOQ offer several opportunities for the field of ESD to further expand empirical research. For instance, a quantitative study linking school-level data to data at the classroom, teacher or student levels, would enable researchers to make empirically backed claims about the influence and effectiveness of the school organisation on ESD outcomes. Furthermore, the ESD-SOQ makes it possible to identify strong ESD-schools allowing for in-depth, qualitative, investigations of these school organisations. As research on school organisational characteristics and educational outcomes is scarce in the field of ESD, the investigation of how a school organisation may enable ESD effectiveness is much needed. In addition to the scientific value of the ESD-SOQ, there are also several opportunities to transform this instrument into a self-evaluation tool for school organisations. As school organisations provide a firm ground for ESD-implementation (Scott, 2013), feedback on organisational characteristics via objective and valid information sources is needed (Creemers & Reezigt, 1997). Providing schools the tools to map their organisational traits, provides them with insight in the organisational characteristics that influence their schools' effectiveness towards ESD can provide valuable information for school improvement.

To conclude, if ESD research aims to investigate how to improve ESD and monitor its impact, the ability to map school organisational characteristics facilitating ESD is imperative. The development and validation of the ESD-SOQ have opened up the opportunity for research that further investigates how schools can improve their impact and ESD effectiveness.

7. Acknowledgements

8. Funding and ethical considerations

9. References

Aelterman, A., Engels, N., & Verhaeghe, J. (2002). *Het welbevinden van de leerkracht: de ontwikkeling van een bevragingsinstrument voor het meten van het welbevinden van leerkrachten basis- en secundair onderwijs*.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. Freeman.

Beaujean, A. (2014). *Latent variable modeling using R. a step-by-step guide* [Book]. Routledge, Taylor & Francis Group.

Bennell, S. J. (2015). Education for sustainable development and global citizenship: Leadership, collaboration, and networking in primary Schools. *International Journal of Development Education and Global Learning, 7*(1), 5-32.

Biesta, G. (2009). Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability, 21*(1), 33-46. https://doi.org/10.1007/s11092-008-9064-9

Boeve-de Pauw, J., Gericke, N., Olsson, D., & Berglund, T. (2015). The effectiveness of education for sustainable development. *Sustainability, 7*(11), 15693-15717. https://doi.org/10.3390/su71115693

Bormann, I., & Nikel, J. (2017). How education for sustainable development is implemented in Germany: Looking through the lens of educational governance theory. *International Review of Education, 63*(6), 793-809. https://doi.org/10.1007/s11159-017-9683-9

Bottery, M. (2012). Leadership, the logic of sufficiency and the sustainability of education. *Educational Management Administration & Leadership, 40*(4), 449-463. https://doi.org/10.1177/1741143212438220

Breiting, S., Mayer, M., & Mogensen, F. (2005). *Quality Criteria for ESD-Schools: Guidelines to enhance the quality of Education for Sustainable Development*.

Carmines, E., & Zeller, R. (1979). *Reliability and Validity Assessment*. Sage. https://doi.org/10.4135/9781412985642

Cohen, L., Martin, S., McCulloch, G., O'Sullivan, C., Manion, L., Morrison, K., & Bell, R. (2011). Validity and Reliability. In *Research Methods in Education* (7 ed., pp. 179-215). Routledge.

Creemers, B., & Kyriakides, L. (2010). Using the Dynamic Model to develop an evidence-based and theory-driven approach to school improvement. *Irish Educational Studies, 29*(1), 5-23. https://doi.org/10.1080/03323310903522669

Creemers, B., & Reezigt, G. (1997). School Effectiveness and School Improvement: Sustaining Links. *School Effectiveness and School Improvement, 8*(4), 396-429. https://doi.org/10.1080/0924345970080402

Gericke, N., Boeve-de Pauw, J., Berglund, T., & Olsson, D. (2019). The Sustainability Consciousness Questionnaire: The theoretical development and empirical validation of an evaluation instrument for stakeholders working with sustainable development. *Sustainable Development, 27*(1), 35-49.

Hair, J. F., Jr., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). *Multivariate data analysis : a global perspective* (7th ed. ed.). Upper Saddle River (N.J.) : Pearson education. http://lib.ugent.be/catalog/rug01:001321386

Hargreaves, A. (2007). Sustainable leadership and development in education: Creating the future, conserving the past. *European Journal of Education, 42*(2), 223-233. https://doi.org/10.1111/j.1465-3435.2007.00294.x

Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. Jossey-Bass.

Holm, T., Sammalisto, K., Grindsted, T. S., & Vuorisalo, T. (2015). Process framework for identifying sustainability aspects in university curricula and integrating education for sustainable development. *Journal of Cleaner Production, 106*, 164-174. https://doi.org/https://doi.org/10.1016/j.jclepro.2015.04.059

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic journal of business research methods, 6*(1), 53-60. https://doi.org/10.21427/D7CF7R

Hoy, W. K., Tarter, C. J., & Miskel, C. G. (2013). *Educational administration: Theory, research, and practice* (9 ed.). McGraw-Hill.

Hoy, W. K., Tarter, C. J., & Woolfolk Hoy, A. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425-446.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Iliško, D., & Badyanova, Y. (2014). A case study of ESD implementation: Signs of sustainable leadership. *Discourse and Communication for Sustainable Education, 5*(1), 38-48. https://doi.org/10.2478/dcse-2014-0004

Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., Groot, E. D., Gilbert, M. C., Musu, L., Kempler, T. M., & Kelly, K. L. (2007). Cognitive Processing of Self-Report Items in Educational Research: Do They Think What We Mean? *Educational Psychologist, 42*(3), 139-151. https://doi.org/10.1080/00461520701416231

Kuzich, S., Taylor, E., & Taylor, P. C. (2015). When policy and infrastructure provisions are exemplary but still insufficient: Paradoxes affecting education for sustainability (EfS) in a custom-designed sustainability school. *Journal of Education for Sustainable Development, 9*(2), 179-195. https://doi.org/10.1177/0973408215588252

Laurie, R., Nonoyama-Tarumi, Y., McKeown, R., & Hopkins, C. (2016). Contributions of Education for Sustainable Development (ESD) to Quality Education: A Synthesis of Research. *Journal of Education for Sustainable Development, 10*(2), 226-242. https://doi.org/10.1177/0973408216661442

Leo, U., & Wickenberg, P. (2013). Professional norms in school leadership: Change efforts in implementation of education for sustainable development. *Journal of Educational Change, 14*(4), 403-422. https://doi.org/10.1007/s10833-013-9207-8

Lijmbach, S., Van Arcken, M. M., Van Koppen, C. S. A., & Wals, A. E. J. (2002). 'Your view of nature is not Mine!': Learning about pluralism in the classroom. *Environmental Education Research, 8*(2), 121-135. https://doi.org/10.1080/13504620220128202

MacBeath, J., & Mortimore, P. (2001). School effectiveness and improvement: The story so far. In P. Mortimore & J. MacBeath (Eds.), *Improving school effectiveness* (pp. 1-21). Open University Press.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing

Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Mogren, A., & Gericke, N. (2017a). ESD Implementation at the school organisation level, Part 1 – Investigating the quality criteria guiding school leaders' work at recognized esd schools. *Environmental Education Research, 23*(7), 972-992. https://doi.org/10.1080/13504622.2016.1226265

Mogren, A., & Gericke, N. (2017b). ESD Implementation at the school organisation level, part 2 – Investigating the transformative perspective in school leaders' quality strategies at ESD schools. *Environmental Education Research, 23*(7), 993-1014. https://doi.org/10.1080/13504622.2016.1226266

Mogren, A., & Gericke, N. (2019). School leaders' experiences of implementing education for sustainable development—anchoring the transformative perspective. *Sustainability, 11*(12). https://doi.org/10.3390/su11123343

Mogren, A., Gericke, N., & Scherp, H.-Å. (2019). Whole school approaches to education for sustainable development: a model that links to school improvement. *Environmental Education Research, 25*(4), 508-531. https://doi.org/10.1080/13504622.2018.1455074

Nikel, J., & Lowe, J. (2010). Talking of fabric: a multi-dimensional model of quality in education. *Compare: A Journal of Comparative and International Education, 40*(5), 589-605. https://doi.org/10.1080/03057920902909477

Olsson, D., Gericke, N., Sass, W., & Boeve-de Pauw, J. (2020). Self-perceived action competence for sustainability: the theoretical grounding and empirical validation of a novel research instrument. *Environmental Education Research*, 1-19. https://doi.org/10.1080/13504622.2020.1736991

Rampasso, I. S., Anholon, R., Silva, D., Cooper Ordoñez, R. E., Santa-Eulalia, L. A., Quelhas, O. L. G., Leal Filho, W., & Granada Aguirre, L. F. (2019). Analysis of the perception of engineering students regarding sustainability. *Journal of Cleaner Production, 233*, 461-467. https://doi.org/https://doi.org/10.1016/j.jclepro.2019.06.105

Rönkkö, M., & Cho, E. (2020). An Updated Guideline for Assessing Discriminant Validity. *Organizational Research Methods*, 1094428120968614. https://doi.org/10.1177/1094428120968614

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2). https://doi.org/10.18637/jss.v048.i02

Rudsberg, K., & Öhman, J. (2010). Pluralism in practice – Experiences from Swedish evaluation, school development and research. *Environmental Education Research, 16*(1), 95-111. https://doi.org/10.1080/13504620903504073

Sass, W., Boeve-de Pauw, J., De Maeyer, S., & Van Petegem, P. (2021). Development and validation of an instrument for measuring action competence in sustainable development within early adolescents : the action competence in sustainable development questionnaire (ACiSD-Q). *Environmental Education Research*. https://doi.org/https://doi.org/10.1080/13504622.2021.1888887

Schelly, C., Cross, J. E., Franzen, W., Hall, P., & Reeve, S. (2012). How to go green: Creating a conservation culture in a public high school through education, modeling, and communication. *Journal of Environmental Education, 43*(3), 143-161. https://doi.org/10.1080/00958964.2011.631611

Scott, W. (2013). Developing the sustainable school: thinking the issues through. *The Curriculum Journal, 24*(2), 181-205. https://doi.org/10.1080/09585176.2013.781375

Sinakou, E., Donche, V., Boeve-de Pauw, J., & Van Petegem, P. (2019). Designing Powerful Learning Environments in Education for Sustainable Development: A Conceptual Framework [Review]. *Sustainability, 11*(21), 23. https://doi.org/10.3390/su11215994

Singer-Brodowski, M., Brock, A., Etzkorn, N., & Otte, I. (2019). Monitoring of education for sustainable development in Germany – Insights from early childhood education, school and

higher education. *Environmental Education Research, 25*(4), 492-507. https://doi.org/10.1080/13504622.2018.1440380

Sleegers, P. J. C., Thoonen, E. E. J., Oort, F. J., & Peetsma, T. T. D. (2014). Changing classroom practices: the role of school-wide capacity for sustainable improvement. *Journal of Educational Administration, 52*(5), 617-652. https://doi.org/10.1108/JEA-11-2013-0126

Starratt, R. J. (2007). Leading a Community of Learners: Learning to be Moral by Engaging the Morality of Learning. *Educational Management Administration & Leadership, 35*(2), 165-183. https://doi.org/10.1177/1741143207075387

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Teddlie, C., & Reynolds, D. (2006). *The international handbook of school effectiveness research*. Routledge.

Teddlie, C., & Yu, F. (2007). Mixed Methods Sampling: A Typology With Examples. *Journal of Mixed Methods Research, 1*(1), 77-100. https://doi.org/10.1177/1558689806292430

Thoonen, E. E. J., Sleegers, P. J. C., Oort, F. J., & Peetsma, T. T. D. (2012). Building school-wide capacity for improvement: the role of leadership, school organizational conditions, and teacher factors. *School Effectiveness and School Improvement, 23*(4), 441-460. https://doi.org/10.1080/09243453.2012.678867

UNESCO. (2014). *Shaping the Future We Want. UN Decade of Education for Sustainable Development (2005-2014): Final Report.* S. a. C. O. U. United Nations Educational. http://www.unesco.org/ulis/cgi-bin/ulis.pl?catno=230302&set=005BB4C0BD_2_171&gp=1&lin=1&ll=1

UNESCO. (2020). *Education for sustainable development: a roadmap* https://www.gcedclearinghouse.org/sites/default/files/resources/200782eng.pdf

Van Petegem, P., De Maeyer, S., & Mahieu, p. (2010). *De participatiebarometer. ontwikkeling van een instrumentarium voor het meten van leerkracht en ouderparticipatie* [Book]. Acco.

Varela-Losada, M., Pérez-Rodríguez, U., Lorenzo-Rial, M., Vega-Marcote, P., & Reid, A. (2020). Dealing with global environmental change: the design and validation of the GEC attitude scale. *Environmental Education Research*, 1-22. https://doi.org/10.1080/13504622.2020.1822990

Verhelst, D., Vanhoof, J., Boeve-de Pauw, J., & Van Petegem, P. (2020). Building a conceptual framework for an ESD-effective school organization. *The Journal of Environmental Education*, 1-16. https://doi.org/10.1080/00958964.2020.1797615

Verhelst, D., Vanhoof, J., & Van Petegem, P. (2021). School effectiveness for education for sustainable development (ESD): What characterizes an ESD-effective school organization? *Educational Management Administration & Leadership*. https://doi.org/10.1177/1741143220985196

Wals, A. E. J. (2009). A Mid-DESD Review: Key Findings and Ways Forward. *Journal of Education for Sustainable Development, 3*(2), 195-204. https://doi.org/10.1177/097340820900300216

Waltner, E.-M., Rieß, W., & Brock, A. (2018). Development of an ESD indicator for teacher training and the national monitoring for ESD implementation in Germany. *Sustainability, 10*(7), 1-17. https://EconPapers.repec.org/RePEc:gam:jsusta:v:10:y:2018:i:7:p:2508-:d:158522

Willis, G. B. (2005). *Cognitive interviewing: a tool for improving questionnaire design* [Book]. Sage.