methods, data, analyses | Vol. 9(1), 2015, pp. 111-136 DOI: 10.12758/mda.2015.004

How to Measure Participation of Pupils at School. Analysis of Unfolding Data Based on Hart's Ladder of Participation

Daniela Wetzelhütter & Johann Bacher
Johannes Kepler University (JKU) of Linz

Abstract
A renaissance of the scientific study of participation has fortunately taken place in recent years. As a consequence of the Convention on the Rights of the Child, forms of participation of children were implemented in different fields of society, especially in the community and at the school level. These programs and attempts are described by science. However, little attention has been paid to the measurement of participation. Therefore, this paper proposes and analyses a measurement instrument for the frequently cited Ladder of Participation by Hart (1992). His model assumes that different levels (degrees) of participation exist, whereby the extent increases with each level (stage).

Keywords: Multidimensional Scaling, Correspondence Analysis, CATPCA, Unfolding, Ladder of Participation, Degree of Participation

1 Introduction

Lewin and Lewin (1982 [1941]) demonstrated in their essay "Democracy and School" that practical experience of democracy is important for learning democratic attitudes. They dealt with the participation in school lessons and pointed out that such participation is potential in primary children. However, the participation task should be defined manageably. According to the authors, democracy in school leads to a more amiable climate in groups and improves individual academic performance.

Childhood participation became the focus of social science research in the 1990s. At that time it was related to the emergence of childhood research (James, Jenks & Prout, 1998; Zinnecker, 1999; Wilk & Bacher, 1994; Bacher, Gerich, Lehner, Straßmair & Wilk, 1999). In this connection, scientists intensively discussed questions such as the relationship between project-based participation and formal democratic structures.

In the past few years (cf. Wetzelhütter, Paseka & Bacher, 2013) the topic has once again been addressed, but this time from the perspective of school development research. One such research focus was the pupil's perspective of involvement or rather participation in school. In this connection, research on participation often refers to the Ladder of Participation developed by Hart (1992) or comparable models from Oser and Biedermann (2006) or Arnstein (2011[1969]). Regarding Hart's construct, the intensity of participation is the core element, where at each "level" the extent of participation of children increases (see Chapter 2). The model suggests a one-dimensional (degree of participation) unfolding scale, whose items are characterized by different intensities of participation. It is assumed that the indicators neither mutually exclude each other (this means that they may occur concurrently) nor mutually depend on each other (this means that reaching a stage does not require the presence of another step – as would be the case for a Guttman scale).

However, Hart makes no clear statement about the dimensional structure of his concept. Therefore, the paper addresses the following question:

- Can participation in school be measured one-dimensionally as assumed by Hart?

In part 2, the paper describes the theoretical background of Hart's Ladder of Participation. The methodical approach (operationalization, databases and data analysis)

Direct correspondence to
Daniela Wetzelhütter, Johannes Kepler University (JKU) of Linz, Department of Empirical Social Research, Austria.
E-mail: Daniela.Wetzelhuetter@jku.at

Acknowledgment: We would like to thank the two anonymous reviewers for their important suggestions.

is discussed in part 3. The results (descriptive statistics, examination of dimensionality, and validation of the outcome) are presented in part 4. Finally, part 5 provides a summary followed by conclusions.

2 Theoretical Framework (Measurement Model): Hart's "Ladder of Participation"

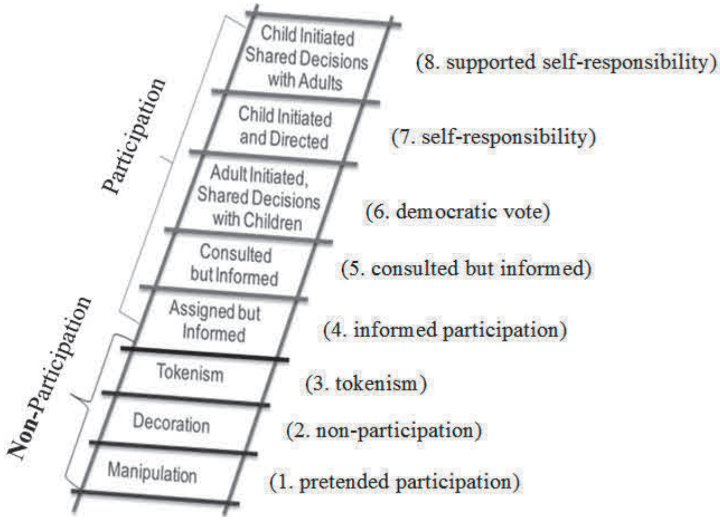
Hart (1992) defines participation as a "process of sharing decisions which affect one's life and the life of the community in which one lives" (p. 5). Therefore, participation is (for him) a *process* of collective decision making. Based on an investigation of participation projects sponsored by UNICEF (programs for street children or prevention programs), Hart (1992) developed an eight-stage Ladder of Participation, based on a model by Arnstein published in 1969 (Hart, 1992). This model generally differentiates between non-participation and participation (Figure 1), though every stage describes a different "process" of children's involvement. Hart further divides non-participation into:

1. "Manipulation" (pretend participation),
2. "Decoration" (children are "used," but participation is not pretend),
3. "Tokenism" (token policy – e.g. when children are given a voice, but are badly/not prepared for the topic on which they are voting).

Participation begins with the fourth stage – the intensity is divided into:

4. "Assigned but Informed" ("just" informed participation of children),
5. "Consulted but Informed" (informed participation, where the opinions of children are treated seriously),
6. "Adult Initiated, Shared Decisions with Children" (adults initiate projects but the decision making is shared by children and adults),
7. "Child Initiated and Directed" (initiated and directed by children), and
8. "Child Initiated, Shared Decisions with Adults" (projects which are initiated by children, but the decision making is shared by children and adults).

"Real participation" is placed at the top position. The lower the degree of participation, the lower the scope and the involvement. At this point it has to be mentioned that the ladder metaphor is an unfortunate choice, since it could be assumed that reaching the top step implies climbing the previous stages. This would imply that the scale is constructed in the logic of a Guttman scale. As Hart pointed out in a later paper (2008), despite an increasing degree of participation for each level, the ladder does not stand for a developmental process: "In fact the ladder is primarily about the degree to which adults and institutions afford or enable children to participate" (p. 23). This means each stage measures (as outlined above) the degree of



According to: Hart (1992, p. 8) – subsequently used definitions in brackets

Figure 1 Hart's Ladder of Participation

participation. Depending on the project/topic and participant, the level may be of different degrees of involvement. For instance, a project may directly start at level 7 ("Child Initiated and Directed") without passing the previous stages of 1 to 6.

This paper examines the question of whether the degree of participation can be measured one-dimensionally using a simple battery of eight questions, outlined in more detail in the next chapter.

Scientifically, such an instrument would be important to develop and test theories concerning the influence of participation on the individual, class and school level. According to Lewin (e.g. Lewin, Lippitt & White, 1939) it would be interesting to analyze if participation results in higher abilities and more effectiveness as well as in less dissocial (aggressive) behavior of school children.

3 Methodical Approach

3.1 Operationalization: Degree of Participation

In the "(Do) students develop school!?" project (Altrichter, Bacher, Langer, Gamsjäger & Wetzelhütter, 2012), a scale was developed to measure the degree of participation in a school as indented by Hart's Ladder of Participation. The final version – a rating in the form of a five-point Likert scale – can be seen in Table 1.

Table 1 Rating scale to measure participation of students in school

What applies to your school and how much?	fully applies	largely applies	uncertain	does not apply much	does not apply at all	concrete example in a school
We may propose our own projects and implement them independently.	Level 7 "self-responsibility" (V7)					Students propose to publish an online school newspaper; they can use the infrastructure and produce the newspaper alone
We may propose our own projects and teachers/directors help us with the implementation.	Level 8 "supported self-responsibility" (V8)					As above, but teachers and the director support them

Table 1 continued

How are choices made at your school?	fully applies	largely applies	uncertain	does not apply much	does not apply at all	concrete example in a school
Although we are asked about our opinion on important decisions, it will not be considered.	Level 1 "pretend participation" (V1)					The school board wants to renovate the school garden. The students are asked for their ideas, but these ideas are ignored in the final decision.
We don't get enough information about important decisions, but we are allowed to vote.	Level 3 "tokenism" (V3)					The students were informed that two different textbooks are available and that they should decide which of them they want to use. No further information was given to them.
We get informed about important decisions but we are not allowed to vote.	Level 4 "informed" (V4)					The school informed their students that the school garden will be renovated.
We get informed about important decisions and our opinion is asked for.	Level 5 "consulted but informed" (V5)					The school informed their students that the school will renovate the school garden and asked the students for their ideas.
When it comes to important decisions, several options are presented to us that we then vote on.	Level 6 "democratic vote" (V6)					The school intends to renovate the school garden. Different concepts were presented to the students. Each student could vote for one.
Important decisions are made without our involvement.	Level 2 "non-participation" (V2)					The school got new computers in the holidays. The students were not asked.

Note. Translated version based on the original German version.

Each item (every statement) represents one level of Hart's ladder. For instance, the statement "When it comes to important decisions, several options are presented to us that we then vote on" corresponds to stage 6 (democratic vote). The five-point Likert scale provided the following potential responses: "fully applies," "largely applies," "uncertain," "does not apply much," and "does not apply at all."

This scale belongs to the group of subject-oriented scaling procedures (see Likert, 1932). In our case, the subjects are schools and school classes. The scale consists of eight items (objects) which may be ranked on one dimension (degree of participation) at which each subject has an ideal point (perceived degree of participation) as suggested in Coombs' (1964) unfolding model (for details see Section 3.3).

3.2 Data Sets and Sampling Design

The data analysis is based on two data sets (see Table 2). The first data set was generated as a pre-test study for calibration (calibration sample) and the second as a nationwide sample for validation (validation sample). The sample is representative for vocational colleges ("Berufsbildende mittlere und höhere Schulen")[1] and upper level grammar schools ("Allgemeinbildende höhere Schulen, Oberstufe") with respect to sex and language spoken at home. In both cases, the selection of the pupils followed a stratified cluster sampling (see for example Bacher, 2009 or Sturgis, 2004). During the first stage, of the calibration sample, entire classes were selected per school grade. This procedure resulted in 22 (64.7%) out of 34 school classes, with n=382 (=86.8% return rate) surveyed pupils out of 440 selected pupils.

During the first stage of the validation sample, 282 schools (=36.2%, including 65 alternatively drawn schools) of several types of schools[2] were selected in each province of Austria (see Table 2). Forty-five percent of these schools were willing to participate (n=127). Within these schools, two classes (altogether 254 classes) were randomly defined based on the Kish selection grid (see Kish 1965). This approach was chosen to avoid a conscious choice by headmasters, for example of "laptop classes" or classes with a specialization in IT topics. Ultimately, 127 Austrian schools (=45% of the gross sample) with 253 classes (=99.6% of the selected classes) and 4,101 pupils (=67.8% of the drawn pupils) took part from all over the country. Due to the fact that participation in the survey was acquired exclusively via telephone and conducted online, the return rate is classified as good.

1 The OeAD (2014) provides information about the Austrian educational system.

2 AHS (upper level of grammar schools) and BMHS (vocational colleges) including A&F (schools and colleges for agriculture and forestry) and BAKIP (training institutions for kindergarten teachers), see OeAD (2014)

Table 2 Total population with drawn and realized samples

Sample	Stage	Total Population N	Drawn Sample		Realized Sample	
			N	%(a)	n	%(b)
Calibration Sample	1 = classes	34	22	64.7	22	100.0
	2 = pupils	698	440	63.0	382	86.8
Validation Sample	1 = schools	779	282	36.2	127	45.0
	2 = classes	1577(c)	254	16.1	253	99.6
	3 = pupils	37494(d)	6045(e)	16.1	4101	67.8

Note. (a) Proportion of the drawn sample in terms of the basic population.
(b) Proportion of the realized sample in terms of the drawn sample.
(c) Total population of classes of the drawn sample of schools (secondary education level).
(d) Total population of pupils of the drawn sample of classes (secondary education level).
(e) Estimated number of pupils based on the average number of pupils per class of the individual school.

3.3 Data Analysis

3.3.1 Model Specification

As mentioned, the Hart ladder does not constitute a Guttman scale. To reach a certain degree of participation in a school, it is not necessary to pass stages with lower levels of participation. A school may immediately start, for example, with informed participation (stage 4) without passing stages 1 (pretend participation) to 3 (tokenism). According to Coombs (1964), the suitable scaling procedure for our scale is an "unfolding model." Figure 2 illustrates the idea.

This model presumes that one dimension is measured at which the items are sorted according to one criterion and every person (stimulus scaling) or every object (object scaling) has one ideal point on the scale. The ideal point (cf. for example de Leeuw, 2005) represents the point of maximum preferential choice of one person or the point that best represents the "required" object.

In the present study, the appropriate model may be realized as follows. In every school or in each class a certain degree of participation (ideal point) exists, which is judged by the surveyed pupils. The level that comes closest to this (perceived) degree receives the highest agreement, the level that is second closest the second largest agreement, and so on. Based on these individual rank orders (so the assumption in regard to Hart's ladder), a common solution (order of the participation items) can be determined. This solution reflects the extent of participation

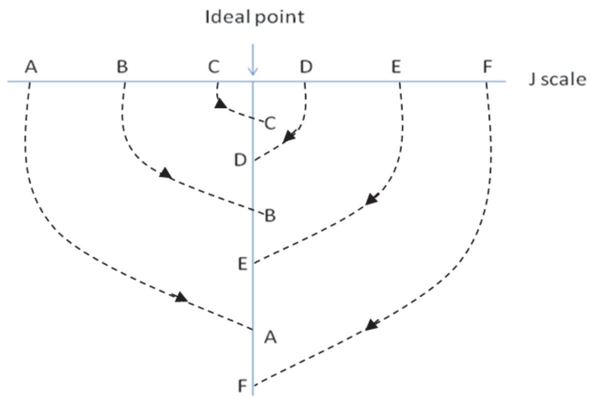


Figure 2 Unfolding model according to Coombs (1964, p. 80)

along one dimension. Practically, an unfolding analysis can be performed within multidimensional scaling (MDS) procedures. Taking these considerations into account, and referring to the narrow definition of multidimensional scaling of Cox and Cox (2001), it is intended to locate the items within a one-dimensional or at least a low-dimensional space that best matches the original (dis)similarities and distances between the items (objects). For the present paper, PROXSCAL (Meulman & Heiser, 2011) was used to perform the multidimensional scaling[3]. This procedure enables the user to generate the distances and dissimilarities (cf. Borg, Groenen & Mair, 2013) of the items for several groups as a so-called "Three-Way Model" (cf. for example Kruskal & Wish, 1978). This means the analyses are based on distance matrices generated per "school class" (calibration sample) or per "school" (validation sample). For PROXSCAL (Meulman & Heiser, 2011; IBM-SPSS, 2011b; Borg et al., 2013), the following specification was chosen:

- *Proximities*: Euclidean distance
- *Starting configuration*: classical (Torgerson)
- *Condition*: matrix (distances are exclusively comparable within one matrix)
- *Transformation*: ordinal (ties: keep ties)
- *Model*: Generalized Euclidean: each school or class has an individual space, which is a rotation of the common space with subsequent weighting of the dimensions

3 For this analysis cases were excluded with missing values and/or without variation in the responses.

The specification follows the recommendation in the literature (IBM-SPSS, 2011b; Borg & Groenen, 2005; Borg et al., 2013). We selected ordinal transformation for ordinal data and kept ties for formal reasons, because a rating scale with a small number of levels may result in the same proximity values for some items (Borg et al., 2013). We chose Torgerson as the starting configuration because "this option tends to give better quality solutions" (Borg & Groenen, 2005) and defined "matrix" as the condition for indicating that a separate transformation has to be found for each matrix (ibid). We chose with "Generalized Euclidean" a dimension-weighting model, to account for individual differences (see for example Borg & Groenen, 2005) between classes (calibration sample) and schools (validation sample) as implied by unfolding data.

However, the specifications are arbitrary to a certain degree as is the case for all applications of MDS. Therefore, we used correspondence analysis, specifically CATPCA (categorical principal component analysis, see Meulman & Heiser, 2011), as an additional method (see below). The specification of an ordinal transformation implies a non-metric MDS. Non-metric MDS requires a specific number of objects (in our case items) for each dimension (Bacher, Pöge & Wenzig, 2010). In this regard, the ratio of 5:1 (items : dimensions) for a complete dissimilarity matrix (ibid) is recommended, or Kruskal and Wish's (1978) rule of thumb of at least more than fourfold the number of items per dimension.

Due to the problem of choosing an appropriate specification for MDS and due to the small number (eight) of analyzed items, the dimensionality is also tested by a correspondence analysis (CATPCA). This procedure tolerates a smaller number of items. Two items per dimension are sufficient. In addition, the user does not have to make specifications concerning starting configuration, distance measures, transformation, handling of ties etc. All these decisions follow from the statistical model of correspondence analysis (of course with the disadvantage of being less flexible).

The correspondence analysis is similar to MDS (e.g. Borg & Groenen, 2005; Greenacre, 2013). It enables the interpretation of distances (Chi^2 and weighted Euclidean distances are calculated) between variables (in this context items) as well as objects (in this context pupils). In the literature, the similarity of these two methods is mentioned frequently (cf. Hoffman & de Leeuw 1992; Cherkassky & Mulier, 2007; Young, 1985). The multiple correspondence analysis is applied in equivalent situations to the MDS (Meulman, van der Kooij & Heiser 2004) or in combination with MDS (e.g. Green, 2010).

In this paper, CATPCA was performed. This procedure (cf. Blasius & Thiessen, 2012) is an appropriate technique for analyzing ordered categorical variables, and was performed with the option of optimizing the relationship between variables (Normalization: VPRINCIPAL). The level of analysis was pupils; the order of the variables was interpretable and the representation comparable with the MDS solutions. The data analysis was performed with IBM SPSS Version 22 (IBM-SPSS, 2011b). Applying factor analysis would not be appropriate from a theoretical point of view because a bipolar concept often results in two interpretable but unrelated factors (see van Schuur & Kiers, 1994) which would not correspond to the assumed theoretical concept (see Chapter 2). Nonetheless, explorative factor analysis can be applied to our data.[4]

3.3.2 Analyzing Procedure

Data analysis was being performed in three steps:

Firstly: In order to test whether our measurement instrument of Hart's ladder is one-dimensional, one- and two-dimensional solutions were computed based on the calibration sample, using MDS (PROXSCAL) and CATPCA.

For the decision regarding which configuration fits better, the following criteria were applied (for a justification, see below):

- MDS: stress level: Stress 1 <0.2; Stress 2<0.4; congruence coefficient: >0.9
- CATPCA: Eigenvalue >1 and Cronbach's alpha >0.7
- MDS and CATPCA: substantive interpretation

Concerning the stress level, it has to be mentioned that Borg et al. (2013) already emphasized that "evaluating a given Stress value is a complex matter" (p. 23). This means that several considerations have to be made. Stress decreases, for instance, with the number of dimensions but increases, for instance, with the number of items (points/objects) or proportion of error components (noise) in the data (Borg et al., 2013). Therefore, the frequently cited benchmark of Kruskal (1964) or Kruskal and Wish (1978), which states that Stress 1 has to be lower than 0.2, and the benchmark of Fahrmeir et al. quoted in Gediga (1998), which says that Stress 2 should be lower than 0.4, are both applied conditionally. The congruence coefficient (which, in regard to Borg et al., (2013), can be interpreted as a correlation coefficient) should be close to the value of 1. Concerning CATPCA, it can be noted that the eigenvalue(s) should be considerably higher than 1 (according to Kaiser's criterion, cf. for example Hardy & Bryman, 2009) and Cronbach's alpha should be higher than 0.7 (cf. for example Nunnally, 1978). Additionally, the correlation of the distance matrices (see CPCC below) obtained by the MDS and CATPCA, based on the one- and two-dimensional solution, is calculated in order to verify the findings.

4 In our case, two factors (Eigenvalue >1) can be extracted for each sample (validation and calibration). Based on the validation sample, variables representing the first four steps of the ladder (V1–V4) load on the first factor, the next two variables (V5, V6) on both factors, and the last two (V7, V8) on the second factor. The outcome for the calibration sample is slightly different. V3 loads on both factors, while V5 loads on the first instead of both factors. Neither result corresponds to the theoretical concept (Chapter 2). The result partially reproduces the grouping of items in three clusters (see Chapter 4).

Lastly, substantive interpretation will be considered, which means that a solution is supported if the arrangement of the items is explicable.

Secondly: The validity of the scale was examined. Initially it was tested to which extent the results of the calibration sample are replicated in the Austria-wide validation sample. Afterwards, the cophenetic correlation coefficient (CPCC) was computed in order to judge how well the validation sample reproduces the calibration sample (see Romesberg, 2004). The CPCC is defined as (linear correlation coefficient):

$$CPCC = \frac{\sum_{i<j}(d_x(i,j)-\overline{d}_x)(\mathrm{d}_y(i,j)-\overline{d}_y)}{\sqrt{\left[\sum_{i<j}(d_x(i,j)-\overline{d}_x)^2\right]\left[\sum_{i<j}(d_y(i,j)-\overline{d}_y)^2\right]}}$$

$d_x(i,j)$ = Euclidean distance between the i^{th} and j^{th} items of the calibration sample
$d_y(i,j)$ = Euclidean distance between the i^{th} and j^{th} items of the validation sample
$\overline{d}_x$ = average of $(d_x(i,j)$; $\overline{d}_y$ = average of $(d_y(i,j)$

The CPCC measures the similarity between two dissimilarity matrices – in our case between the computed dissimilarity matrix for the calibration and the computed dissimilarity matrix for the nationwide validation sample. Both matrices were calculated based on representation in a one- or two-dimensional space. According to Romesberg (2004), a value from 0.8 upwards is acceptable, since "the distortion is not great" (p. 27).

Thirdly: Two content variables were included in the scaling process. Validity was examined by two criteria: the location of the items in the dimensional space and the stability of the obtained scaling model by comparison of the model with and without content variables. The content variables are "co-decision to decorate the classroom" and "co-decision of school rules." Regarding decisions about rules at school, they are mostly made by laws, teachers, or school heads. Merging the results of the mentioned sparkling-science project (Altrichter et al., 2012) shows that on unique occasions or with informal participation possibilities, pupils often do not get enough background information or time to develop/determine their position. Therefore, it is reasonably assumed, if pupils are involved in decision making, that they become informed just before the vote is carried out in order to simplify the procedure. Accordingly, the hypothesis is:

H1: The item "co-decision of school rules" is located next to the level "tokenism" (level 3, V3), and the obtained structure (without the criteria variable) doesn't change.

Furthermore, the sparkling-science project (Altrichter et al., 2012) showed that participation most likely takes place in "peripheral" issues. For instance, the "co-decision to decorate the classroom" is usually based on a joint decision of teachers and pupils. Accordingly, the hypothesis is:

H2: The item "co-decision to decorate the classroom" is located next to the item "consulted but informed" (level 5, V5), and the obtained structure of the model (without the criteria variable) doesn't change.

The results of the described data analysis will now be presented in Chapter 4, with descriptive results provided in the appendix.

4 Results

4.1 Dimensionality of the scale in the calibration sample

In order to determine the dimensionality of the scale, one- and two-dimensional solutions were computed by MDS (PROXSCAL) and by CATPCA. Table 3 summarizes the results of the calibration sample.

The one-dimensional MDS solution doesn't fulfill the defined stress criteria of 0.20 (Stress 1) or 0.40 (Stress 2), which indicates a serious proportion of noise (errors) in the data. As expected, a two-dimensional model decreases the stress level, but the stress criteria are not met. However, stress 1 is only slightly higher than the threshold of 0.20 and in both cases the congruence coefficient is close to a perfect solution (value of 1). The results of CATPCA support a one-dimensional solution more clearly. The second eigenvalue is sufficiently smaller than the first eigenvalue. The consistency of the second dimension is poor (Cronbach's alpha=0.206), whereas the Cronbach's alpha of the first dimension is 0.773. Furthermore, the correlation between the distances (r=0.930) obtained by MDS (PROXSCAL) and CATPCA also supports the one-dimensional solution. The poor correlation (r=0.501) based on the two-dimensional models indicates that the respective solutions are less stable and robust. They depend on the applied method (see also Figure 3 A–D).

From a substantive point of view, both MDS solutions are acceptable (see Figure 3 A and C). In either case, the items can be divided into two groups of items at the zero point of the x-axis, which may be interpreted as the scope of action. This result is similar to Hart's outline – only item V4 is allocated in the "wrong group." The left group can be named "effective/adequate participation" and the right group may be identified as "ineffective/inadequate participation." However, the two-dimensional solution is not interpretable as a "ladder of participation." Based on Figure 3C, the scope of action is further distinguishable due to the initiative of involvement in the dependence of students (upper half) vs. teachers (lower half). Nonetheless, the rank order of the items (reading from left to right)

Table 3 Summarized results for determining dimensionality using PROXSCAL and CATPCA (calibration sample)

	PROXSCAL			CATPCA		
Model	Stress 1	Stress 2	Congruence coeff.	Eigenvalue (Variance)	Cronbach's alpha[b]	CPCC[a]
One-dim. solution	0.316	0.616	0.95	3.1 (38.65%)	**0.773**	**0.930** (p=0.000)
Two-dim. solution	**0.203**	0.556	0.98	1.2 (15.26%)	0.206	0.501 (p=0.007)

Note. (a) n=28 distances (=8 times 7/2); CPCC between distance matrices obtained by the coordinates produced by PROXSCAL and CATPCA. A correlation of 1 would imply that the distances between items computed by PROXSCAL perfectly fit (in a linear sense) with distances between items computed by CATPCA. For the one-dimensional solution, the value of the CPCC is near 0.930, whereas the CPCC is 0.501 for the two-dimensional solution. Hence, the two-dimensional solution is less robust, whereas the one-dimensional solution is more stable and robust.

(b) Cronbach's alpha (see Cronbach, 1951) is based on all items. Pupils are the level of analysis. Cronbach's alpha is defined as (see Meulman et al., 2004):

$\alpha = \frac{m}{m-1} \cdot \frac{\sum \lambda - 1}{\sum \lambda}$: where the number of categories is m and the eigenvalues are λ^2

deviates in both cases from the order that Hart recommends. The levels "consulted but informed" (V5) and "democratic vote" (V6) depict the highest degree of participation, while (supported) self-determination (V7 and V8) shows a slightly lower degree. In addition, the level "pretend participation" (V1) should depict the lowest degree. This result contradicts the theoretical concept but is comprehensible in terms of "involvement" (as synonym for participation). For instance pupils are at least involved even if participation is pretend (V1) in contrast to simply being informed (V4) or not participating at all (V2).

The one-dimensional solution of CATPCA (Figure 3B) refers to the existence of three clearly identifiable groups (see below): "sufficient participation" (V5, V6, V7, V8), "symbolic participation" (V3, pupils can vote, but are not prepared or well informed, it can be assumed that their votes are ignored, too) and "deficient participation" (V1, V2, V4). The items of each group are located close together. This explains the poor fit of MDS (PROXSCAL), which seems to require that the objects are distributed in the whole space. If objects are located close together, this results in errors in MDS, whereas this fact improves the results of CATPCA. CATPCA can be seen as a factor analysis for categorical scaled items. Two or more measures with the same properties (the same location in the space) improve the scaling.

Figure 3A

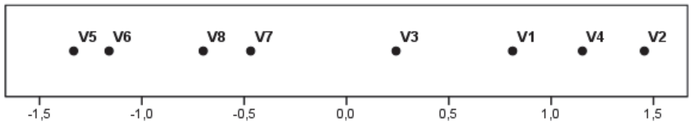


Figure 3B

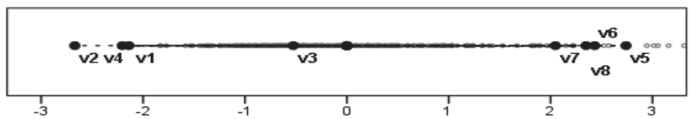


Figure 3C

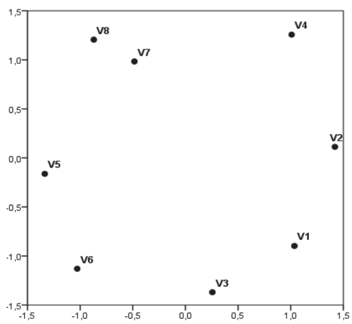
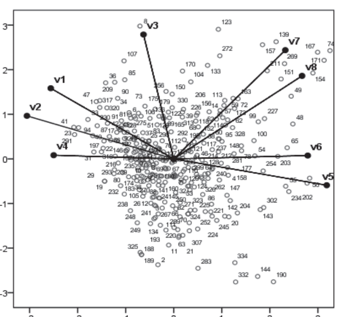


Figure 3D



Note. V1 = pretend participation, V2 = non-participation, V3 = tokenism, V4 = informed, V5 = consulted but informed, V6 = democratic vote, V7 = self-responsibility, V8 = supported self-responsibility

Figure 3 A–D One- and two-dimensional solutions (PROXSCAL (3A and 3C) vs. CATPCA (3B and 3D)) – calibration sample

A further reason for the poor stress is the fact that PROXSCAL uses (in this case) individual distance measures, whereas CATPCA is applied for an aggregated similarity matrix. Hence, individual differences influence the overall fit of MDS less. The substantive reason for these differences is presumably personalized participation (see Wetzelhütter et al., 2013), which depends on different perceptions of participation by pupils.

Apart from that, the (more or less ambiguous) label "uncertain" of the mid response category of the five-option rating scale might have been ambiguous for the respondents and might have caused errors, resulting in measuring a different dimension (see for example Rost, Carstensen & von Davier, 1999).

As an interim conclusion, it can be noted that the results, especially those of CATPCA, support the one-dimensional model rather than the two-dimensional. Arguments for this solution are:

- The one-dimensional model is replicable by different analytical methods (MDS/PROXSCAL and CATPCA) as can be seen, for example, by a high CPCC (which is not the case for the two-dimensional solution).
- The eigenvalue and Cronbach's alpha for the first dimension are acceptable (which is not the case for the second dimension).
- The one-dimensional solution is, in accordance with Hart's concept of a "ladder of participation," substantively interpretable as a degree of participation. Hence, the one-dimensional solution has a theoretical basis.

However, the assumed eight different levels of participation cannot be differentiated. The findings suggest that it is only possible to distinguish three groups of participation.

4.2 Model Validation: Reproduction of Dimensional Space in the Validation Sample

The model validation is carried out by two approaches. First, it is tested as to what extent the above-described solutions are replicable using the Austria-wide data. In a second and final step, relevant content variables are added as criteria variables in order to test content related hypotheses (see below).

4.2.1 Replication of the Model

As before, one- and two-dimensional solutions were computed using MDS (PROXSCAL) and CATPCA in order to determine the dimensionality of the developed scale. Table 4 shows that the results of the validation sample confirm the previous outcome of the calibration sample (cf. Table 3). Moreover, the stress values of the MDS are smaller compared with the calibration sample. Stress 1 fulfills the threshold in the two-dimensional case but Stress 2 still has a value higher than the threshold. The fact that the stress values are higher than the threshold indicates that the MDS approach has problems. Again the reason is that there are groups of items and the items of each group are located close together, which results in errors. CATPCA supports the one-dimensional solution. Once more, the CPCC (CPCC r=0.910) obtained by MDS (PROXSCAL) and CATPCA also supports the one-dimensional solution. Hence, the one-dimensional solution is more robust against the method applied.

From a substantive point of view (see Figure 4A-D), as before, the one- and two-dimensional MDS solutions show a possible distinction at the continuum of the scope of action (from left to right on the x-axis). Therefore, "effective/adequate participation" (left) and "ineffective/inadequate participation" (right) may, regarding the two-dimensional model, be further distinguishable regarding the initiative of involvement (y-axis). The two-dimensional model differs from the earlier one

Table 4 Summarized results for determining dimensionality using PROXSCAL and CATPCA (validation sample)

	PROXSCAL			CATPCA		
Model	Stress 1	Stress 2	Congruence coeff.	Eigenvalue (Variance)	Cronbach's alpha[(b)]	CPCC[(a)]
One-dim. solution	0.289	0.557	0.96	3.2 (40.1%)	**0.787**	**0.910** (p=0.000)
Two-dim. solution	**0.174**	0.440	0.98	1.3 (16.07%)	0.254	0.606 (p=0.001)

Note. (a) n=28; Correlation between distance matrices obtained by the coordinates produced by PROXSCAL and CATPCA. (b) Cronbach's alpha is based on all items (formula for calculation, see Table 3). Pupils are the level of analysis.

Figure 4A

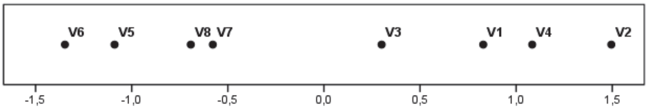


Figure 4B

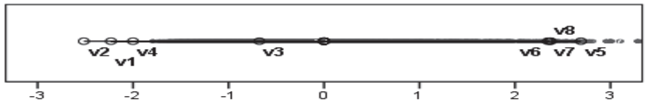


Figure 4C

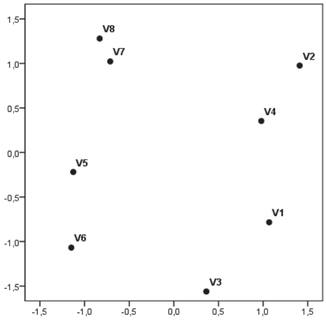
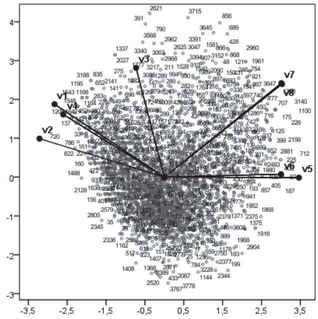


Figure 4D



Note. V1 = pretend participation, V2 = non-participation, V3 = tokenism, V4 = informed, V5 = consulted but informed, V6 = democratic vote, V7 = self-responsibility, V8 = supported self-responsibility

Figure 4 A–D one- and two-dimensional solutions (PROXSCAL (4A and 4B) vs. CATPCA (4C and 4D)) measuring the degree of participation (validation sample)

Table 5 CPCC for one- and two-dimensional distance matrices

Calibration with Validation Sample				
Procedure	Dimensionality	CPCC[(a)]	p	n
PROXSCAL	One-dimensional solution	**0.978**	0.000	28
	Two-dimensional solution	0.838	0.000	28
CATPCA	One-dimensional solution	0.995	0.000	28
	Two-dimensional solution	**0.945**	0.000	28

Note. (a) Correlation between distance matrices of the calibration sample with the validation sample obtained by the coordinates produced by PROXSCAL and CATPCA

presented in two points. Firstly, the items v5 and v6 can hardly be differentiated concerning the scope of action. Secondly, the items v4 and v2 are switched.

Once again, CATPCA differentiates between three groups: "sufficient participation" (V5, V6, V7, V8), "symbolic participation" (V3), and "deficient participation" (V1, V2, V4).

As described in Section 3.2.2, the CPCC was computed in order to measure the similarity between the calibration and the validation sample. In our case, distance matrices were calculated for each sample (calibration and validation) based on the representation in a one- and two-dimensional space. Table 5 shows that the distance matrix of the validation sample represents the similarity structure of the calibration sample almost perfectly (MDS: CPCC=0.978; CATPCA: CPCC=0.995) in the one-dimensional case. For the two-dimensional case, the CPCC is good (MDS: CPCC=0.838; CATPCA: CPCC=0.945), but the coefficients are smaller. Consequently, this result also reinforces the one-dimensional solution.

To summarize, the one-dimensional solution is supported by:

- Substantive interpretation: the results suggest that three levels of participation can be distinguished
- Model robustness: the one-dimensional MDS (PROXSCAL) solution is replicated by a different analytical method (CATPCA)
- Model replication: the one-dimensional solution can be reproduced with the validation sample. The CPCC is nearly perfect.

Therefore, the outstanding analyses are based on the one-dimensional solution.

4.2.2 Criteria Validation

In a second and final step, relevant content variables are added as criteria variables. It is assumed that they are connected with specific levels of participation. It is expected that the predictors will be positioned next to those items which are influenced or rather determined by their occurrence. Concurrently, the basic (one-dimensional) model should not significantly change due to this strategy.

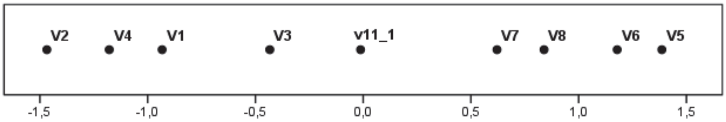
Figure 5 A–B shows the results for the first hypothesis. As expected, co-decision of school rules (v11_1) is located next to "Tokenism," which means that pupils are allowed to vote but are badly prepared concerning the topic. This may be caused for example by the aim to simplify such procedures. The structure remains basically unchanged, so the grouping of the items in the three groups is not affected.

In accordance with the second hypothesis, "co-decision to decorate the classroom" (V11_2) is placed next to the item which implies an (nearly) "equal" involvement in decision making (V5). In fact, the item V11_2 is, independently of the position of V5, placed next to it. This result reinforces the stability of the solution (see Figure 6 A–B). Only the items V4 and V1 are interchanged in the validation sample – the structure remains basically unchanged. The grouping of the items in the three groups is not affected.

In addition, we computed the mean of the two criteria variables in dependence of the scale values (object quantifiers, obtained by CATPCA) of the pupils. If the scale measures different degrees of participation, a higher scale value should correspond to a stronger agreement in the presence of the two criteria variables at school. The results are shown in Table 6. For significance testing, the multilevel structure was taken into account using the SPSS module MIXED (IBM-SPSS, 2011a). The results confirm our interpretation. A higher degree of participation on the scale corresponds to a more positive evaluation of presence of the two criteria.

To sum up, the structure of the model remained basically unchanged regardless of the used data set. The above-mentioned hypotheses were confirmed, as the included criteria variables are placed as expected next to related items. In addition, a clear correspondence exists between the derived scale values and the criteria variables.

Figure 5A



Note. Stress 1: 0.40; Stress 2: 0.76; congruence coefficient: 0.92

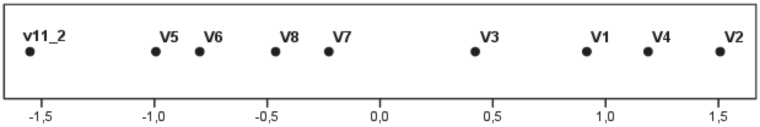
Figure 5B

Note. Stress 1: 0.38; Stress 2: 0.73; congruence coefficient: 0.92

v11_1 = co-decision of school rules, V1 = pretend participation, V2 = non-participation, V3 = tokenism, V4 = informed, V5 = consulted but informed, V6 = democratic vote, V7 = self-responsibility, V8 = supported self-responsibility

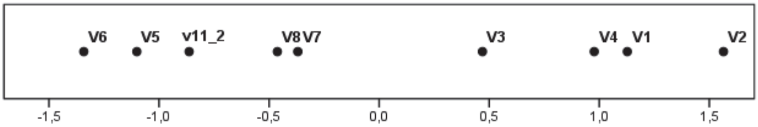
Figure 5 A–B One-dimensional solution (PROXSCAL) in connection with context variable I (calibration sample (5A) vs. validation sample (5B))

Figure 6A



Note. Stress 1: 0.34; Stress 2: 0.65; congruence coefficient: 0.94

Figure 6B



Note. Stress 1: 0.34; Stress 2: 0.64; congruence coefficient: 0.94

V11_2 = co-decision to decorate the classroom, V1 = pretend participation, V2 = non-participation, V3 = tokenism, V4 = informed, V5 = consulted but informed, V6 = democratic vote, V7 = self-responsibility, V8 = supported self-responsibility

Figure 6 A–B One-dimensional solution (PROXSCAL) in connection with context variable II (calibration sample (6A) vs. validation sample (6B))

Table 6 Means of criteria variables in dependence of scale values of pupils (obtained by CATPCA)

	V11_1		V11_2	
Degree of participation	Mean	p	Mean	p
low (scale value below -1)	4.80 (n=604)	Reference Group	3.44 (n=604)	Reference Group
middle (scale value between -1 and 1)	4.42 (n=2539)	0.000	2.85 (n=2539)	0.000
high (scale value larger than 1)	4.04 (n=606)	0.000	2.37 (n=606)	0.000

Note. V11_1 = co-decision of school rules, V11_2 = co-decision to decorate the classroom, Scale values from 1 = always (strongly present) to 5 = never (not present),
Significance from using multilevel model (MIXED), random effects for schools were assumed

5 Summary and Conclusions

In the past few years, a stronger scientific focus on childhood, youth, and educational research on participation has taken place, and Hart's Ladder of Participation (1992, 2008) is frequently used in this context. Hart distinguishes eight levels of participation, each of which is characterized by a different degree of involvement and self-responsibility of children. No developmental process where a previous level must be successfully passed before the next level can be approached is assumed. From a methodological point of view, Hart assumes a one-dimensional unfolding model.

This paper addresses the question of whether participation in school can be measured one-dimensionally, as assumed by Hart. In order to answer the question, a rating scale was developed measuring those eight steps of the ladder from a low (pretend participation) to a high (supported self-responsibility) degree of participation and implemented in an Austrian study on school participation. MDS procedure (PROXSCAL) and CATPCA were applied for data analysis.

The results suggest that participation in school can be measured with a one-dimensional scale in accordance with Hart. The one-dimensional solution reveals higher robustness towards different data analysis and higher consistency as measured by Cronbach's Alpha than a two-dimensional solution. In addition, the one-dimensional model is reproduced with the validation sample, and remained unchanged when validation criteria are included and tested hypotheses with criteria variables are confirmed.

However, in contrast to Hart (1992) we have not been successful in distinguishing eight different levels of participation. The results reveal three groups of participation: “sufficient participation,” “symbolic participation,” and “deficient participation.”

Two primary reasons may be considered as explanations for this outcome. Firstly, the result may reflect the fact that democracy in school is less evident than assumed (see for example Wetzelhütter et al., 2013), so that only three groups can be distinguished. Secondly, the measurement instrument may have been deficient and too insensitive, not able to differentiate the assumed different degrees of participation.

In addition, the content validity of several items can be placed into question. Hart’s ladder (2008), for example, positions “pretend participation” at the bottom (Level 1) and “supported self-responsibility” at the top (Level 8). Empirically, “pretend participation” (Level 1) is placed above, for example, “informed participation” (Level 4) and “non-participation” (Level 2), while “supported self-responsibility” (Level 8) is placed below, for example, “democratic vote” (Level 6). This suggests that, from a pupil’s point of view, any form of participation, even pretense (Level 1), implies a higher degree of participation than just being informed (Level 4) or not participating at all (Level 2).

Hence, a first conclusion of the study is to try to improve the instrument in order to be able to distinguish between the two mentioned explanations (only three groups of participation are seen by school children versus the instrument lacks sensitivity).

With reference to data analysis, CATPCA (categorical principal component analysis) seems to be more appropriate than MDS. In addition, the former has the advantage that the user must specify fewer parameters to which no clear guidelines exist.

Despite the above-described weakness of the proposed instrument, it can already be used in order to test theories about the influence of participation in school on the individual, class, and school level and to improve school practice in order to enable the productive development of schoolchildren.

References

Arnstein, S. (2011[1969]). A ladder of citizen participation. In R. T. LeGates & F. Stout (Eds.), *The City Reader*. Fifth edition (pp. 238-250). New York, NY: Routledge.

Altrichter, H., Bacher, J., Langer, R., Gamsjäger, M. & Wetzelhütter, D. (2012). *Schüler/innen entwickeln Schule?! – Endbericht*. Linz: Johannes Kepler Universität.

Bacher, J. (2009). Analyse komplexer Stichproben. In M. Weichbold, J. Bacher, J. & C. Wolf (Eds.), *Umfrageforschung. Herausforderungen und Grenzen*. Österreichische Zeitschrift für Soziologie. Sonderheft 9/2009 (pp. 253-272). Wiesbaden: VS Verlag.

Bacher, J., Gerich, J., Lehner, R., Straßmair, M. & Wilk, L. (1999). Auswirkungen kommunaler Jugendbeteiligungsprojekte. In C. Friesl, M. Hahn, B. Heinzlmaier & C. Klein (Eds.), *Erlebniswelten + Gestaltungsräume* (pp. 213-238). Graz & Wien: Verlag Zeitpunkt.

Bacher, J., Pöge, A. & Wenzig, K. (2010). *Clusteranalyse. Anwendungsorientierte Einführung in Klassifikationsverfahren.* 3. Auflage. München: Oldenbourg Verlag.

Blasius, J. & Thiessen, V. (2012). *Assessing the Quality of Survey Data.* Chichester, UK: Sage.

Borg, I., Groenen, P. J. F. & Mair, P. (2013). *Applied Multidimensional Scaling.* Heidelberg: Springer Verlag.

Borg, I. & Groenen, P. J. F. (2005). *Modern Multidimensional Scaling. Theory and Application. Second edition.* New York, NY: Springer.

Bryman, A. & Cramer, D. (2009). Constructing variables. In M. Hardy & A. Bryman (Eds.), *Handbook of Data Analysis* (pp. 17-34). London: Sage Publications, Inc.

Cherkassky, V. & Mulier, F. (2007). *Learning from Data. Concepts, Theory, and Methods.* Second edition. Hoboken, NJ: Wiley & Sons, Inc.

Coombs C. H. (1964). *A Theory of Data.* New York, NY: Wiley & Sons, Inc.

Cox, T. F. & Cox, M. A. A. (2001). *Multidimensional Scaling.* Second edition. Boca Raton, FL: Chapman & Hall/CRC.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16 (3), 297-334.

De Leeuw, J. (2005). Multidimensional Unfolding. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science.* Vol. 3 (pp. 1289-1294). New York, NY: Wiley & Sons, Inc.

Gediga, G. (1998). *Skalierung: Eine Einführung in die Methodik zur Entwicklung von Test- und Messinstrumenten in den Verhaltenswissenschaften.* Münster: Lit Verlag.

Green, R. J. (2010). *Coastal Towns in Transition. Local Perceptions of Landscape Change.* Dordrecht: Springer.

Greenacre, M. J. (2013). Correspondence Analysis. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods. Statistical Analyses.* Vol. 2 (pp. 142-153). New York, NY: Oxford University Press.

Hart, R. A. (1992). *Children's participation. From tokenism to citizenship.* Innocenti essays, No. 4. Florence: UNICEF, United Nations Children's Fund.

Hart, R. A. (2008). Stepping back from 'the ladder': Reflections on a model of participatory work with children. In A. Reid, B.B. Nikel & V. Simovska, (Eds.), *Participation and Learning. Perspectives on Education and the Environment, Health and Sustainability* (pp. 19-31). Dordrecht: Springer.

Hoffman, D. L. & de Leeuw, J. (1992). Interpreting multiple correspondence analysis as a multidimensional scaling method. *Marketing Letters,* 3 (3), 259-272.

IBM-SPSS (2011a). *IBM-SPSS. Advanced Statistics 22.* Chicago: IBM-SPSS.

IBM-SPSS (2011b). *IBM-SPSS. Categories 22.* Chicago: IBM-SPSS.

James, A., Jenks, C. & Prout, A. (1998). *Theorizing Childhood.* Cambridge: Polity Press.

Kish, L. (1965). *Survey Sampling.* New York, NY: Wiley & Sons, Inc.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29 (2), 115-129.

Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling. Sage University paper series on quantitative applications in the social sciences.* Number 07-11. Beverly Hills: Sage Publications, Inc.

Lewin, K. & Lewin, G. (1982 [1941]). Democracy and school. *Understanding the child, 10,* 7-10. (KLW 6; 285-291, "Demokratie und Schule").

Lewin, K., Lippit, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology, 10*, 217-299.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140 (22),* 5-55.

Linting, M., Meulman, J. J., Groenen, P. J. F. & van der Kooij, A. J. (2007). Nonlinear principal components analysis: Introduction and application. *Psychological Methods, 12 (3),* 336-358.

Meulman, J. J. & Heiser W. J. (2011). *IBM SPSS Categories 20.* IBM: https://www.csun.edu/sites/default/files/categories20-64bit.pdf

Meulman J. J., van der Kooij A. J. & Heiser W. J. (2004). Principal component analysis with nonlinear optimal scaling transformations for ordinal and nominal data. In D. Kaplan (Ed.), *Handbook of Quantitative Methodology for the Social Sciences* (pp. 49-70). Thousand Oaks, CA: Sage Publications, Inc.

Nunnally, J. C. (1978). *Psychometric Theory.* Second edition. New York, NY: McGraw-Hill.

OeAD (2014). The Austrian Educational System. https://www.oead.at/welcome_to_austria/education_research/the_austrian_education_system/EN/

Oser, F., & Biedermann, H. (2006). Partizipation – ein Begriff, der ein Meister der Verwirrung ist. In C. Quesel & F. Oser (Eds.), *Die Mühen der Freiheit: Probleme und Chancen der Partizipation von Kindern und Jugendlichen* (pp. 17-37). Zürich: Ruegger.

Romesberg, H. C. (2004). *Cluster Analyses for Researchers.* Place: NC: LULU Press.

Rost, J., Carstensen, C. H. & von Davier, M. (1999). Sind die Big Five Rasch-skalierbar? Diagnostica, 45 (3), 119-127.

Sturgis, P. (2004). *Analysing Complex Survey Data: Clustering, Stratification and Weights.* Social Research Update, Issue 43. Guildford: University of Surrey.

Van Schuur, W. H. & Kiers, H. A. L. (1994). Why factor analysis often is the incorrect model for analysing bipolar concepts, and what model to use instead. *Applied Psychological Measurement, 18 (2)*, 97-110.

Wetzelhütter, D., Paseka, A. & Bacher, J. (2013). Partizipation in der Organisation Schule aus der Perspektive der Schülerinnen und Schüler. In S. Weber, M. Göhlich, A. Schröer & C. Fahrenwald, (Eds.), *Organisation und Partizipation. Beiträge der Kommission Organisationspädagogik* (pp. 157-166). Wiesbaden: VS Verlag.

Wilk, L. & Bacher, J. (Eds.). (1994). *Kindliche Lebenswelten. Eine sozialwissenschaftliche Annäherung.* Opladen: Leske + Budrich.

Young, F. W. (1985). Multidimensional scaling. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences*, Vol. 5 (pp. 649-659). New York, NY: Wiley & Sons, Inc.

Zinnecker, J. (1999). Forschen für Kinder – Forschen mit Kindern – Kinderforschung. Über die Verbindung von Kindheits- und Methodendiskurs in der neuen Kindheitsforschung zu Beginn und am Ende des 20. Jahrhunderts. In M. Honig, A. Lange, & H.-R. Leu (Eds.), *Aus der Perspektive von Kindern? Zur Methodologie der Kindheitsforschung* (pp. 69-80). Weinheim & München: Juventa.

Appendix A

Table A1 descriptive results: participation-scale

	fully applies	largely applies	uncertain	does not apply much	does not apply at all	n [a]
Participation Item (level at the scale)			Calibration sample			
V1: Pretend Participation (Level 1)	6.0	13.7	33.7	34.9	11.6	335
V2: Non-participation (Level 2)	10.7	21.2	29.6	25.7	12.8	335
V3: Tokenism (Level 3)	3.9	23.3	36.4	27.8	8.7	335
V4: Informed participation (Level 4)	5.7	21.8	32.5	24.2	15.8	335
V5: Consulted but informed (Level 5)	11.9	33.7	33.1	14.9	6.3	335
V6: Democratic vote (Level 6)	12.2	34.3	26.0	17.6	9.9	335
V7: Self-responsibility (Level 7)	11.3	30.1	34.0	19.1	5.4	335
V8: Supported self-responsibility (Level 8)	13.1	36.7	28.4	14.3	7.5	335
Participation Item (level at the scale)			Validation sample			
V1: Pretend Participation (Level 1)	7.2	22.6	28.4	33.8	8.0	3779
V2: Non-participation (Level 2)	14.2	23.0	29.0	22.6	11.2	3779
V3: Tokenism (Level 3)	3.5	17.5	31.2	34.6	13.3	3779
V4: Informed participation (Level 4)	7.8	25.0	30.6	26.6	10.1	3779
V5: Consulted but informed (Level 5)	6.4	33.0	29.7	22.8	8.2	3779
V6: Democratic vote (Level 6)	8.3	27.0	27.9	22.1	14.7	3779
V7: Self-responsibility (Level 7)	9.2	30.5	30.7	21.5	8.1	3779
V8: Supported self-responsibility (Level 8)	11.9	32.8	27.1	19.2	9.1	3779

Note. (a) Exclusively cases with missing values or without variation in the response