



# Education Quarterly Reviews

**Metsämuuronen, Jari. (2019), Educational Assessment and Some Related Indicators of Educational Equality and Equity. In: *Education Quarterly Reviews*, Vol.2, No.4, 770-788.**

ISSN 2621-5799

DOI: 10.31014/aior.1993.02.04.105

The online version of this article can be found at:  
<https://www.asianinstituteofresearch.org/>

Published by:  
The Asian Institute of Research

The *Education Quarterly Reviews* is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research, and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of Education.



ASIAN INSTITUTE OF RESEARCH  
Connecting Scholars Worldwide

# Educational Assessment and Some Related Indicators of Educational Equality and Equity

Jari Metsämuuronen<sup>1,2</sup>

<sup>1</sup>Finnish National Education Evaluation Centre (FINEEC), Finland

<sup>2</sup>NLA University College, Bergen, Norway

Correspondence: Jari Metsämuuronen, Finnish National Education Evaluation Centre (FINEEC), P.O. Box 28 (Mannerheiminaukio 1 A, 6th Floor), FI-00101 HELSINKI. Tel: +358 400 579 848 E-mail: jari.metsamuuronen@gmail.com

## Abstract

The article asks three questions of equality in education. First, what kind of equality we are talking about - numerical equality, proportional equality, formal equality, or moral equity? Second, how to categorize the indicators in a practical manner? Third, how to use the equality indicators based on the national level assessment of student achievement? The article starts with a theoretical treatment and gives a framework for further studies of educational indicators. The indicators are divided into absolute- and relative indicators as well as binding- and preferable indicators. In the empirical part, 26 parity indicators are introduced and discussed based on national level datasets from student achievement.

**Keywords:** Equality, Equity, Student assessment, Educational indicators, Parity

## 1. Concepts of Equality and Equity

### 1.1 Complexities in Theorizing Equality

Equity is a value to do with justness and fairness; it is generally accepted that all the humans should be treated in a just and fair manner as implied also in the *Universal Declaration of Human Rights* (UN, 1948; see also Arneson, 2014). This could be called also as the ‘basic equality’ (see Nathan, 2014; Arneson, 2014). Within the educational realm, we may talk about educational equity. Levin (2010) defines educational equity as fairness in access to opportunities to benefit from education. While equity is a value and principle, equality, on its behalf, is the degree or state of being equal especially with status, rights and opportunities. Equality makes sense if the equity is accepted as a value and principle.

Though all persons, by virtue of being persons, have equal basic dignity and worth, it may be good to note some obvious oddities in equality. First, the world is *not equal* in absolute sense. Two non-identical objects are never

completely equal (Gosepath, 2007; Tugendhat & Wolf, 1983, p. 170). There really are differences, which are a basis for uniqueness, which supposedly should be taken as a good thing; ‘equality’ needs to be distinguished from ‘identity’ and it needs to be distinguished from ‘similarity’. Thus, to say that children are equal is not to say that they should be identical. On the other hand, ‘equality’ implies ‘similarity’ rather than ‘sameness’ (Gosepath, 2007).

Second, it seems that some persons just happen to be born under a *luckier* star than some other. The same input circumstances may lead to totally different outcome results depending on random acts. Is this randomness injustice or inequality or just “luck”? On the other hand, the “happen to be born”-argument cannot be used as rationale for inequality between races, castes, sexes, or between disabled and “abled” child (see Hurley, 2001; Arneson, 2001). Why? Because when two persons have *equal* status in at least one normatively relevant respect (such as being citizens in the same country), they must be treated equally regarding this respect – “treat like cases as like” (Aristotle, *Nicomachean Ethics*, V.3. 1131a10-b15; *Politics*, III.9.1280 a8–15, III. 12. 1282b18–23).

Third, it seems that some people *use their capacity* and recourses *better* than others (of the discussion of capabilities and equality, see Sen, 2005; Rawls, 1971, pp. 60–65). Even if we could, magically, organize the world so that all the persons would have totally equal education and knowledge base, it would lead to a situation, within a year, that some of us have learnt more and some would have lost part of their knowledge. As individuals, we just have different motivational structures, interests, talents or skills, and, for example, mental capacities, which, evidently, has something to do the seemingly unequal end-product for individual students.

Fourth, in the real life, we just do not follow the rule of ‘absolute equality’ between humans, because of differences in the individual capabilities, resources, responsibilities, powers, or individual characteristics such as motivation, interests, or talents. This can be seen in our normal family life: in our families the small children are not given equal possibilities to affect the family decisions compared with the parents – maybe because of a fear of under-developed sense of consequences of the decisions. When the children grow older they get more responsibilities and possibilities.

Fifth, it’s worth noting Levin’s (2010, p. 4) note that “*There is no natural state of educational equity, but one defined by each society on the basis of its values and the imperative that it sets for the issue as a moral commitment. ... Each society may define educational equity in different ways and use different criteria for assessing equity in the overall population and among different groups.*” Hence, it is difficult to claim that, in every society, there should be equal values behind the educational equity. However, in the modern globalized world, it’s difficult to claim seriously, for instance, that in our society the boys are valued so much more over the girls that we are not willing to give a proper education for girls. The same can be said with some disadvantage group based on family background, religion, or place of birth. This issue is discussed deeper when introducing the binding indicators for educational equity.

Guiton and Oakes (1995) showed the intimate relation between values on equality and measures of equality. In the educational settings, all the children should have equity to be treated fairly even though there may be inequality in their family background. On the other hand, equality does not mean much if there is not equity; just after the equity is guaranteed as a principle and practice, equality make sense. It is notable that some of the global actors, such as OECD and UNESCO are prone to use term “equity” instead of “equality” in their background papers and recent reports (OECD 2013a; Sherman & Poirier, 2007) though, in many cases, “equality” may also be an appropriate concept. Pupil-teacher ratio, for example, can be taken as an indicator for *equity* (see Sherman & Poirier, 2007, p. 58) – the students should have equal opportunities for the decent class sizes with a decent number of contacts with the teacher. Gender parity, for example, on its behalf is a typical example of an indicator for *equality*; it makes sense after confessing that boys and girls should be treated equally.

### 1.2 What kind of equality we then are seeking in education?

When we refer to equality, what *kind* of equality we actually are talking about? Asks Rae with colleagues (1981, p. 132), equal in *what* respect? Continues Temkin (1993), what *is* equality and what is *inequality*? Gosepath (2007) discusses – based on Dann (1975, p. 997), Menne (1962), and Westen (2014, pp. 39, 120) – the terms ‘equality’, ‘identity’, and ‘similarity’. ‘Equality’ signifies correspondence between a group of different objects, persons, processes or circumstances that have the same qualities in at least one respect, but not all respects, i.e., regarding one specific feature, with differences in other features. Thus, ‘equality’ needs to be distinguished from ‘identity’, which refers to differences between individuals regarding one or more specific features, and it needs to be distinguished from ‘similarity’ which approximates ‘correspondence’.

The classic theorists of equality (Aristotle, Plato, Lock, Rousseau, Kant) have divided equality and justice into several categories though Rae and colleagues (1981, 132) noted that in any real historical context, no single notion of equality is superior over the others:

- (1) numerical equality (“give all the same load”) (Aristotle, *Nicomachean Ethics*, 1130b–1132b; cf. Plato, *Laws*, VI.757b–c),
- (2) proportional equality (“give all what they can carry”) (Aristotle, *Nicomachean Ethics*, 1130b–1132b; cf. Plato, *Laws*, VI.757b–c),
- (3) formal equality (“give the like cases the same load”) (Aristotle, *Nicomachean Ethics*, V.3. 1131a10– b15; *Politics*, III.9.1280 a8–15, III. 12. 1282b18–23), and
- (4) moral equity (“everyone deserved the same dignity and the same respect”) (Locke, 1690; Rousseau, 1755; Kant, 1785; 1797).

A treatment is *numerically* equal when it treats all persons as *indistinguishable*. That is not always just. A blunt example of an unjust numerical equality in the educational settings can be as follows: Because the 15 years old differs from 6 years old remarkably in perseverance, for example, we can require more of 15 years old than 6 years old. Hence, it might be unjust to require a six years old school-child be apt 8 hours in school, while we can require it from 15 years old teen. Same hours would be numerically equal but not fair for the younger child. In the educational settings, we can claim that all the students at the *same* grade could be treated numerically equal if there is no other rationale (such as retardation or serious learning difficulty) for expecting different output from different students.

A treatment is *proportional* and *formally* equal when it treats all relevant persons in relation to their due. When factors speak for unequal treatment or distribution, because the persons are unequal in relevant respects, the treatment or distribution proportional to these factors is just (Gospath, 2007). Unequal treatment or distribution must be considered proportionally: that is the prerequisite for persons being considered equally. In the educational settings, we can claim that all the students at the *different* grade should be treated proportionally equal and students with *the same* grade as *formally* equal. Proportional equality is relevant also when expecting different output from different students at the same grade if there is relevant rationale for that (such as retardation or serious learning difficulty). The differences between proportional equality (“give all what they can carry”) and formal equality (“give the like cases the same load”) is somewhat dim. According to Gospath (2007), the formal postulate remains quite empty if it remains unclear when or through what features two or more individuals should be considered equal. When seeking justice, one need to discuss which cases are equal and which unequal (Aristotle, *Politics*, 1282b 22). In the educational settings, we usually think that the students at the same grade should be treated as ‘like’ cases and the students at the same grade are assessed according to same rules. Hence, we seem to seek formal equity within the grades. However, some students at the same age cohort, shown very low performance, are sometimes given a possibility to be assessed according to a “personal” curriculum which means that the requirements are remarkably lowered. In the case, we seem to seek the proportional equality.

### ***1.3 Conclusion: on what ground we should seek the equality in education?***

We may agree that all children have the moral right to get equal opportunities even though all children are different and unique. We are thus willing to give all children equal possibilities to be what they would be. It is the same as all the children should have equal moral rights to get clean water to drink. This political will of the government (and thus the society as a whole) is recorded in laws and acts and it usually operationalized in the national curricula(-um) for different subjects and topics.

It seems that, in the practical educational settings, we are willing to see the *numerical equality in some extent*: for example, we expect the equal length of the school year within a country from all schools and the numerically equal minimum number of teaching hours to be attended. However, in most cases, we expect to find *proportional and formal equality on the basis on moral equity*. That is, in each age group all the children should be given equal opportunities to learn because learning is one of the basic rights for the children.

The question is, how to measure the possible inequality in the educational system. The latter part of the article handles practical indicators for parity: In the empirical part, these are divided into four categories based on two dimensions: “binding”/ “quality” indicators and “absolute”/ “relative” indicators. The practical part introduces 26 relative preferable indicators for educational equality based on national student achievement.

## **2. Educational indicators – from Chemistry to the global actors**

### ***2.1 Concepts related with Indicators***

The concept of ‘indicator’ has been borrowed from Chemistry to the other fields such as Social Sciences and Econometrics. In Chemistry, an indicator is a substance which show characteristic change in its color when comes in contact with acid or base. Then, an indicator is used to determine the degree of acidity or basicity of any solution. Generalizing the idea: an indicator is “a measurement or value which gives you an idea of what something is like” (Collins Cobuild, 2014). Hence, an educational indicator is a pointer which shows the state of art in the educational realm. Indicators are characterized primarily by the fact that they provide information in summary form, are communicable and are subject to relative consensus (Delorme & Chatelain, 2011, p. 8). On the top of the different Philosophical domains, the theoretical framework for using indicators in the socio-economical settings and policy making comes either from the Management- or Econometric theories. Notes Delorme and Chatelain (2011), the concept of indicators can be linked to objective-based planning related with the Management domain. On the other hand, Guiton and Oaks (1995), based on Guiton (1992), have connected the equality with the econometrical theories: Libertarian, Liberal, and Democratic Liberal conceptions. These theories are not discussed here, though. This article relies more of the ideas coming from Philosophy.

Within the social sciences, the concepts regarding the indicators, indices, statistics, and datasets are not always clear. Mannis (2014) has condensed the different levels of data for policy purposes as the Information Pyramid. The fundamental basis of the pyramid are *data*; when unprocessed, data are of little value for policy purposes. If data are processed into *statistics* or *tables*, they can be used in reports, but still they can be difficult to use for policy. *Indicators* are statistics directed specifically towards policy concerns; they point towards successful outcomes for policy. These are usually highly aggregated and have easily recognizable purposes. Claims Mannis, the classical socio-economic indicators, such as the unemployment rate or Gross Domestic Product (GDP) growth, are “*numbers which are such powerful and recognizable indicators of performance that they may cause governments to fall*”. At the highest level in the pyramid are *indices*, such as the gender parity index or human development index, which combine different indicators into a single number useful for comparison over time and space. Nevertheless, though it may be possible, in theory, to separate indices from indicators, it seems that the indices and indicators may overlap each other; in all cases, the indices can be used as indicators and, in many cases, indicators could be called as indices. In this article, the two are handled as unseparated.

## **2.2 Educational Indicators in the Use of International Stakeholders**

In literature, the indicators are categorized various ways. Delorme and Chatelain (2011), for example, use the term ‘performance indicators’; they categorize the indicators as Input-, Output-, Outcome-, Impact-, and Context indicators. Mannis (2014) categorizes indicators as Driving Force Indicators (such as real GDP per capita), State Indicators (GDP per capita in dollars), and Response Indicators (Investment share in GDP in percentages). Also, three great players in the global educational field, World Bank (WB), Organization for Economic Cooperation and Development (OECD), and United Nations (UN) and specifically United Nations Educational, Scientific and Cultural Organization (UNESCO) have developed their own indicators for pointing the state of art of Education. World Bank indicators for the Education are divided into indicators for Efficiency, Inputs, Outcomes, and Participation (see <http://www.indexmundi.com/facts/topics/education>).

The numerous list of OECD indicators in the *Education at a Glance* (OECD, 2014b) are classified, on one hand, on the basis of actors in education systems (individual learners and teachers, instructional settings and learning environments, education service providers, and the education system as whole) and, on the other hand, according to whether they address learning outcomes for individuals or countries, policy levers or circumstances that shape these outcomes, or to antecedents or constraints that put policy choices into context (see OECD, 2014b, p. 17).

The 52 educational indicators of UNESCO are classified into six baskets in the framework of Education for all (EFA): Early childhood care and education (ECCE), Universal primary education (UPE), Lifelong learning and life skills, Adult literacy and basic education, Gender equality, and Quality of Education (see UNESCO, 2011). Hence, frameworks for educational indicators and indices are many, which makes the comparison of the outcomes difficult if not impossible. On the other hand, a rich bank of indicators may enrich the view of the status of the educational settings. It is notable that, generally, the above-mentioned educational indicators of WB, OECD, and UNESCO do not include the learning outcomes as indicators for the system.

## **2.3 Educational indicators based on learning outcomes**

The quality aspect of the education – especially the learning outcomes – interests us from the international comparison viewpoint. Though the sets of educational indicators of the global actors, seen above, do not include indicators based on learning outcomes, the average student achievement is used widely as an indicator of the systems. The new results from PISA (Programme for International Student Assessment, e.g., OECD 2013a; 2016a), TIMSS (Trends in International Mathematics and Science Study, Martin et al, 2012; Mullis, Martin, Foy, & Arora, 2012; Mullis, Martin, Foy, & Hooper, 2016), PIRLS (Progress in International Reading Literacy Study, Mullis, Martin, Foy, & Drucker, 2012; Mullis, Martin, Foy, & Hooper, 2017), or PIAAC (Programme for the International Assessment of Adult Competencies, OECD, 2013b; 2016b) are keenly waited news (or feared, depending on ranking) in many ministries of education, editorial boards of newspapers, and research centers. There seems to be an interest – maybe even an over-interest – in comparisons of the learning outcomes in the participating countries. The main indicator of the successful educational policy seems to be the ranking of the countries though the background information of the students, teachers, and principals gives us a firm educational context for the results. The ambitious goals to be at the best sequence of the ranking and, apparently, the seemingly low results in certain countries have raised worried voices against PISA type of global tests (see, *Guardian*, 6 May 2014, an open letter to Andreas Schleicher – notably, over 80% of the undersigned writers came from USA). One of the worries in the open letter may be worth highlighting: in many countries, because of the ranking in the international comparison, the governments have begun to amend the educational practices; countries are overhauling their education systems in the hopes of improving their rankings.

PISA results include interesting and valuable opening to glance equity in educational opportunities. First time, the equity matters are handled in a full report. The treaty in *PISA 2012 Results: Excellence through Equity* is largely based on the viewpoint of how school systems allocate their resources for education and how that allocation is related to student or school characteristics, such as socio-economic status, immigrant background or school location is a good start in raising the treatment into the next level from ranking lists to a more practical

policy issues (OECD 2013a, p. 28). Hence, the educational equity seems to be geared toward the economical perspective. However, it is notable that though the basic viewpoint in the report is economical, it allows such extensions as school characteristics, family structure, immigrant background, language spoken at home. Other volumes of PISA results include such equity discussion as gender parity in learning (OECD, 2014a), attitudes, behavior and approaches to learn which are associated with inequities in the acquisition of knowledge and skills (OECD, 2013c), and policies and practices adopted in schools and school systems and their relation to performance and equity (OECD, 2013d).

The equity indicators of OECD draw heavily from Levin (2010, see OECD, 2013a, p. 17). Basing on educational attainment and educational achievement Levin categorizes the types of equity into the baskets of educational access, educational resources, educational processes, intermediate outcomes, and educational outcomes (Levin, 2010, p. 5 ff.). Levin sees the *educational access* as a necessary indication of equity. The *educational resources* include such elements for equity indicators as teachers and leadership talent, facilities, instructional materials or the number of days and hours of instruction in a school year as well as availability of textbooks, computers and decent class sizes. The *educational processes* include such elements as types of programs, curriculum, and instruction that the schools offer. The *intermediate outcomes* include indicators like failure and grade repetition rates. Levin also identifies general groups of students for whom we are expecting to see the educational equity (Levin, 2010, p. 8 ff.): Typical groups whose educational status differs from other groups and the educational mainstream are: gender, socioeconomic status, ethnicity, race, immigrants, language, region, and disability. Notes Levin, it is important to note that students are often found in multiple categories that exacerbate their educational challenges.

Based on the rich database in TIMSS and PISA type of inquiries, it would be possible to expand the analysis of the equity and equality in the educational settings even with the existing database. In what follows, several indicators are suggested. Some of them are already introduced above and some more are introduced and suggested. Some of the indicators above can be classified strictly as *equity indicators*; they measure whether the systems can guarantee equal opportunities for all citizens. Some other indicators may be more *equality indicators*; they measure the equality and parity between different groups based on the student achievement. The latter treaty here concentrates on equality indicators though relevant equity indicators are kept in discussion. Before suggesting some further possibilities for equality indicators, a simple framework is introduced for categorizing the indicators.

### 3. Equality Indicators based on the (Inter)National Assessment of Student Achievement

#### 3.1 Absolute and relative indicators for educational equality

In the simple framework for categorizing the equality indicators in education, two dimensions are used in classifying the indicators into absolute/relative set of indicators and binding/preferable set of indicators (Table 1). The absolute and relative are clear as concepts. The binding and preferable sets of indicators may need some explanation. The binding indicators ask: “are we doing what we *should* do” and the preferable indicators ask: “are we doing what we should do with *good results*, providing all citizens *equal possibilities* with *efficient processes* with *economically sustainable way*”. It is worth noting Levin’s (2010) claim that there is no natural state of educational equity, but one defined by each society based on its values and the imperative that it sets for the issue as a moral commitment. Hence, the “binding” set of indicators may vary between the societies. However, there are over-cultural values shared with different societies which may be used as a basis for comparing the educational possibilities in different countries. One set of these is the UN universal declaration of human rights (UN, 1948). Though these values are ratified by most countries, it has faced critique from the Islamic countries as well as Asian countries because of a possible Western orientation (of the critique, see condensed in [http://en.wikipedia.org/wiki/ Universal\\_Declaration\\_of\\_Human\\_Rights](http://en.wikipedia.org/wiki/Universal_Declaration_of_Human_Rights)). The bottom line is, however, that there *are* universal human values and beliefs on which it is possible to build some objective criteria for “binding” indicators. One of these is the common basic worth and value of each individual.

**Table 1. A schematic division of the types of indicators**

		Absolute indicators			
		Absolute binding	Absolute preferable		
Binding indicators		<ul style="list-style-type: none"> <li>Schools/teaching available in a decent distance</li> <li>Students' and teachers' absence/presence in the schools</li> <li>Number of schooldays follows the norm</li> <li>School enrollment</li> <li>Accessibility for disabled students</li> <li>No bullying in schools</li> <li>No racism in schools</li> <li>...</li> </ul>	<ul style="list-style-type: none"> <li>The general average level of student achievement is decent or good</li> <li>The absolute level of student achievement raises in the country through the years</li> <li>the processes for the learning outcomes are proper and efficient</li> <li>The learning outcomes are produced economically sustainable way</li> <li>Decent student/teacher ratios</li> <li>No teacher's favoring in schools</li> <li>...</li> </ul>	Preferable indicators	
		Parity in binding indicators between... <ul style="list-style-type: none"> <li>rural/urban schools,</li> <li>private/community schools,</li> <li>geographical areas</li> <li>boys/girls</li> <li>language- or ethnic groups,</li> <li>different SES groups</li> <li>parents occupational-, educational- and economic background</li> <li>...</li> </ul>	Parity in preferable indicators between... <ul style="list-style-type: none"> <li>rural/urban schools,</li> <li>private/community schools,</li> <li>geographical areas</li> <li>boys/girls</li> <li>language- or ethnic groups,</li> <li>different SES groups</li> <li>parents occupational-, educational- and economic background</li> <li>processes used in education</li> <li>economical resources of the schools</li> <li>...</li> </ul>		
		Relative binding	Relative preferable		
		Relative indicators			

The *absolute binding indicators* are mainly legislative orders for education: those must be fulfilled or, if not, the educational provider, school, or individual acts criminally or otherwise irresponsible. Such banal indicators as availability of a school with a decent distance, equal possibility of girls and boys to entrance the school, and absence/presence in the school, may be typical absolute binding indicator. In the modern discourse, one may say that the idea of “Education for all” is real and active when absolute binding indicators are active. Also, such obvious matters as bullying or racist behavior fall in this category. Mostly, the absolute binding indicators are hard-fact indicators; sometimes they may be taken as naive or banal – but, if there are evidently some problems in these areas, we need to monitor those. The absolute binding indicators reflect the normative equality (“same for all”) and moral equity; they measure fairness in access to and to enjoy of the benefits from education as defined by Levin (2010). Many of the WB, OECD, and UNESCO indicators fall into this category.

The *relative binding indicators* point the relative differences in the binding indicators between specific groups such as geographical location of the school, school type (private/community), or schools management structures. There should not be any difference in students' and teachers' absence/presence between rural- and urban- or private- and community schools, or between the geographical areas within the country; children's possibilities in education should not be determined by an unlucky “accident” of being born in a certain city, in rural or urban area, or as boy or girl. If problems are found, such as inequality in reaching the education in different parts of the country, something should be amended in the system. Sometimes, the relative binding indicators may be more interesting than the absolute ones because they go strictly to the essence of proportional- and formal equality and moral equity; no difference should be found between the cases where there is something common between the cases and because everyone deserved the same dignity and the same respect. In the educational settings, the common element is the citizenship; all the children within one nation should to be offered the same possibilities to become what they can be.



The *absolute preferable indicators* are to do with the absolute facts of the learning outcomes and economic and efficient processes. These may be called also as quality indicators of the system related to the absolute matters. The learning outcomes and their increase in time seem to be the most used indicator of the processes in the general education (see PISA, TIMSS, PIRLS, and PIAAC results above); they can be classified as absolute preferable indicators. The economic aspect seems to be highlighted in PISA results (see OECD, 2013a). It seems that the quality of the *processes* is – for some reason – slightly on defensive. Also, decent student/teacher ratio can be taken as one of the absolute preferable indicators though there is no universal law of how many students there should or could be optimally for one teacher. The absolute preferable indicators reflect the Normative equality; “same for all”, however, not that strictly as the binding absolute indicators. These indicators could be classified as equity indicators rather than equality indicators – we need to offer equal opportunities for all children to reach the preset goals for education, nuanced in the local curriculum.

The *relative preferable indicators* relate to schools, teachers and students: the schools should be given equal opportunities to provide student high level teaching, the teachers should be provided with the same kind of competencies and working conditions, and the students should be given equal opportunities to reach the same goals regardless the home background. Here the students’ possibilities are highlighted over the schools’ and teachers’ possibilities. There should not be any difference between the learning outcomes between different geographical areas, boys and girls, or between the ethnic groups; children’s learning outcomes should not be determined by an “unlucky accident” of being born as a boy or girl, in a certain ethnic group, or in low- or high SES group. If inequalities are found – such as inequality in reaching the educational goals between the language- or ethnic groups – the government should react adequately to amend in the system.

A practical difference between the relative binding and relative preferable indicators is that, while the society can do something with the access to school and with the action within the school, the society cannot change easily the reality in the students’ families. It is known that the socioeconomic status (SES) of the pupil’ family seems to explain the achievement level quite well but it cannot be changed easily; the poor families cannot be made rich easily, and it may be unethical to try to make the rich families poor. We cannot make the highly educated parents less educated to reach the equality of SES. Hence, one cannot do much for the phenomenon. However, if the reason for poorer results of the students from the low-SES families is the low educational level or illiteracy of the parents (part of the SES), the government can do something to increase the adult literacy rate in the country; if the reason is the low economical input from the families for their children, the government could compensate this, for example, by offering free lunches, travelling, textbooks, and so on to reduce the gap between the groups.

As a conclusion, though the absolute- and binding indicators are important in assessing the state of art of the educational equity in the country, the relative- and preferable indicators may be more interesting because they go strictly to the essence of proportional- and formal equality and moral equity; no difference should be found between the cases where there is something common between the cases and because everyone deserved the same dignity and the same respect. In the educational settings, the common element is the citizenship; all the children within one nation should to be offered the same possibilities to be what they can be regardless their home background. In what follows, mainly the *relative preferable* indicators are in focus, that is, the parity indicators based on the learning outcomes.

### ***3.2 Some suggestions for parity indicators based on student assessment***

The general aim of the national level assessment of student achievement is to produce objective, accurate, and comparative information of the state of art of achievement of the students in the nation to scan the possible weaknesses in the educational system. The aimed learning outcomes are usually nuanced in the national curricula.

The national level student assessment is used as an operating tool to evaluate the current state of the educational system to produce these aimed results. When there is no examination system (like in Finland) or there is a

separate student assessment and evaluation on the side of the examinations (like in Nepal), the aim of the student assessment is primarily *not* to assess the students themselves. Instead, based on the primary observations from the students, the aim is to assess the efficiency and equality in the educational *system*. The sample-based assessment is a relatively cheap way to acquire this information. If problems are found – such as inequality to reach the educational goals in different parts of the country or between the language- or sex groups – something should be amended in the system. If the Ministry of Education and the other stakeholders, like teachers' unions, District Education Offices or teachers' trainers and curriculum developers, react adequately and efficiently to the assessment results, and the educational system is changed to reduce the inequality gaps between different groups, the sample-based student assessment has “barked its price”. Though the discussion here, apparently, seems to focus on national assessment it does not exclude the international comparisons; the point is that the national level student assessment is organized to get information of the educational settings in the society. Very local assessment settings, like continuous assessment in schools, are not in focus here.

The indicators from here on, are called parity indicators; they can be compared with Gender parity; one expects no differences between the groups. In Tables 2a–2d, more nuanced parity indicators or indicator areas are compiled. The indicators are divided into four sets: parity indicators related to the basic learning outcomes (Table 2a), parity indicators related to strict equality in learning outcomes (Table 2b), parity indicators related to other interesting factors that may have relevance in learning outcomes (Table 2c), and a selection of absolute parity indicators related to the learning outcomes (Table 2d). The suggestions are not exhaustive ones; more indicators can be found. Also, the names and abbreviations are arbitrary though justified. Except the last set, the indicators are practically *relative* ones; different groups are compared with each other. They are mainly preferable indicators; the learning outcomes are in focus. It may be possible, though out of the focus of this article, to create also indices based on the indicators.

#### ***Indicators related to Elementary basic results of learning outcomes***

The indicators related to the elementary basic results (see nuanced in Table 2a) can be divided into three. The indicators in the first set are based on the total national mean and the shape of the national distribution. These indicators are *Population Parity Indicator* (PPI, there should not be several student populations) and *International/National Parity Indicator* (INPI, the national results do not differ radically from the average international results). In language proficiency, where the criterion-based evaluation can be utilized (that is, standards such as the Common European Framework, CEFR – see also discussion of a parallel systemic in Mathematics in Metsämuuronen, 2018), the absolute ability level in the nation could be used also as an important indicator. It is not a parity indicator though.

The other set of indicators is related to content areas and item type wise results, such as *Content area Parity Indicator* (CAPI, no remarkable difference in achievement between the content areas, such as arithmetic and algebra in Mathematics), *Item type Parity Indicator* (ITPI, no remarkable difference in achievement between the subjective and objective type of items), and *Hierarchical level Parity Indicator* (HLPI, no remarkable difference in achievement between the different hierarchical cognitive levels of items, such as application type and recall type of items). The last of these (HLPI) can be questioned because the tendency in the real life is that the items reflecting higher skills (synthesizing and analyzing) are usually more demanding than the items for recalling the facts. Hence, we are expecting differences in achievement between the items reflecting the different hierarchical levels. However, there should not be any differences between sexes or school types in this regard.

The third type of indicator is *Continuity toward Parity Indicator* (CTPI, disparities are getting narrower during the years) which is, most of all, a longitudinal indicator and hence, maybe called an index.

#### ***Indicators related strictly to equality in learning outcomes***

The indicators related strictly to the equality in learning outcome results (see nuanced in Table 2b) can be divided into two. Obvious indicators related to the equality results are *Gender Parity Indicator* (GPI, no

differences between males and females), *Ethnicity Parity Indicator* (EPI, no differences between ethnic-, caste-, or religious groups), and *Home language Parity Indicator* (HLPI, no difference between the language groups).

The other set of indicators, related with geographical- and physical elements, include *District Parity Indicator* (DPI, no difference between districts or municipalities), *Region Parity Indicator* (RPI, no regional differences), *Geographical zone Parity Indicator* (GZPI, no differences between different living areas, such as in mountains, hills, and plains), *School type Parity Indicator* (STPI, no differences between community schools and private schools), *School location Parity Indicator* (SLPI, no difference between rural and urban schools), and *School language Parity Indicator* (SLPI, no differences in the results based on the administrative and instructional language of the schools).

### ***Some relative indicators related to other factors connected with learning outcomes***

The indicators related with the other interesting factors (see nuanced in Table 2c) may be arguable. The indicators introduced here are based on the national assessment results from Nepalese reality; the large-scaled national assessments of student achievements in 2012 (Metsämuuronen & Kafle, 2013) and 2013 (ERO, 2014; Metsämuuronen & Illic, 2018) are based on very convincing datasets of more than 100.000 students from grade 3, 5 and 8 to analyze which factors seem to explain the learning outcomes in Nepal. Based on the results, some possibly interesting factors are raised here as indicators of educational parity and imparity.

The indicators include such parity indicators or indicator areas as *Parents' education Parity Indicator* (PEPI, there should be no difference in student achievement between the educational groups of the parents), *Parents' occupation Parity Indicator* (POPI, no difference in student achievement between the occupational groups of the parents), *Home possessions- and -accessories Parity Indicator* (HPAPI, no difference in student achievement with different amount of home possessions and -accessories), *Socioeconomic status Parity Indicator* (SESPI, no difference in student achievement between the SES groups), *Age Parity Indicator* (API, Students are studying with their normal age group), *Help in studies Parity Indicator* (HSPI, no difference in student achievement between the different stakeholders giving help in studies), and *Homework given and checked Parity Indicator* (HGCPI, no difference in teachers' actions in giving and checking the homework).

**Table 2a.** Parity indicators for the basic student achievement results

Indicator area	Rationale	A possible result	Practical note on the rationale and interpretation based on Nepalese dataset
Population Parity Indicator (PPI)	Only one (normally distributed) population should be found.	"There are two distinctive student populations."	Several populations may indicate inequality in the student population. Two distinctive student populations (or widened normal distribution) may indicate differences between boys and girls, different geographical areas or school types.
International/ National Parity Indicator (INPI)	The national results do not differ radically from the average international results	"The average reading proficiency in grade 5 in English is much lower than the international average in PIRLS standard of grade 4."	The challenge in the international ranking lists and scores behind the ranks is that the international tests cannot measure the national goals. However, they tend to give indicative information of crude challenges in the national results. It is possible to combine the international item banks with the national tests by using linking items. Without linking the test scores with Item Response Theory (IRT) or Rasch modeling, the comparison does not make much sense.
Content area Parity Indicator (CPI)	No remarkable differences in achievement between the content areas should be found	"The learning outcomes in the content areas of Reading and Writing are 7–8 percent units lower than in Grammar and Vocabulary."	In Mathematics, for example, the learning outcomes may be lower in the content areas of Algebra and Numeracy in comparison with Arithmetic and Geometry. This may mean that the teaching in Algebra and Numeracy may be inefficient or that the curriculum favors Arithmetic and Geometry. This indicator may tell inequality between the geographical areas or school types.
Item type Parity Indicator (ITPI)	No remarkable difference in achievement between the subjective and objective type of items should be found	"In most cases, the students started to do the productive task, but the skills were not high enough for the highest marks."	Students tend to be better in multiple choice type of items (MCQ) than in the tasks requiring ability to produce something themselves. Hence, it is understandable that the proportion of correct answers within the open-ended items is somewhat lower than within the MCQ type of items. This indicator may tell strictly what kind of tasks the students cannot manage, which, secondary, may tell inequality between the geographical areas or school types.
Hierarchical level Parity Indicator (HLPI)	No remarkable difference in achievement between the different hierarchical levels of items should be found.	"8% of students were not able to solve any of the tasks requiring higher ability."	Students tend to be better in recall type of items than in the tasks requiring ability to solve problems requiring higher skills. Hence, it is understandable that the proportion of correct answers within complex items is somewhat lower than within the tasks requiring only recalling type of ability. When, in the test suitable for the age level, 10–18% of students are not able to solve any of the tasks requiring higher ability, there seems to be a problem in the educational system. Maybe, all students should be able to solve, at least, one task requiring ability to solve complex problems.
Continuity toward Parity Indicator (CTPI)	Disparities are getting narrower during the years	"Compared with the 1998 results, the students in urban schools score remarkably higher in 2012."	This indicator needs a historical and longitudinal dataset to be used. When the disparity in any parity indicator such as between boys and girls or between geographical areas gets smaller, it is a good sign from CTPI viewpoint. On the other hand, widened differences or no difference, when originally noticed, is a sign of a need for the system to react adequately.

Table 2b. Party indicators for the basic equality results

Indicator area	Rationale	A possible result	Practical note on the rationale and interpretation on the basis of Nepalese dataset
Gender Party Indicator (GPI)	No differences in achievement between males and females	"Though the differences between boys and girls in proficiency are statistically significant, in fact, they are very mild (0.6 percent units)."	In many societies, the education of girls has not been, historically, profitable for the family because the girls, in any case, turn to be part of the husband's family. The modern discourse, however, requires educational equity between boys and girls. Girls tend to be better in language while boys used to be better in mathematical studies. The latter is not necessarily reality in the developed countries any more.
Ethnicity Party Indicator (EPI)	No differences in achievement between ethnic-, caste-, or religious groups	"There are statistically significant though not remarkable differences (5.6 percent units at the highest) between the castes in English."	As an example of ethnic background, the caste is given here. Though the caste system is officially abandoned it still lives in the mind-sets of most people in Hindu religion settings. Historically, the Brahmins and Chhetris have been heavily involved in education, but Dalits, "untouchables", for example, have been practically outside of the educational system. Hence, the modern society has made lots of efforts to make the education possible and accessible for all children. Any differences in learning outcomes are a sign of inequality in the society.
Home language Party Indicator (HLPI)	No differences in achievement between the language groups	"The difference in English between the lowest and highest language groups is 53.0 percent points (out of 100%)."	The student achievement may depend on the language spoken in their homes i.e., the mother tongue of the students. The mother tongue reflects, in many cases, the ethnical background and hence any difference may be taken as a possible source for inequality in society. The ability of reading and writing in the major language may be one reason why these students seems to perform lower in the tests administered with the second language.
District Party Indicator (DPI)	No differences in achievement between the districts	"There are wide differences (51.0 percent points at the highest) between the districts"	Districts are geographical areas between the regional- and municipal level areas. In some cases, these could be called municipalities also. The point is that there should not be any differences between the educational opportunities between the larger or restricted entities in the country. District wise differences may indicate economical or educational differences (lack of high-SES families, for example).
Region Party Indicator (RPI)	No regional differences in achievement	"There is a wide difference between in the capital city and in the rest country (29 percent points)."	Regions are larger geographical areas in comparison with Districts (see DPI). In some cases, these could be called counties or provinces also. There should not be any differences between the educational opportunities between the larger entities in the country. Region wise differences may indicate economical or educational differences (lack of high-SES families, for example) in the different parts of the country.
Geographical zone Party Indicator (GZPI)	No differences in achievement between the geographical zones	"There are moderate differences (5.3 percent units at largest) between the student performances in three geographical zones when the Capital city is omitted"	The Mountain, Hill and Plain are examples of geographical features in certain countries. These kinds of geographical zones may include, from the population point of view, the mixed ethnicities, weather conditions, economic activities, more or less aggressive development, as well as differences in the density of the human capacity. The differences between the zones indicate unequal opportunities for the children.
School type Party Indicator (STPI)	No differences between community schools and private schools	"Students in the institutional schools outperform remarkably the students in the community schools. The difference is the widest in Reading (36 percent points)."	In some countries, like in Finland, the private schools are established because of some philosophical reason; the achievement results may be lower than in the standard community schools. In some other countries, attending a private school indicates higher socioeconomic status and, hence, the results may be higher than in the community schools. The reason is not necessarily the poorer teachers nor processes but the very intensive selection of the students. In any case, wide differences between the institutional and community schools indicate educational inequality in the society.
School location Party Indicator (SLPI)	No difference between rural and urban schools	"Students in the urban community schools gain 12 percent point more than the students in the rural areas."	In many countries the universities are located and the aggressive development happens in cities. This means that the country's intellectual and economical capacity may be condensed in cities. In developing countries, this may lead to uncontrolled urbanization if the families, on a large scale, send their children to big cities to study and move later themselves to seek a better life. Ultimately, if this trend continues in the future, it may lead to a wider inequality in society between rural and urban areas as well as between the geographical zones (see GZPI).
School language Party Indicator (SLPI)	No differences in the results based on the administrative and instructional language of the schools	"The students in the English media schools perform in English, naturally, much higher (36 percent points) than in non-English media schools."	In some countries, like in Finland and Canada, there are several official languages and the instruction language differs school wise. In that case, it is important to take care that the learning results do not differ in these schools. When testing a specific language, like English, in the English media schools, the better results are awaited. However, if there are differences in mathematics achievement, for example, it indicates educational inequality between the students.

**Table 2c.** Parity indicators based on interesting factors related to learning outcomes

Indicator area	Rationale	A possible result	Practical note on the rationale and interpretation based on Nepalese dataset
Parents' education Parity Indicator (PEPI)	No difference in student achievement between the educational groups of the parents	"36.4% of the students had an illiterate mother and 16.4% an illiterate father."	In Nepalese context, parents' educational level predicts well the children's future achievement level in English. Especially harmful for the achievement level seems to be the situation where the father or mother is or both are illiterate.
Parents' occupation Parity Indicator (POP)	No difference in student achievement between the occupational groups of the parents	"54.9% of the mothers and 35.3% of the fathers worked in agriculture or only home."	In Nepalese context, economic- and intellectual capacity at home seems to be connected strictly with the children's achievement levels. If the father or mother or both are coming from an agricultural or related occupation, the students' achievement in Mother language, Mathematics, and English is significantly and remarkably lower than with the other occupational groups.
Home possessions- and accessories Parity Indicator (HPAP)	No difference in student achievement with different amount of home possessions and accessories	"2.3% of the student did not have any of the 11 home possessions asked and 26.4% had none of the 3 accessories."	In Nepalese context, when children have very few home possessions (like a desk for homework or a dictionary) the achievement level is remarkably lower than the national average. When none or only one accessory (of a mobile phone, television, and computer) out of three is met, the results are lower than there are two or more met.
Socioeconomic status Parity Indicator (SESP)	No difference in student achievement between the SES groups	"13.9% of the students are at the lowest level of SES."	SES is usually a combined statistic of parents' education (see PEPI), occupation (see POP), and economical status (see HPAP) in the society. In Nepalese context, the difference in student achievement between the lowest and highest SES groups is remarkable (40 percent points). A special structural problem seems to be the mothers' high illiteracy rate and low educational level (see PEPI). If the problems in parents' low educational level would be solved, the results in these groups may raise remarkably.
Help in study Parity Indicator (HSP)	No difference in student achievement between the different stakeholders giving help in studies	"6.4% of the student get private tuition, 18.4% was helped by their teacher. The others were mainly helped by their sisters, brothers, or parents."	In the Nepalese dataset of 5 <sup>th</sup> grade English, there is about 8 percent point difference between those who don't get any kind of help and those who receive (private) tuition. The help given by the mother and by brother and sister seems to raise the achievement level more than help given by the father or teacher. The high percentage of private tuition indicates that there is very high competition in educational system. It is possible that the group with private tuition also spent more time on their homework, which may explain the higher score.
Homework given and checked Parity Indicator (HGCP)	No difference in teachers' actions in giving and checking the homework	"3.0% of the student expressed that they did not get homework or those were not checked."	In the Nepalese context, if the teacher gave homework and checked them systematically, the achievement level of the students was statistically significantly higher than without checking or issuing of homework. This may indicate lack of teacher training, or that these students are not very serious in going to school (and hence, they did not notice that teacher, in fact, always gave the homework).
Age Parity Indicator (API)	Students are studying with their normal age group	"25.4% of the students of 5 <sup>th</sup> grade fell aside 10–12 years."	In the Nepalese dataset of 5 <sup>th</sup> grade English, the highest performance is with those students studying with their normal age group, that is, at the age of 10 to 12 years. Otherwise the achievement decreases as the age increases. The too high or low age of the students may indicate repeating classes, insufficient special- and supporting education, and irregularities in entering the school.

**Table 2d.** Selected absolute parity indicators related to the learning outcomes

Indicator area	Rationale	A possible result	Practical note on the rationale and interpretation based on Nepalese dataset
Availability of textbook Parity Indicator (ATPI)	All students should have access to a proper textbook	"4.3% of the students lack the proper textbook in English."	In the Nepalese context, the achievement level of those students with no proper text book is significantly, though not necessarily remarkably, lower than those who have access to the textbook. Lack of textbook may indicate systematic problems in delivery systems in remote areas. In any case, lack of basic textbooks is an absolute indicator of inequity in the educational system.
Working after school Parity Indicator (WPI)	No paid work nor too much household chores should be required from children	"31.3% of the student worked in the paid capacity and 22.6% spent more than 2 hours in the household chores."	In Nepalese context, either working in a paid capacity or for four hours per day unpaid on household work outside school reduces statistically significantly the school achievement of the student. Though most of the low-grade students do not usually work many hours per day in the paid capacity – mostly one hour or less – their volume is too much. In any case, the child labor is prohibited by the law; something profound needs to be done to reduce the need for the school children to work in a paid capacity.
Student behavior Parity Indicator (SBPI)	No student should be bullied by the students nor teachers	"56.0% of the students have encountered bullying in school within the last month and 5.3% of students are experiencing a severe kind of bullying."	Bullying is one of the problems in the school that worsens the learning environment for the students. In Nepalese context, bullying seems to be quite common in schools though extreme cases of severe bullying are rare (5% of the students in Nepal). This negative phenomenon causes needless harm to young children and has to be rooted out from the schools.
Teacher behavior parity Indicator (TBPi)	No student should be favored or neglected in school by the teacher	"8.6% of the students feel that their teacher does not treat them fairly."	Many students may feel that their teachers are not treating the students fairly. If this appears to be reality, teachers need to change their behavior; all the students should be treated equally regardless their background, motivation, or achievement level. All the teachers, teachers' unions and teachers' trainers should be aware of this potential threat of equality.

### ***Some absolute indicators related variables related to learning outcomes***

Four absolute indicators which should show zero disparity are the *Availability of textbook Parity Indicator* (ATPI, all students should have access to a proper textbook), *Working after school Parity Indicator* (WPI, no paid work nor too much household chores should be required from children), *Student behavior Parity* (SBPI, no student should be bullied by students or teachers) indicating the frequency of bullying in school and *Teacher behavior Parity* (TBPI, no student should be favored or neglected in school) indicating the sense of unfairness of the teachers among the students.

In the Nepalese context, the achievement level of those students with no proper text book is significantly, though not necessarily remarkably, lower than those who have access to the textbook. Lack of textbook may indicate systematic problems in delivery systems in remote areas. In any case, lack of basic textbooks is an absolute indicator of inequity in the educational system. In Nepalese context, also, either working in a paid capacity or for four hours per day unpaid on household work outside school reduces statistically significantly the school achievement of the student. Though most of the low-grade students do not usually work many hours per day in the paid capacity – mostly one hour or less – their volume is too much. In any case, the child labor is prohibited by the law; something profound needs to be done to reduce the need for the school children to work in a paid capacity. Most probably the need for working in the paid capacity or need to participate more than 2 hours in the household chores is only one part of a complex knot of problems involved with the low SES affecting the low learning results. Bullying is one of the problems in the school that worsens the learning environment for the students. In Nepalese context, bullying seems to be quite common in schools (round 50% of the students expressed at least one kind of bullying out of five asked incidents) though extreme cases of severe bullying are rare (5–8% of the students expressed all 5 kinds of incidents). This negative phenomenon causes needless harm to young children and must be rooted out from the schools. Many students in Nepal feel that their teachers are not treating the students fairly. If this appears to be reality, teachers need to change their behavior; all the students should be treated equally regardless their background, motivation, or achievement level. All the teachers, teachers' unions and teachers' trainers should be aware of this potential threat of equality.

### **4. Concluding remarks**

The starting point of the treaty was the notion that all children have the moral right to get equal opportunities in education even though they are different and unique. The theoretical section discussed of different types of equality and equity and concluded that we do not seek numerical equality, but we seek proportional- and formal equality on the basis on moral equity. That is, in each age group all the children should be given the same opportunities to learn because learning is one of the basic rights for the children.

The empirical section of the article asked how to measure the possible equality and inequality in the educational system. A simple framework for categorizing the indicators suggested four categories based on two dimensions: binding/quality indicators and absolute/relative indicators. The binding indicators ask: “are we doing what we *should* do” and the preferable indicators ask: are we doing what we should do with *good results*, providing all citizens *equal possibilities with efficient processes with economically sustainable way*”. The absolute- and binding indicators are important in assessing the state of art of the educational *equity* in the country and the relative- and preferable indicators indicate the proportional- and formal *equality*; no difference should be found between the cases where there is something common between the cases and because everyone deserved the same dignity and the same respect.

The section proposing possible indicators for further use introduced 26 parity indicators based on learning outcomes and discusses their rationale and relevance based on National assessment of Student achievement in Nepal. These indicators were divided into four categories: parity indicators related the basic learning outcomes, parity indicators related with strict equality in learning outcomes, parity indicators related with interesting factors that may have relevance in learning outcomes, and a selection of absolute parity indicators related with



the learning outcomes. The suggestions are not exhaustive ones; more indicators can be created. However, these alone would enlarge our knowledge of the state of art of equity and equality in educational settings.

An obvious question is how the suggested equity indicators relate with the equity indicators published by the global players, WB, OECD, and UNESCO? The basic indicators of WB, OECD, and UNESCO seems to be absolute binding indicators based on numerical facts of the educational system while the 26 indicators here are relative preferable indicators based on learning outcomes. OECD has published also equity indicators based on learning outcomes; small portion of the 26 indicators suggested here tangent the ones used in OECD equity report (OECD, 2013a) but they enlarge notably the possibilities to monitor the equality in the educational system. The indicators suggested here tangent also quite well with Lewin's (2010, p. 8 ff.) identification of general groups of students for whom we are expecting to see the educational equity: gender, socioeconomic status, ethnicity, race, immigrants, language, region, and disability. This article provides us practical tools for assessing the possible inequalities and understanding *what* kind of equality we are seeking with the selected indicators.

One mild disclaimer of the international testing of student achievement comes, however, from the fact that the test items need to be constructed such way that they cannot perfectly match the national curriculum of any country. The international comparisons should be used very carefully as a basis of any changes in the education. Namely, the *national* student assessment is used as a tool to evaluate the state of the current national educational system to produce the aimed results nuanced in the national curricula. We need to keep in mind that the national curriculum is the political will of the country; those contents in the curriculum must be fulfilled – not any international testing setting's goals. If a country is famous of – for example – poetry and music and the country is willing to give that inheritance to the next generation, why should they change those specific topics to more mathematical and science-oriented subject just because those happened to interest other countries? On the other hand, the selection should be conscious: what consequences there may be not to concentrate on those subjects which internationally are lifted high. This mild disclaimer does not mean that the international testing settings, such as PISA, TIMSS, PIRLS, and PIAAC, are worthless – quite an opposite: they *are* important because they give comparable information of student achievement over the countries. Their possibilities in tackling the national questions are just limited and, hence, we need also national level assessment to ask whether the national goals are achieved.

International comparisons of equity and equality are challenged by Levin's (2010) note that there is no natural state of educational equity, but one defined by each society based on its values and the imperative that it sets for the issue as a moral commitment. If really each society may define educational equity in different ways and use different criteria for assessing equity in the overall population and among different groups, the question is: how, and on what basis, we would create comparable measurement instruments for all societies for educational equity? Maybe the generic equality indicators based on national learning outcomes would be one part of this kind of toolbox.

## References

- Aristotle, *Nicomachean Ethics*, in *The complete works of Aristotle*, ed. Jonathan Barnes, Princeton: Princeton University Press.
- Aristotle, *Politics*, in *The complete works of Aristotle*, ed. Jonathan Barnes, Princeton: Princeton University Press.
- Arneson, R. J. (2014). Basic Equality: Neither Rejectable nor Acceptable. In Uwe Steinhoff (ed.), *Do All Persons Have Equal Moral Worth? On 'Basic Equality' and Equal Respect and Concern*, Oxford: Oxford University Press.
- Arneson, R. J. (2001). Luck and Equality. *Proceedings of the Aristotelian Society*, supp. vol. pp. 73–90.
- Collins Cobuild (2014). Collins COBUILD Advanced Learner's Dictionary of British English. Retrieved from <http://www.collinsdictionary.com/dictionary/english-cobuild-learners/indicator?showCookiePolicy=true> (accessed February 26, 2019).
- Dann, O. (1975). Gleichheit. In V. O. Brunner, W. Conze, & R. Koselleck (Eds), *Geschichtliche Grundbegriffe*. Stuttgart: Klett-Cotta. pp. 995–1046.

- Delorme, P. & Chatelain, O. (2011). Policy Steering. The Role and Use of Performance Measurement Indicators. Aid Delivery Method Program. EuropeAid. Retrieved from <https://europa.eu/capacity4dev/dear-programme/document/adm-policy-steering-role-and-use-performance> (accessed February 26, 2019).
- ERO (2014). *Student achievement in Mathematics, Nepali and English at grade 3 and 5 in 2011*. Educational Review Office, Ministry of Education, Kathmandu, Nepal.
- Gosepath, S. (2007). Equality. *Stanford Encyclopedia of Philosophy*. First published Tue Mar 27, 2001; substantive revision Wed Jun 27, 2007. Retrieved from <http://plato.stanford.edu/entries/equality/> (accessed February 26, 2019).
- Guardian*, 6<sup>th</sup> May 2014, an open letter to Andreas Schleicher. Retrieved from <http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics> (accessed February 26, 2019).
- Guiton, G. W. (1992). *Developing indicators of educational equality: conceptual and methodological dilemmas*. Los Angeles: University of California, Los Angeles, Center for Research on Educational Standards and Student Testing.
- Guiton, G. & Oakes, J. (1995). Opportunity to Learn and Conceptions of Educational Equality. *Educational Evaluation and Policy Analysis*, 17(3), 323–336. <http://dx.doi.org/10.3102/01623737017003323>.
- Hurley, S., (2001). Egalitarianism and Luck. *Proceedings of the Aristotelian Society*, supp. vol. pp. 51–72.
- Kant, I. (1785), *Grundlegung zur Metaphysik der Sitten*, in: *Kants Gesammelte Schriften*, ed. by Preußischen Akademie der Wissenschaften, Berlin 1902 ff., vol. IV.
- Kant, I. (1797), *Metaphysik der Sitten*, In: *Kants Gesammelte Schriften*, ed. by Preußischen Akademie der Wissenschaften, Berlin 1902 ff, vol. VI.
- Levin, H. (2010). *A Guiding Framework for Measuring Educational Equity*. INES Network for the Collection and the Adjudication of System-Level Descriptive Information on Educational Structures, Policies and Practices, EDU/EDPC/INES/NESLI(2010)6, March 2010.
- Locke, J. (1690). *The Second Treatise of Government*, ed. C.B. MacPerson, Indianapolis: Hackett 1980.
- Mannis, A. (2014). Indicators of Sustainable Development. In Gaia: Environmental Information System. Retrieved from <http://www.ess.co.at/GAIA/Reports/indics.html> (accessed February 26, 2019).
- Martin, M. O., Mullis, I.V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 International Results in Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Menne, A. (1962). Identität, Gleichheit, Ähnlichkeit: Eine logisch-philosophische Untersuchung. *Ratio*, 4, 44–53.
- Metsämuuronen, J. (2018). Common Framework for Mathematics – Discussions of Possibilities to Develop a Set of General Standards for Assessing Proficiency in Mathematics. *International Electronic Journal of Mathematics Education*, 13(2), 13–39. DOI: <https://doi.org/10.12973/iejme/2693>.
- Metsämuuronen, J. & Ilic, J. (2018). Why some low resourced community schools in Nepal perform well in English. In D. Hayes, (Ed.) *English Language Teaching in Nepal: Research, Reflection and Practice*. Nepal: British Council. (pp.107–132). ISBN 978-0-86355-888-7 Retrieved from [https://www.researchgate.net/publication/323759197\\_Why\\_Some\\_Low\\_Resourced\\_Community\\_Schools\\_in\\_Nepal\\_perform\\_Well\\_in\\_English](https://www.researchgate.net/publication/323759197_Why_Some_Low_Resourced_Community_Schools_in_Nepal_perform_Well_in_English) (accessed February 26, 2019).
- Metsämuuronen, J. & Kafle, B. R. (Eds) (2013). *Where Are We Now? Student achievement in Mathematics, Nepali and Social Studies in 2011*. Ministry of Education, Kathmandu, Nepal.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/isc/publications.html>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/isc/publications.html>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/isc/publications.html>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/isc/publications.html>.
- Nathan, C. M. J. (2014). What is Basic Equality? In Uwe Steinhoff (ed.), *Do All Persons Have Equal Moral Worth? On 'Basic Equality' and Equal Respect and Concern*, Oxford: Oxford University Press. Retrieved from <http://ukcatalogue.oup.com/product/9780198719502.do>. (accessed February 26, 2019).
- OECD (2013a), *PISA 2012 Results: Excellence Through Equity: Giving Every Student the Chance to Succeed (Volume II)*, PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264201132-en>.
- OECD (2013b). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD Publishing. <http://dx.doi.org/10.1787/9789264202456-en>.

- OECD (2013c). *PISA 2012 Results: Ready to Learn: Students' Engagement, Drive and Self-Beliefs (Volume III)*. OECD Publishing. <http://dx.doi.org/10.1787/9789264201170-en>.
- OECD (2013d). *PISA 2012 Results: What Makes Schools Successful: Resources, Policies and Practices (Volume IV)*. OECD Publishing. <http://dx.doi.org/10.1787/9789264201156-en>.
- OECD (2014a). *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014)*. Student Performance in Mathematics, Reading and Science. OECD Publishing. <http://dx.doi.org/10.1787/9789264208780-en>.
- OECD (2014b). *Education at a Glance 2014: OECD Indicators*. OECD Publishing. <http://dx.doi.org/10.1787/eag-2014-en>.
- OECD (2016a). *PISA 2015 Results (Volume I). Excellence and Equity in Education*. OECD Publishing. DOI: <http://dx.doi.org/10.1787/9789264266490-en>.
- OECD (2016b). *Skills Matter: Further Results from the Survey of Adult Skills*. OECD Skills Studies. OECD Publishing. DOI: <http://dx.doi.org/10.1787/9789264258051-en>.
- Oppenheim, F. (1970). Egalitarianism as a Descriptive Concept. *American Philosophical Quarterly*, 7, 143–152, reprinted in L. Pojman & R. Westmoreland (eds.), *Equality. Selected Readings*, Oxford: Oxford University Press 1997, pp. 55–65.
- Plato, *Laws*, in: Plato, *Complete Works*, ed. John M. Cooper, D.S. Hutchinson, Indianapolis: Hackett 1997.
- Rae, D. & Yates, D. (& Hochschild, J., Morone, J., and Fessler, C.) (1981) *Equalities*. Cambridge, Mass.: Harvard University Press.
- Rawls, J. (1971). *Theory of Justice*. Oxford, Clarendon Press; and Cambridge, Massachusetts, Harvard University Press.
- Rousseau, J.-J. (1755), *A Discourse on Inequality*, London: Penguin 1984, partly reprinted in L. Pojman & R. Westmoreland (eds.), *Equality. Selected Readings*, Oxford: Oxford University Press 1997, pp. 36–45.
- Sen, A. (1980). Equality of what? In S. M. McMurrin, (ed.), *Tanner Lectures on Human Values*. Vol. I. Cambridge, Cambridge University Press. pp. 197–220
- Sen, A. (2005). Development as Capability Expansion. In S. Fukuda-Parr & A. K. Shiva Kumar, *Readings in Human Development: Concepts, Measures and Policies for a Development Paradigm*. 2<sup>nd</sup> Edition. New Delhi and New York: Oxford University Press.
- Sherman J. D. & Poirier J. M. (2007). *Educational Equity and Public Policy: Comparing Results from 16 Countries*. UIS Working Paper No. 6. UNESCO Institute for Statistics, Montreal. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000149523> (accessed February 26, 2019).
- Temkin, L. (1993). *Inequality*. Oxford: Oxford University Press.
- Tugendhat, E. & Wolf, U. (1983), *Logisch-Semantische Propädeutik*, Stuttgart: Reclam.
- UN (1948). Universal Declaration of Human Rights. Retrieved from <http://www.un.org/en/universal-declaration-human-rights/index.html> (accessed February 26, 2019).
- UNESCO (2011). *Systematic Monitoring of Education for All. Training Modules for Asia-Pacific*. UNESCO Bangkok. Asia and Pacific Regional Bureau for Education. Thailand. Retrieved from <https://uis.unesco.org/sites/default/files/documents/systematic-monitoring-of-education-for-all-training-modules-for-asia-pacific-en.pdf> (accessed February 26, 2019).
- UNESCO (1994). *World Declaration on Education for All and Framework for Action to Meet Basic Learning Needs. Guidelines for implementing the World Declaration on Education for All*. Third Printing. Adopted by the World Conference on Education for All Meeting Basic Learning Needs Jomtien, Thailand, 5-9 March 1990. UNESCO for the Secretariat of the International Consultative Forum on Education for All. Paris. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000127583> (accessed February 26, 2019).
- Westen, P. (2014). *Speaking of Equality: An Analysis of the Rhetorical Force of 'Equality' in Moral and Legal Discourse*. Original 1991. Princeton: Princeton University Press.