Building data literacy with the teaching profession at global scale

Jón Torfi Jónasson, professor emeritus at the University of Iceland

[jtj@hi.is](mailto:jtj@hi.is) <https://uni.hi.is/jtj/en/>

My interest in this topic has three strands. The second is perhaps most directly related to the current project (BDL) even if I see all three as relevant, as they are closely related.

## A critical probe: The directional value of data

It is important to explore, from a scientific perspective, the directional value of data within education. The primary question is this: given that one has solid, valid data, how much direction for action can be implicit therein, within either policy, or practice. I have argued (in Icelandic) that despite the crucial and multifaceted value of data and evidence (assuming these to be of the highest quality) the directional value, both for policy and practice is generally vastly overestimated, by teachers, policy makers and academics. An issue that is largely neglected by all concerned.

Reference to my Icelandic paper on this issue (with an English abstract) and to a number of my recent English presentations on a similar topic are listed in Appendix 1.

## A research project: The views of the educational professional –– the topic for March 18th

Given the emphasis on data by policy-makers and administrators it is important to explore what value teachers attach to data. This may be informal data collected during the daily teaching routine, more formal tests in the class, formal (often standardised – even national) tests or examinations, diagnostic data, or perhaps international large scale tests. There is a need to gauge how they want to harness the available data, in particular, what directional value they attach to it and on what basis they would act.

With some colleagues, I lead in 2019, an application to research this issue. This was not successful. A part of the background framework is presented in Appendix II. The most important section, Research issues (questions), is on the five different foci of the proposed research, presented in Appendix II. [The first focus was on clarifying what the participants saw as the principal aims of the education for which they were responsible. The second, to explore how they saw available data as enabling them to achieve these aims. The third about what data they saw as most useful and who should be responsible for collecting it. The fourth on examples from their professional practice about the use of data and the fifth on the caveats or challenges of data use related to their practice. All the discussion would be supported by a selection of relevant examples.]

## A different voice? Ingredients of professional development.

Given the value attached to data by the educational world, it is important to deliberate what information and skills would be sensible to impart to the professional teaching force in this area. Having been responsible for statistics and quantitative methods courses in the social sciences and then also qualitative methods – also for teachers, on and off, for three decades I have developed strong views on this issue, but they have certainly changed quite notably during this period. I argue that the focus should be on the construct and consequential validity of the available data (rather than on collecting and analysing) and that the focus should be on in-service professional learning rather than on pre-service input.

I have not formally developed this point, except in notes.

Uncertainty about the directional value of data:

Data and its formative value for educational aims, policy and practice – for systems or students

The point of departure is that data can be immensely valuable for education at all levels, ranging from testing, assessing or diagnosing students, to doing the same for schools or national education systems. The data is supposed to be good, solid, and transparent, which invites objectivity, equity and fairness. Without underpinning this crucial point further, I only stress the immense importance of data for describing any situation, showing strengths or weaknesses or disruptions at different levels of education, both overall and down to the tiny details.

My main argument is that, despite its strength, data does by no means solve all our problems. The main reason is the lack of directional value inherent in the data: an assumed value that is vastly overestimated. I may present my argument from two sides.

This involves going from data to action, using the largely uncontested notion of formative assessment (of students, schools or systems) as an example. Here the implicit metaphor seems to be the cybernetic feedback loop. When one knows the situation, it is assumed that it is essentially built into the system (represented by a competent teacher, a school principal or a government official) what the reaction should be. The problem is, that there are, potentially, a vast array of possible sensible actions a competent creative professional might choose from – which is not typical, of a normal cybernetic loop. I claim that the guidance provided by data is very limited, even in the well-argued position taken by Lorrie Shepherd and her colleagues (Shepard, 2018; Shepard et al., 2017) where they argue that the data has a guiding value only when adopting an elaborate learning model. This needs, of course, to be clarified, by a number of examples.

Coming to the second approach one may ask if nothing sensible can be done within the educational arena without data. Of course, and most of education in the hands of a professional can very fruitfully be undertaken without much data. Having a discussion with students or policy makers or reading an essay or other assignments calls for constant reaction, guiding and suggestions. A true educational activity is relentlessly being conducted, at all levels of education, without recourse to data in the sense the term is used in the modern data discourse?

* Presentations – PowerPoint slides are accessible

**2019.4 April.** The uncertain directional value of data in education. Or – what is the value of data for everyday education, its policy and practice? Or – for what, exactly, is data relevant? Presented the University of Aarhus, April 3rd 2019. [Aarhus presentation on data](https://uni.hi.is/jtj/files/2019/06/Aarhus-presentation-on-data.pptx)

**2019.3 March.** The uncertain directional value of data in education. Presented at the NERA 2019 in Uppsala on March 6th 2019. [NERA JTJ paper-v1](https://uni.hi.is/jtj/files/2019/03/NERA-JTJ-paper-v1.pptx)

**2018.12 September.** In the direction of good education, teachers navigate the seas of data. Presented at the [Northern Lights conference](https://karvi.fi/en/event/northern-lights-conference/) Espoo, Finland, 27.-28. September 2018. [JTJ Northern lights Espoo Sept 2018](https://uni.hi.is/jtj/files/2018/10/JTJ-Northern-lights-Espoo-Sept-2018.pptx)  [See Youtube link 1:50:00-2:20:00](https://www.youtube.com/watch?v=wRHU-aEA3ww)

# Appendix I Substantive background

* Paper – see English abstract – see below.

Jón Torfi Jónasson. (2019). [Óvissa um leiðsagnargildi gagna? Notkun gagna við mótun menntastefnu og skólastarfs](https://ojs.hi.is/tuuom/article/view/3045). [Tímarit um uppeldi og menntun,](https://ojs.hi.is/tuuom/issue/view/324) 28(2), bls. 161-180. [Uncertainty about the directional value of data. Data and their formative value for setting educational aims, policy and practice.]

The paper abstract in English:

Uncertainty about the directional value of data:

Data and their formative value for setting educational aims, policy and practice

In recent years there seems to be an increasing discussion of, and emphasis on, the use of data in every sphere of life, and certainly in education. The growth and improved availability of data is obvious and questions are raised about its role and impact on life in general, as demonstrated by the speculations presented by Pentland (2019) and Kallinikos (2018) on how data are changing our lives. The paper demonstrates that several different worlds of data are developing in addition to our more traditional worlds of test scores and student social background. We have the big data and AI world (Jóhann Ari Lárusson og White, 2014; Luckin et al., 2016; Williamson, 2017) and the emerging bio-social data world, including brain based education (Youdell og Lindley, 2019). These focus more on the individual student or on accumulated data relating to large groups which, however, allow individual analysis and response. There are also the large-scale studies run by the OECD and IEA, partly combined in the WIDE data base (Cooper et al., 2009; UNESCO, n.d.). Describing these advances, it is emphasized that the discussion of data has been with us for a long time but seems to be attracting increased attention, apparently for very good reasons. We address the key question of how useful data really are and note that several authors and institutions do not think this is an issue, because of their obvious importance and power. We cite several different sources to underpin this view (Cooper et al., 2009; Custer et al., 2018; World Bank, 2018). We then develop the question by asking what exactly are data good for? Can we do nothing without them? Moreover, can we do practically everything with them? In particular, in what sense can data guide our policies and actions? We then defend the view that educational aims are based on personal and social values; data may be useful for clarifying their importance, but not for pinpointing them. Next we go on to discuss the notion of formative evaluation and again claim that the guidance provided by data is limited, even in the well-argued position taken by Lorrie Shepherd and her colleagues (Shepard, 2018; Shepard et al., 2017). Similarly, the cybernetic model of feedback does not apply well to education; even though it is clear that guidance is valuable, the data do not tell which way to go. Thus, the notion of assessment for learning, formative assessment, may be misleading. This is because even if the intention is that the assessment should serve learning it does so at best indirectly and is thus far less useful than the phrase seems to indicate. Throughout the paper it is emphasised that we do not question the usefulness of data due to their transparency, e.g., as exemplified by their relationship to operational definitions, objectivity and how they invite equity and fairness. Data certainly have their very important place. In the last part of the paper it is suggested that a serious problem with the emphasis on data is that they potentially crowd out important discussion about the aims of education. Several references to important classical and Icelandic texts are noted which should have a dominant place in the professional discourse on education. As an example, two Icelandic authors are discussed for their relevance in this context. Taking this perspective, it is interesting that even when discussing the complicated present and future our young people face, it may well be that classical old ideas may serve us well in deciding where education should take us. In educating our professionals, we must ensure that the balance is not tilted away from the aims and conduct of education in favour of other somewhat less important technical matters.

# Appendix II The research strand – a part of a research application

# Specific aims of the project, research questions/hypotheses, feasibility, originality and impact

Given the general and formal emphasis on data at all levels of the educational discourse and practice, it is imperative to understand how it interacts with the daily practice of education. The aim of the study is to explore how education professionals understand the usefulness of data in the policymaking and practice of education at various levels. We will define data in the background section, but in general terms we are talking about student data, school data and system data. These data may be either qualitative or quantitative, or of some hybrid type. Thus we are referring to data on individual students which may be summative or formative assessment, results from screening, diagnostic tests or portfolios. Furthermore, there are data at the school level, such as internal or external evaluations or at the system level such as PISA scores or dropout measures. As the focus of this research is to describe and understand how various professional groups within education view data, we ask: What do they consider as data? How are data used – actually and ideally – and how important do they think data are for education in their experience? Why? What do they consider the important challenges to address posed by data that they have, if any? This has scientific, pragmatic and practical implications as discussed in the value section.

The background to the study is the ongoing – and somewhat unquestioned – emphasis on collecting and using data for decisions within all levels of education. The impetus for this emphasis is diverse. Ranging from the strive to better understand the nature of education and its practices and how these unfold, to better informed policymaking, systematic and transparent program evaluation and, not least, providing fair, rich and transparent information for the teacher. The conviction, that data are an inevitable and crucial part of goal setting and policy-making both at macro (system) and micro (school and student) levels of education – prompts the question, which is the topic of this study: How do education professionals see and experience data’s role in educational practice and policy? Moreover, to be sure, the indisputable usefulness of data for scientific, descriptive, analytic and communicative purposes is not at issue here.

In education, a major task is to set goals: for the system, for each school, a class and for individual students. Once the goals have been set, the appropriate action must be initiated, at the respective levels, to ensure inclusion, to reduce dropout, to enhance reading skills or to fulfil other worthy goals. Every day each teacher steers her students towards certain goals that have been set. Thus, acknowledging or setting goals and then initiating reasoned action are substantive everyday tasks of education. Thus, we return to the question: what is the role of data in making the necessary decisions. It seems a simple question, yet it is a complex one.

Here the apparently ubiquitous (and perhaps increasing) function of data meets a number of challenges, two of which we focus on in this study: What is – broadly speaking – the value of data for setting educational aims in the eyes of the professionals concerned? And, what is, to them, the value of data to plan direct action at various levels of education? In order to understand these challenges, we turn to the data users themselves, the teachers, school leaders, counsellors and local and central administrators – and students – in order to find out what they think about data in their life and work. The main focus is on educators at the compulsory and upper-secondary school levels but pre-primary and tertiary levels are also addressed. The limitation is solely for the sake of manageability while still respecting the breadth of the field.

This study aims to adopt a wide perspective by involving teachers working in the classroom, but also school leaders, counsellors and the administrators of a system of education, and indeed the students. We will characterize the views towards data at each level (more solid description for compulsory and upper-secondary). Thereby we will be able to point to important differences taking also into account the level of the data (student data, school data, and system data). We also determine to which extent the concerns and ideas about the relevance and use of data (sometimes the same data, such as PISA scores) are the same at different working levels. For this reason, we do not confine our study to either the micro or to the macro level. We emphasize that although the study is structured by grouping the participants, individual interviews will allow us to gauge the extent to which the discourse is similar in the details among the different individuals and groups but also in what way it is different within different spheres of education and for different levels or types of data.

Research issues (questions)

The study explores several related aspects of the understanding of the use of data, among professionals within the field of education. Five research issues or foci of different levels of generality will guide the discussion, both in the focus groups and interviews. The most general focus is presented first and then the progressively narrower foci, with headings that give an indication of the content. See also the method section on focus groups and interviews. More elaboration is presented in the theoretical background text.

*Education meets the world of data.* The first focus is on the participants’ general views on the main aims of education at the level at which they mainly work (e.g. as related to rationale presented in the law or national curriculum). This is a first step towards understanding how the participants might see data of any kind as relevant, either to define or to clarify these educational aims. Here we will refer to the national curriculum, and recent discussion related to it. As concrete examples, we will bring up the 21st century skill discourse, the PISA global competencies, the UNICEF right of the child discourse, and the well-being discourses and how data may be relevant in defining or prioritizing aims. We will attempt to find out participants’ views on the way education shall be conducted and how these views relate to data by presenting examples for discussion. Do they favour competition, selection, streaming, and standardization all of which require data? How do they, in this context, see the balance or relationship between their role as ensuring standards and developing the individual?

*General views of the use of data – addressed through examples.* The second focus is on how professionals within the arena of education, understand the importance of data, in particular its guiding or directional value. How and to what extent will data tell them where to go next (both in which direction and which steps to take). We will explore this by using a variety of examples that we construct from the known world of the professionals in question. A policy maker has data on the outcome of PISA and there is a consensus that something must be done because the outcome is worse than desired. How does she regard the data and what steps does she feel are, at least implicitly, does the data advise? A teacher has data (e.g., from an interim data gathering) on how her pupils understand a certain topic that she has been teaching and feels that they should get a better grip. How does she think and talk about the data and what implicit (or explicit) advice does she feel is present in the data she has?

*The necessary data – the interesting or welcome data.* The third focus is on what decisions the professionals think should be based on some formal data use and collection and who might or should be responsible for the collection and use of data when they seem to be desirable. The decisions we have in mind, are for example related to daily teaching, selection or streaming, or distribution of resources. What data do the participants think they absolutely need; what data do they feel the system suffers from, by not having? The other side will also be addressed: what data that are already available do they consider beneficial to have? In the research interviews, a list will be presented to each participant after a brief open discussion at the beginning of the interview.

*Into the practice itself.* The fourth focus is on examples of how the professionals actually use data and look at examples of where they get the data from and how they are used. How does a teacher use outcome of screening or examinations, either informal tests or interim examinations? Or, how do they actually use any other kind of information or data. How do administrators advise the use of international or national tests or evaluations?

*Some formal challenges presented by data use.* The fifth and final focus is on various challenges or concerns that are related to the use of data. Good data are difficult to collect and inferences are fraught with problems. This is a particularly complex field to explore and we will approach it with carefully chosen examples for discussion, and not enter an abstract discussion. The most important concerns here, as discussed in the background section, are: the questions of causal attribution, how the data defines the significant topics to address or probe and questions of validity. We will carefully monitor that we do not burden the participants with too much material and it is even possible that these latter issues must be explored in a separate adjunct study, and thus only marginally be included here.

Educational arenas and foci of inquiry. The professional groups we will address operate within two arenas or levels within the education system. The first refers to teachers, counsellors and school principals (with focus on students and teaching), where we explore the ways they use and view data – be they summative, formative, screening or diagnostic, qualitative or quantitative, or are of some hybrid type. The second arena, refers to the administrative professionals, including both the top level in the municipal and the ministerial apparatus, also looking at different categories of data. For the school arena (with focus on teaching), we plan to have five groups, i.e., from the pre-primary, compulsory (primary and lower-secondary), the upper secondary levels and tertiary level (where we concentrate on those attached to teacher education). For the school levels that are in focus (compulsory and upper-secondary) we intend to talk to school heads, students and teachers, but separately.

Value and impact

Within education, every professional has to deal with a complex interaction of values (related to the aims of education), good teaching practice and a wealth of information for instructional purposes. It is important from the scientific perspective to grasp how these interconnected worlds relate in the minds and work of practicing professionals and thus be able to describe in what way data influences how professionals understand and conduct education.

Financial and professional resources are committed at all levels of the educational system in an effort to have a positive impact. Thus, a light cast on the question of distribution of resources and of the necessary expertise is very important from a pragmatic perspective.

The better we understand the challenges faced by the professionals involved the better we might prepare and guide them in their complex tasks. This may influence their basic education, their professional development and the support systems we develop for their aid.

# Present state of knowledge in the field

A world of data and the data imperative

….

Data in education and the professional community

…

Documentary analysis

….

Summary

Drawing on a wide spectrum of research and discussion literature, we have shown that data are playing an important role in education and this is likely to both to increase and change. We have emphasised the apparent change in the data landscape and pressures to use more data and also emphasising previous research in the area we address. Jónasson (2016) has argued that if education changes slowly we can partly explain this if we understand the attitudes of those within the system. In the present project we gauge how professionals view and harness the wealth of data that are available to them, and can potentially anticipate how they deal with new tools and data landscape. When we discuss the value of data in the context of this research, we always assume that we are talking about data of high quality (i.e. we make the assumption that the data in question are good). We have shown that various authors want to use data for different purposes and there is no question in our minds that understanding, analysing and talking about much of education and certainly motivating action is difficult without data. However, what we think is not at issue here. We have opened up questions about how data are relevant to professionals working within the field of education in order to allow them do all the important things they are expected to do and want to do. Moreover, we want to understand how they see the role of data in setting goals and determining action, and to get some idea of how they go about this. We want to understand their perspective independent of academic debates about these issues.

Time plan

....

Method

The pilot study. We intend to ask six experienced teachers and another of six headmasters who are thus well versed in the discourses about data to participate twice in a focus group. They will come from different levels in the system. We will also interview them individually. By doing this we will be able to explore the framework we develop, obtain a concrete idea about how to progress from one focus group to the next and an interview, and how this should be organized in detail and probe what unexpected problems we may have to deal with.

The documentary study. We will analyse legal and related formal texts looking for terms that in some way indicate the use of data or evidence (see the background section). Connected to the different levels we will analyse the terminology used in these policy documents. We will determine to what extent data defines educational aims or the steps to be taken. This analysis will serve as background to our focus group discussions.

The different groups of participants. See Table 2. Our main foci are the compulsory and upper-secondary levels and thus students, teachers, school leaders, counsellors and administrators who deal with those levels. We will augment this with inclusion of both participants from pre-primary and tertiary levels. As far as the professional groups are concerned (except the administrators), we will include both those that are reported to be enthusiastic about the use of data and those who may have reservations. We will focus on professionals who have between five and fifteen years of experience; they have developed their ideas but still represent “recent” views. This will not apply to the students nor necessarily to those at the administrative level.

Collection of data. The research paradigm is based on focus group interviews and a follow up with individual interviews, with approximately six people per group and individual interviews, partly scaffolded with a structured interviewing list. We will explain to the participants with examples what we are talking about when we talk about data (“gögn” in Icelandic) and that our aim is to understand how professionals within the field of education understand the role of data. We will emphasise that there is no right answer but given the emphasis on data in educational discourses, we think it is very important to understand how professionals think about data. We envisage a two-stage process. There will be two meetings of each focus group and then a meeting with each individual participant. At both meetings we will allow open discussion, but primarily focus on concrete examples. At the individual interview, we will ask the participant to respond to a brief questionnaire in order to obtain the relevant background information and then engage in a short discussion on the nature of education, the use of data and further exploration of the views emerging from the focus group.

Table 2. Overview of the groups for which data will be collected and an estimate of the time needed for the data collection, the raw collection of the data itself and its primary analysis. In the hands of RA, PhD and MAs.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Data collection and analysis | Focus group  Once | Focus group  Twice | Individual interviews |  | Approximate total time for focus groups and interviews (preparation, data collection and analysis).  40 hours per focus group meting  15 hours per interview |
| Pilot |  | 2 (6) | 12 |  |  |
| Teachers pre-primary |  | 1 (6) | 6 |  |  |
| Students pre-primary |  |  |  |  |  |
| School leadership pre-primary |  | 1 (6) | 6 |  |  |
| Teachers primary |  | 1 (6) | 6 |  |  |
| Students primary | 1 (6) |  |  |  |  |
| School leadership primary |  | 1 (6) | 6 |  |  |
| Teachers lower secondary |  | 1 (6) | 6 |  |  |
| Students lower secondary | 1 (6) |  |  |  |  |
| School leadership lower secondary |  | 1 (6) | 6 |  |  |
| Teachers upper secondary |  | 1 (6) | 6 |  |  |
| Students upper secondary | 1 (6) |  |  |  |  |
| Teachers tertiary-secondary |  | 1 (6) | 6 |  |  |
| School leadership upper secondary |  | 1 (6) | 6 |  |  |
| Students tertiary | 1 (6) |  |  |  |  |
| Administration of education |  | 2 (6) | 12 |  |  |
| Total number of focus group meetings | 4 | 26 |  | 30 | 30 weeks |
| Total number of interviews |  |  | 78 | 78 | 20 weeks |
| Total time used collecting and primary analysis of the data |  |  |  |  | 50 weeks (RA, PhD, MA) |

The selection of the participants is not straightforward. First, the question is what type of randomness should be chosen? We are not attempting to construct a descriptive normative picture of what Icelandic professionals think, but rather to find out, by the way of examples and by going to considerable depth, what a selected group of professionals think about fundamental issues, which concern the use of various types of data in education. From this perspective, we seek a divergence of views at all levels (which may, not exist however). Thus, we intend to solicit views from a knowledgeable group, which we will define with reference to experience, i.e. number of years in a particular context, what emphasis they place on data and on what grounds. Thus, we plan on a version of purposeful sampling taking quotas from the levels we define. We acknowledge that by focussing on different groups we are anticipating differences that we will discuss if they emerge. Second, some people may be reluctant to participate as they may feel that they are being “tested” in some way, even though this is of course not the intention and care will be taken not to convey such impression. We will of course present this for what it is: a scientific exploration of how people think about data and their use, irrespective of what enthusiasts or critics may think.

We will ensure the anonymity of the participants. Thus, we will only specify in our description of those that participate, at which level of the system they are working, and what their role is (sometimes in general terms). This applies in particular to those who are not teachers as they may hold posts that are easier to identify. We will not list the names of the participants in our reporting and we will present them all in the same gender. We will erase the actual recordings of the focus groups or the interviews as some of the participants may feel, by hindsight, they are being “tested”. This erasure will take place right after the transcription. Our intention is also to erase the transcripts four years from the time of collecting the data. We will seek the advice of the UI ethics committee on the study and in particular on this handling of the data.