The Pulse of Learning Analytics
Understandings and Expectations from the Stakeholders

Hendrik Drachsler
Open University of the Netherland
Valkenburgerweg 177
6419AT Heerlen, Netherlands
+31 (0)45 576-21-74
hendrik.drachsler@ou.nl

Wolfgang Greller
Open University of the Netherlands
Valkenburgerweg 177
6419AT Heerlen, Netherlands
+31 (0)45 576-21-74
wolfgang.greller@ou.nl

ABSTRACT
While there is currently much buzz about the new field of learning analytics [19] and the potential it holds for benefiting teaching and learning, the impression one currently gets is that there is also much uncertainty and hesitation, even extending to scepticism. A clear common understanding and vision for the domain has not yet formed among the educator and research community. To investigate this situation, we distributed a stakeholder survey in September 2011 to an international audience from different sectors of education. The findings provide some further insights into the current level of understanding and expectations toward learning analytics among stakeholders. The survey was scaffolded by a conceptual framework on learning analytics that was developed based on a recent literature review. It divides the domain of learning analytics into six critical dimensions. The preliminary survey among 156 educational practitioners and researchers mostly from the higher education sector reveals substantial uncertainties in learning analytics.

In this article, we first briefly introduce the learning analytics framework and its six domains that formed the backbone structure to our survey. Afterwards, we describe the method and key results of the learning analytics questionnaire and draw further conclusions for the field in research and practice. The article finishes with plans for future research on the questionnaire and the publication of both data and the questions for others to utilize.

Categories and Subject Descriptors

General Terms

Keywords
Learning analytics, survey, understanding, expectations, attitude, privacy, learning technologies, innovation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

INTRODUCTION
Despite the great enthusiasm that is currently surrounding learning analytics, it also raises substantial questions for research. In addition to technically-focused research questions such as the compatibility of educational datasets, or the comparability and adequacy of algorithmic and technological approaches, there remain several ‘softer’ issues and problem areas that influence the acceptance and the impact of learning analytics. Among these are questions of data ownership and openness, ethical use and dangers of abuse, and the demand for new key competences to interpret and act on learning analytics results.

This motivated us to identify the six critical dimensions (soft and hard) of learning analytics, which need to be covered by the design to ensure an appropriate exploitation of learning analytics in an educationally beneficial way. In a submitted article to the special issue on learning analytics [5], we developed the idea of a conceptual framework encapsulating the design requirements for the practical application of learning analytics. The framework models the domain in six critical dimensions, each of which is subdivided into sub-dimensions or instantiations. Figure 1. below graphically represents the framework. In brief, the dimensions of the framework contain the following perspectives:

- Stakeholders: the contributors and beneficiaries of learning analytics. The stakeholder dimension includes data clients as well as data subjects. Data clients are the beneficiaries of the learning analytics process who are entitled and meant to act upon the outcome (e.g. teachers). Conversely, the data subjects are the suppliers of data, normally through their browsing and interaction behaviour (e.g. learners). In some cases, data clients and subjects can be the same, e.g. in a reflection scenario.

- Objectives: set goals that one wants to achieve. The main opportunities for learning analytics as a domain are to unveil and contextualise so far hidden information out of the educational data and prepare it for the different stakeholders. Monitoring and comparing information flows and social interactions can offer new insights for learners as well as improve organisational effectiveness and efficiency [23]. This new kind of information can support individual learning processes but also organisational knowledge management processes as describe in [25]. We distinguish two fundamentally different objectives: reflection and prediction. Reflection [14] is seen here as the critical self-evaluation of a data client as indicated by their own datasets in order to obtain self-knowledge. Prediction [11] can lead to earlier intervention (e.g. to prevent drop-out), or adapted activities or services.
- **Data**: the educational datasets and their environment in which they occur and are shared. Learning analytics takes advantage of available datasets from different educational systems. Institutions already possess a large amount of student data, and use these for different purposes, among which administering student progress and reporting to receive funding from the public authorities are the most commonly known. Linking such available datasets would facilitate the development of mash-up applications that can lead to more learner-oriented services and therefore improved personalization [24]. However, most of the data produced in institutions is protected, and the protection of student data and created learning artefacts is a high priority for IT services departments. Nevertheless, similar to Open Access publishing and related movements, calls for more openness of educational datasets have already been brought forward [13]. Anonymisation is one means of creating access to so-called Open Data. How open educational data should be, requires a wider debate but, already in 2010, several data initiatives (dataTEL, LinkedEducation) began making more educational data publicly available. A state of the art overview of educational datasets can be found in [15].

- **Method**: technologies, algorithms, and theories that carry the analysis. Different technologies can be applied in the development of educational services and applications that support the objectives of the different educational stakeholders. Learning analytics takes advantage of so-called information retrieval technologies like educational data mining (EDM) [20] [17], machine learning, or classical statistical analysis techniques in combination with visualization techniques [18]. Under the dimension ‘Methods’ in our model, we also include theoretical constructs by which we mean different ways of approaching data. These ways in the broadest sense “translate” raw data into information. The quality of the output information and its usefulness to the stakeholders depend heavily on the methods chosen.

- **Constraints**: restrictions or potential limitations for anticipated benefits. New ethical and privacy issues arise when applying learning analytics in education [16]. These are challenging and highly sensitive topics when talking about datasets, as described in [13]. The feeling of endangered privacy may lead to resistance from data subjects toward new developments in learning analytics. In order to use data in the context of learning analytics in an acceptable and compliant way, policies and guidelines need to be developed that protect the data from abuse. Legal data and privacy protection may require that data subjects give their explicit and informed consent and opt-into data gathering activities or have the possibility to opt-out and have their data removed from the dataset. At the same time, as much coverage of the datasets as possible is desirable.

- **Competences**: user requirements to exploit the benefits. In order to make learning analytics an effective tool for educational practice, it is important to recognise that learning analytics ends with the presentation of algorithmically attained results that require interpretation [21] [22]. There are many ways to interpret data and base consecutive decisions and actions on it, but only some of them will lead to benefits and to improved learning [25]. Basic numeric and other literacies, as well as ethical understanding are not enough to realise the benefits that learning analytics has to offer. Therefore, the optimal exploitation of learning analytics data requires some high level competences in this direction, but interpretative and critical evaluation skills are to-date not a standard competence for the stakeholders, whence it may remain unclear to them what to do as a consequence of a learning analytics outcome.

To further substantiate the currently dominant views on the emerging domain, we turned to the wider education community for feedback. The aim was to extrapolate the diverse opinions from different sub-groups and roles (e.g. researchers, teachers, managers, etc.) in order to see: (a) what the current understandings and the expectations of the different target groups are and (b) if a common understanding of learning analytics has already been developed.

The mentioned framework was used to structure the questionnaire in order to avoid as much as possible bias toward a single perspective of learning analytics, e.g. the data technologies, and in order to get a balanced overview of the field as a whole. The questionnaire took concrete aspects into focus in the following way: The ‘stakeholders’ dimension inquired about the expected beneficiaries; ‘objectives’ tried to highlight the preference between reflective use of analytics and prediction; It also checked for the development areas where benefits are most likely or are expected; the ‘data’ section looked into stances on sharing and access to datasets; ‘methods’ investigated trust in technology and algorithmic approaches; ‘constraints’ focused on observations on ethical and privacy limitations (so-called soft-barriers); and, finally, ‘competences’ looked into the confidence for exploiting the results of analytics in beneficial ways.

Although we won’t go into this issue in this article, we are aware that there may be cultural, organizational, and personal differences that influence the subjective evaluation of the dimensions of learning analytics.

In section two below, we go on to describe in more detail the set-up of the questionnaire, the participants and the distribution method. In section three, we present and discuss results and statistical effects.

### 2. EMPIRICAL APPROACH

#### 2.1 Method

To evaluate the understanding and expectations of the educator and research community in learning analytics, we decided to use a questionnaire for reasons of ease of distribution and world-wide reach. In a globally distributed learning analytics community, this promised the best effort-return ratio, as opposed to other deeper, but more restrictive and effort intensive methods such as interviews. However, we anticipated the questionnaire as a first
To give the questionnaire an organized structure that would capture the domain in its entirety, where pedagogic and personal perceptions would have equal attendance to technical challenges or legalistic barriers, we divided the instrument into the six dimensions as indicated by the framework (see above). For each dimension, we asked the participants two-three questions and offered the opportunity for open comments. Questions were formulated in a variety of types, including prioritization lists (rank order), Likert scales, matrix tables, and multiple and single choice questions.

Another criterion we felt necessary to adhere to in our evaluation of the current perception of learning analytics as a domain was openness. Rather than selecting a handful of renowned experts in the field, or to involve a particular education sector or even a local school (which would most probably have just revealed a widespread ignorance about this developing research domain), we wanted to compile an overview cutting across national, cultural, sectorial boundaries, and even roles of people involved in learning analytics. Although this would unavoidably lead to a much fuzzier picture, we felt the benefits to our understanding of the interest and hesitations toward learning analytics, in what is a general trend to much wider open educational practices, outweighed such concerns, allowing us to better assess the potential impact of learning analytics to education.

Before publishing it, the questionnaire was validated in a small internal pilot with two PhD students and two senior staff members within the newly founded Learning Analytics research unit in our institution. In order to reach a wide network of a globally distributed Community of Practice, we designed and hosted the questionnaire online, using the free limited version of Qualtrics (qualtrics.com). This online tool is pleasantly designed and easy to use. It provides several sophisticated question-answer types with more being available for premium users. The data and the questionnaire are exportable in a number of popular formats including MS Excel and SPSS. The free version came with a limitation of 250 responses. All excess answers were recorded, but discarded in the analysis and data export. In our case, with a small sampling community, the free version proved to be sufficient.

### 2.2 Reach

We first promoted the questionnaire in a learning analytics seminar at the Dutch SURF foundation, a national body for driving innovation in education in the Netherlands. We then went on to distribute the questionnaire through the JISC network in the UK and via social media channels of relevant networks like the Google group on learning analytics, the SIG dataTEL at TELeurope, the Adaptive Hypermedia and the Dutch computer science (SIKS) mailing lists and to participants in international massive open online courses (MOOCs) in technology enhanced learning (TEL) using social network channels like facebook, twitter, LinkedIn, and XING. This distribution method is reflected in the constituency reached, in that there is, for example, a limited response rate from Romance countries (France, Iberia, Latin America) against a high return from Anglo-Saxon countries. The lack of responses from countries like Russia, China or India, maybe due to a number of factors: the distribution networks not reaching these countries, the language of the questionnaire (English), or a general lack of awareness of learning analytics in these countries. Still, we found that with the numbers of returns, we received a meaningful number of people interested in the domain. 

<table>
<thead>
<tr>
<th>Domains</th>
<th>Questions</th>
<th>Answer types and data range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stakeholders</td>
<td>Q3.2 In your opinion, who will benefit the most from Learning Analytics?</td>
<td>Rank order</td>
</tr>
<tr>
<td></td>
<td>Q3.3 Which relationships between stakeholder groups do you think Learning Analytics will strengthen most?</td>
<td>Side by side correlation</td>
</tr>
<tr>
<td></td>
<td>Q4.2 In your opinion, how much will Learning Analytics influence the following areas?</td>
<td>Matrix table with Likert scale (don’t know - 1, not at all - 2, a little - 3, very much - 4)</td>
</tr>
<tr>
<td></td>
<td>Q3.4 Which added innovation can Learning Analytics bring to educational technologies?</td>
<td>Rank order</td>
</tr>
<tr>
<td></td>
<td>Q4.4 What should the main objective for Learning Analytics be?</td>
<td>Single choice</td>
</tr>
<tr>
<td></td>
<td>Q4.5 When thinking of impact, how much attention should Learning Analytics pay to...?</td>
<td>Slider (0 - 100) (0 = not desirable, 100 = very desirable)</td>
</tr>
<tr>
<td>Objectives</td>
<td>Q5.2 How important do you consider the following attributes of educational datasets?</td>
<td>Matrix table with Likert scale (don’t know - 1, not important - 2, important - 3, highly important - 4)</td>
</tr>
<tr>
<td></td>
<td>Q5.3 Which IT systems does your organization use for teaching and learning?</td>
<td>Multiple choice</td>
</tr>
<tr>
<td></td>
<td>Q5.4 Would you share educational datasets of learners openly in an international research community, provided these datasets are anonymised according to standard principles?</td>
<td>Single choice (‘Yes, No, Don’t know’)</td>
</tr>
<tr>
<td>Data</td>
<td>Q6.2 Learning Analytics is based on algorithms, a, and theories that translate data into meaningful information. How hopeful are you that these methods produce an accurate picture of learning in the following areas?</td>
<td>Slider (0-100) (0 = I don’t think so, 100 = I’m certain of it)</td>
</tr>
<tr>
<td>Method</td>
<td>Q7.2 In your opinion, how much will Learning Analytics influence the following subjects?</td>
<td>Matrix table with Likert scale (don’t know - 1, not at all - 2, a little - 3, very much - 4)</td>
</tr>
<tr>
<td></td>
<td>Q3.3 Learning Analytics utilises data which is protected by data protection legislation. Do you think a national standard anonymisation process would alleviate fears of data abuse?</td>
<td>Single choice (‘Yes, No, Don’t know’)</td>
</tr>
<tr>
<td></td>
<td>Q3.4 Does your institution have and operate ethical guidelines that regulate the use of student data (for example in research)?</td>
<td>Single choice (‘Yes, No, Don’t know’)</td>
</tr>
<tr>
<td></td>
<td>Q5.5 Should institutions internally share all the data they hold on students with all members of staff (academic, technical, administrative, research)?</td>
<td>Single choice (‘Yes, No, Don’t know’)</td>
</tr>
<tr>
<td>Constraints</td>
<td>Q6.6 How much do you feel Learning Analytics and automated data gathering affect the privacy and rights of individuals?</td>
<td>Matrix table with Likert scale (Don’t know -1, not at all -2, a little -3, very much -4)</td>
</tr>
<tr>
<td>Competences</td>
<td>Q8.2 How important do you consider the following skills to be present in learners to benefit from Learning Analytics?</td>
<td>Matrix table with Likert scale (don’t know -1, not important -2, important -3, highly important -4)</td>
</tr>
<tr>
<td></td>
<td>Q9.3 Do you think the majority of learners would be competent enough to learn from their Learning Analytics reports?</td>
<td>Single choice (‘Yes, No, Don’t know’)</td>
</tr>
<tr>
<td></td>
<td>Q8.4 How sure can we be that end users will draw the right conclusions from their data, and decide on the best course of action?</td>
<td>Slider (0-100) (0 = Range not sure, 100= absolutely certain)</td>
</tr>
</tbody>
</table>

For a more representative study the dissemination of the questionnaire should be supported as well over public bodies like school foundations and governmental institutions.
The survey was available for four weeks, during September 2011. After removal of invalid responses we analysed answers from 156 participants, with 121 people (78%) completing the survey in full. In total, the survey now covers responses from 31 countries, with the highest concentrations in the UK (38), the US (30), and the Netherlands (22) (see Figure 2. below).

![Figure 2. Geographic distribution of responses](image)

### 2.3 Participants

Although we tried to promote the questionnaire equally to schools, universities and other education sectors, including e-learning companies, we received a significantly higher response from the tertiary sector (further and higher education) with 74% (n=116). It is probably fair to say that learning analytics as a topic is not yet popular or well-known in other educational sectors with the combined K-12 sector amounting to 9% (n=13) and some 11% (n=17) coming from the adult, vocational, and commercial sectors. The remaining 6% (n=9) in the ‘other’ subgroup includes cross-sector and other individuals, such as retirees from the education sector.

The only other demographic information we collected from participants was their role in the home institution. Here we received a broad variety of stakeholder groups that deal with learning analytics. Multiple answers were possible, taking into account that people may have more than one role in their organisation.

The three largest groups of our test sample were teachers with 44% (n=68), followed by researchers with 36% (n=56) and learning designers with 26% (n=41). With 16.1% (n=32) senior managers too were identified as a representative group of which two thirds (65.6%) came from HE institutions. 40.4% of the 156 participants claimed more than one role in their institution, of which again 40.3% were teacher/researchers (16.7% of the total sample).

Next, we’ll present the most relevant results from the online questionnaire regarding expectations and understanding of learning analytics.

### 3. RESULTS

Our report on the results is organized along the lines of the six dimensions of the learning analytics framework (cf. section 1 above). We paid special attention to mapping opinions against institutional roles in order to identify any significant agreement or discord in each of the dimensions.

One uncertainty underlying the outcomes is the lack of an established domain definition and/or established domain boundaries through practice. The term “learning analytics” is still rather vague, shared practice in the area is only just emerging and a scientifically agreed definition lacking. From on-going research and development work we know that some researchers subsume for example educational business analysis or academic analytics [8], or action analytics [7] under learning analytics [2]. Thus, the domain name itself carries a highly subjective interpretation, which almost certainly influenced the answers in the survey. We have no doubt that as the domain matures further, this interpretation will be narrowed down, leading to a better graspable scope and possibly more congruencies in the responses.

#### 3.1 Stakeholders

In this section, we wanted to know: (a) who was expected to benefit the most from learning analytics, and, (b) how much will learning analytics influence specific bilateral relationships?

Regarding the prioritisation of the stakeholder of learning analytics, the majority of respondents agreed that learners and teachers were the main beneficiaries of learning analytics where 1 was the highest score on the Likert scale. The weighting of the 155 responses shows that learners were rated highest at 1.9 mean rank, followed by teachers with 2.1. However, the ranking distribution and standard deviation for learners was higher (1.12) than for teachers (0.88). Institutions came in third place with an average rank of 2.6. There was also substantial contribution to the ‘other’ category with suggestions for further beneficiaries. Among those and most prominent were government and funding bodies, but also employers and support staff were mentioned.

![Graph 1. Relationships affected (1)](image)

Graph 1 above illustrates the outcomes of question (b) and confirms the findings of question (a) above. The peaks identify the anticipated intensity of the relationship. Relationships with parents are not seen as majorly impacted, which is probably due to the fading influence parents have in tertiary education. It would be interesting to complete this picture with more responses from the K-12 domain. The highest impact is seen in the teacher - student relationship (83.5%, n=111, of respondents emphasised this), whereas the reverse student - teacher connection is strengthened slightly less (63.2%, n=84). Only less than half the participants see peer relationships as being strengthened through learning analytics: learner - learner by 45.9% (n=61), and teacher - teacher by 41.4% (n=55). At roughly the same level comes the relationship between institution and teachers (46.6%, n=62).
In the spider diagram (graph 2 above), the area indicates that it is the relationships of teachers that are expected to be most widely affected, followed by learners, institutions, and parents at a minimal level.

### 3.2 Objectives

In this section, we asked participants in which way learning analytics will change educational practice in particular areas. Of the total answers given in all 13 areas \((n=1543)\), collected from 119 participants, only 10.8% of responses anticipated no change at all. On the other hand, the remaining responses left it open whether the expected changes will be small (43.8%) or extensive (45.4%).

#### Graph 3. Objectives for learning analytics

Looking at the individual areas (cf. graph 3 above), the highest impact was expected in more timely information about the learning progress (item 2), and better insight by institutions on what’s happening in a course (item 8). On the bottom end were expectations with respect to assessment and grading (items 6 and 5), where the least changes were anticipated.

Further, we contrasted the importance of three generic objectives for learning analytics: (a) reflection, (b) prediction, (c) unveil hidden information. 47% \((n=61)\) of the respondents felt that stimulating reflection in stakeholders about their own performance was the most important goal to achieve, while 37% \((n=48)\) expressed the hope that learning analytics would unveil hidden information about learners (cf. graph 4). Both are not necessarily in contradiction to each other, since insights into new information can be seen as motivator for reflection. However the case may be, only 16% \((n=20)\) favoured the prediction of a learner’s performance or adaptive support as a key objective.

### Graph 4. Generic preference

When looking at these objectives from the perspective of the different roles of participants, we find that teachers show a fairly equal interest in unveiling hidden information 44.6% \((n=25)\), and in reflection 37.5% \((n=21)\). This is a reasonable finding as many teachers expect learning analytics to support them in their daily teaching practice by offering additional indicators that go beyond reflection processes. On the other hand, 60.4% \((n=29)\) of researchers indicated a clear preference for reflection.

Translated into technological development, the expectations favoured more adaptive systems (highest rank), followed by data visualisations of learning, and better content recommendations in third place. Further interesting suggestions were “learning paths/styles adopted by students”, the clustering of learning types, and applications for the acknowledgement of prior learning.

A further question surveyed the perception of learning analytics being a formal or less-formal instrument for institutions. In two intermixed sets of three options, one set represented formal institutional criteria: mainstream activities, standards, and quality assurance, all relating to typically tightly integrated domains that are governed by institutional business processes and strategies. The other set contained three less-formal and less monitored areas of pedagogic creativity, innovation, and educational experimentation. All three items represented individual choice of staff members to be innovative, experimental, and creative in their lesson planning and teaching activities. As indicated in graph 5 below, among the 129 responses, there was a noticeable preference towards less formal institutional use of learning analytics at a ratio of 55:45 per cent. Quality assurance ranked highest in importance among the formal criteria, whereas innovation was seen as most important aspect of all criteria.

### Graph 5. Formality versus innovation
One participant summed up the situation of these findings in the following statement: “It would be easy for learning analytics to become a numbers game focused on QA, training/instruction and rankings charts, so promoting its creative and adaptive potential for lifelong HE/professional-life learning is going to be key for the sector - unless learning analytics people want to spend all their lives doing statistical analysis.”

### 3.3 Educational data

The section on data investigated the parameters for sharing datasets in and across institutions. The potential of shareable educational datasets as benchmarking tools for technology enhanced learning is explicitly addressed by the Special Interest Group (SIG) dataTEL of the European Association of Technology Enhanced Learning (EATEL) and has been demonstrated in [11].

Sharing of learning analytics data is impeded by the lack of some standard features and attributes that allows the re-use and re-interpretation of data and their applied algorithms [3]. For researchers, the most important feature was the availability of added context information (n=43, means 3.42) with a maximum value of 4 on the Likert scale. Perhaps, equally unsurprising was that for the manager group sharing within the institution (n=16, means 3.63) and anonymisation (n=19, means 3.53) were the most important values. Teachers, on the other hand, valued context (n=52, means 3.42) and meta-information (n=47 means 3.47) the most. At the other end of the spectrum, version control was the least important attribute across all constituencies (n=106, means 2.93). However, despite ‘version control of educational datasets’ was ranked the lowest, we still believe that this will play an important role in an educational data future. Version controlled datasets will offer additional insights into reflection and improvements through learning analytics by comparing older and newer datasets.

Graph 6 illustrates the importance of the given data attributes. Note that the notion of “important” outweighs the “highly important” overall, which results in a lower means value.

---

**Graph 6. Data attributes**

To get an idea of existing educational data, we asked participants about their institutional IT systems. For learning analytics, the landscape of data systems will play an important part in information sharing and comparison between institutions.

In the tertiary education sector alone (Further and Higher Education), 93.9% (n=92) reported an institutional learning management system, which made this the most popular data platform by far. This was followed by a student information system 62.2% (n=61) and the use of third-party services such as Google Docs or Facebook 53.1% (n=52). Table 2 below shows a summary inventory of institutional systems in use across all sectors of education covered in our demographics.

We assume that the more widely available a type of system is, the more potential it would hold for inter-institutional sharing of data, which could be utilised for comparison of educational practices or success factors. However, such sharing would depend on the willingness of institutions to share educational datasets with each other. When asked this question, a majority of people (86.6%, n=71) were happy to share data when anonymised according to standard principles.

<table>
<thead>
<tr>
<th>Table 2. Data systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer</td>
</tr>
<tr>
<td>Learning management system (e.g., Blackboard, Moodle)</td>
</tr>
<tr>
<td>Student information system</td>
</tr>
<tr>
<td>External services (e.g., Google docs, Facebook, Twitter, slideshare, iTunes U)</td>
</tr>
<tr>
<td>Intranet</td>
</tr>
<tr>
<td>Wiki platform</td>
</tr>
<tr>
<td>Course management system</td>
</tr>
<tr>
<td>Social networking platform</td>
</tr>
<tr>
<td>E-portfolio system (e.g., Mahara)</td>
</tr>
<tr>
<td>Mobile platform (e.g., Apps, e-books)</td>
</tr>
<tr>
<td>other (please specify)</td>
</tr>
<tr>
<td>Sensor data (e.g., location data)</td>
</tr>
</tbody>
</table>

What is slightly contradictory is that people who indicated before that anonymisation was not an important attribute for data are less inclined to share (n=18, 83.3% yes : 16.7% no) than people who felt that it was highly important (n=40, 92.5% yes : 7.5% no).

### 3.4 Methods

Learning analytics is based on algorithms (formulas), methods, and theories that translate data into meaningful information. Because these methods involve bias [1], the questionnaire investigated the trust people put into a quantitative analysis and in accurate and appropriate results. Within the 100% rating range, where 100% would indicate total confidence and 0% no confidence at all, the responses were located at mid-range. Among the given choices, slightly higher trust was placed on the prediction of relevant learning resources. This may be due in analogy to the amazon.com recommendation model, which is well-known and widely trusted. Other recommendations, such as predictions on peers or performance were rated rather low. The percentage on the horizontal axis in graph 7 below shows the level of confidence.

**Graph 7. Confidence in accuracy**
One comment criticised that it was “disappointing that you included institutional markers, rather than personal ones for the learners, e.g. while learning outside the institution, which in my view are much more important and interesting”. We are not aware that the questions actually reflected an institution-centric perspective. At the same time, we still remain sceptical that analytics might currently be able to seamlessly capture learning in a distributed open environment, but mash-up personal learning environments are on the rise [12] and may soon provide suitable opportunities for personal learning analytics as has recently been presented in [6], and [9].

3.5 Constraints

The constraints section focuses on the mutual impact that wider use of learning analytics may have on a variety of soft barriers like privacy, ethics, data ownership, data openness, and transparency of education (see graph 8 below). It should provide more detailed information on potential restrictions or limitations for the anticipated benefits of learning analytics. Most of the participants agree that learning analytics will have some or very much influence on the mentioned characteristics. Only a few did not expect any effects on privacy (10.4%, n=96) and ethics (8.8%, n=102). The majority of the responses believe that learning analytics will have the biggest impact on data ownership (66.4%, n=107) and data openness (63%, n=108) followed by more transparency of education (61.3%, n=111).

Graph 8. Problem areas of learning analytics

After the general weighting of the expected impact on these constraints, we explicitly asked the participants how they estimated the influence of learning analytics and automated data gathering on the privacy rights of individuals by further describing what we mean with privacy rights in four statements (see graph 9 below).

From 123 responses it appears that there is much uncertainty about the influence of learning analytics on privacy rights (cf. graph 9). The answers are widely spread from ‘no effect at all’ until ‘very much effect’. But the majority of participants believe that learning analytics will influence all four privacy dimensions at least a little. By recoding the given answers into a negative voting (will have no effect) and a positive voting (will have an effect) we got a clearer picture of the expectations of the participants. Regarding statement 1, about two thirds (65.8%, n=81) believe that learning analytics will affect ‘privacy and personal affairs’. Equally, in statement 3 - ‘ownership and intellectual property rights’ - we can again see a clear majority (60.1%, n=74) convinced that these will be affected by learning analytics. Statement 2 - ‘Ethical principle, sex, political and religious’ and statement 4 - ‘Freedom of expression’ are close together, but with the majority in the negative, thus expressing they do not think that learning analytics affects these privacy domains (statement 2 negative effect size 54.5%, n=67; statement 4 negative effect size 53.7%, n=66). Taking further into account the large presence of ‘don’t know’ answers, we conclude that to most participants, the impact on privacy is not yet fully determinable.

Graph 9. Soft issues

To get further information on these pressing soft barriers, we wanted to know if the participants have already (a) an ethical board and guidelines that regulate the use of student data for research. Further, we wanted to know (b) if they trust anonymisation technologies, and finally (c) how they rate a concrete example for data access in their own organisation to test the two answers before.

Regarding (a), the majority of the participants 61% (n=75) indicated that they have an ethical board in place. Another 18% (n=22) said that they did not have such a body in place, whereas 21% (n=26) were unsure. Yet to us, such an organisational infrastructure represents an important starting point for more extended learning analytics research that is ethically backed up through proper procedure.

With respect to (b), we went on to ask the participants whether they thought a (national) standard anonymisation process would alleviate fears of data abuse. With 49% (n=60), the majority of the 123 participants showed high trust in anonymisation technologies, whereas 24% (n=29) did not believe that anonymisation would be effective to reduce data abuse. 21% (n=26) indicated they know too little to answer this question. This leads us to the interpretation that in case learning analytics utilises data that is protected by legislation, participants expect further development of effective anonymisation techniques to deal with this issue.

After having asked participants about ethical guidance and their trust in anonymisation, we tested with question (c) how the participants estimated the use of educational data within their organisation. We asked them whether institutions should allow every staff member to view student data internally in the organisation. In this, we received a significant negative response from the participants. 43% (n=53) did not want to allow all staff members to view student data, only 30% (n=37) did not see any problem with shared access. We also received 15% open text responses to this question that mainly emphasised the need for levelled access to student data in compliance with the law and ethical regulation and the strong need to anonymise data. The tenor in the comments strongly pointed to a “need to know” rational. That is to say that participants felt that only people who had good reasons to see such data should be permitted to access them. As one commentary phrased it: “Only if legitimately necessary and only for those who have a need to know”.

3.6 Competences

In our section on the competences dimension, we wanted to identify the key competences connected with learning analytics.
We also asked for the confidence experts have in the independence of learners to exploit learning analytics for their own benefit.

According to the learning analytics framework we suggested the following seven key skills: 1. Numerical skills, 2. IT literacy, 3. Critical reflection, 4. Evaluation skills, 5. Ethical thinking, 6. Analytical skills, 7. Self-directedness. We wanted to know which of these skills the participants find important to benefit from learning analytics? Graph 10 shows the spider diagram of the answers.

Graph 10. Competences

The participants found all mentioned skills rather important for learning analytics. By way of means ranking with a maximum value of 4 on the Likert scale, participants identified self-directedness (means 3.53) and critical reflection (means 3.42) as the most important competences required from beneficiaries. These were rated as ‘highly important’ by 59.3% (n=67) and 48.7% (n=57) respectively. Numerical skills (means 2.83) and ethical thinking (means 2.95) were on the bottom end of the scale. We consider this in line with the previous answers to ethical aspects of learning analytics.

In addition to the required skills, we wanted to know whether participants thought that learners were competent enough to independently learn from learning analytics reports. It turned out that a significant majority did not think that learners would be able to deal with learning analytics reports without additional support (70.2%, n=85). Only 21% (n=26) believed that learners were competent enough to do so. We have to admit that we did not ask the same question with respect to skills required of teachers. This would have been an interesting comparison at this point.

4. SUMMARY AND CONCLUSIONS

The current article reported the results of an exploratory community survey in learning analytics that aimed at extracting the perceptions, expectations and levels of understanding of stakeholders in the domain. Divided up into six different dimensions we came to a number of conclusions which we are going to present below.

- Stakeholders: Participants identified the main beneficiaries in learning analytics as learners and teachers followed by organisations. Furthermore, the majority of respondents agreed that the biggest benefits would be gained in the teacher-to-student relationship and that learners would almost certainly require teacher help to learn from an analysis and for taking the right course of action. This is rather surprising as learning analytics is seen by many researchers as an innovative liberating force that would be able to change traditional learning by reflection and peer support, thus strengthening independent and lifelong learning. This latter opinion on independence could be seen in the ‘objective’ section of the survey (cf. chapter 3.2 above) where the majority expressed a preference for learning analytics to pay special attention to non-formalised and innovative ways of teaching and learning. Yet, respondents expect less potential impact on the student-to-student and the teacher-to-teacher relationships. This current perspective may be affected by the scarcity of learning analytics applications that demonstrate the innovative possibilities for learning and teaching. Thus people may not have a clear point of reference as, for example, is the case for ‘social networks’ where an established group of competitive platforms already exists.

- Objectives: The survey concludes further that research on learning analytics should focus on reflection support. The attained results clearly emphasized the importance of ‘stimulating reflection in the stakeholders about their own performance’. This goal could be supported by revealing hitherto hidden information about learners, which was the second most important objective. At the same time more timely information, institutional insights, and insights into the learning context were other areas of interest to the constituency.

- Data: Our institutional inventory in chapter 3.3 gives an overview of the most widespread IT systems. These could be prioritised by learning analytics technologies to gain an institutional foothold. They also provide the best ground for inter-institutional data sharing. Anonymisation can perhaps be seen as the most important enabler for such sharing to happen. It is emphasised in a number of responses as the second most important data attribute and confirmed in the willingness of people to share if data is anonymised. For a clear majority anonymisation also reduces fears of privacy breaches through sharing (cf. chapter 3.5). On the other hand, when it comes to internal sharing with departments and operations’ units of the same institution, the use of available data will continue to be an uphill struggle, and, according to participants, require good justification. Here, perhaps, a clearer mandate to ethical boards may help. These are already widely in place.

- Methods: Chapter 3.4 on methods revealed that trust in learning analytics algorithms is not well developed. We interpret the mid-range return levels as hesitation towards “calculating” education and learning. What seems interesting to us is that the widely interpretable hope for gaining a comprehensive view on the learning progress was given the highest confidence, but perhaps this shows wishful thinking rather than a real expectation. Overall rather low was the expectations of impact on assessment. A majority of people did not see easier or more objective assessments coming out of learning analytics (cf. chapter 3.2). They were also not fully convinced that it would provide a good assessment of a learner’s state of knowledge (cf. chapter 3.4).

- Constraints: A large proportion of respondents thought learning analytics may lead to breaches of privacy and intrusion. Yet, they ranked privacy and ethical aspects as of lesser importance to consider (cf. chapter 3.5) or as belonging to further competence development (cf. chapter 3.6). However, data ownership was expressed as highly important. This may be interpreted in that way that if ownership of data lies with the learners themselves, there is no perceived risk for privacy or ethical abuse. In any case, it seems that many organisations have ethical boards and
guidelines in place. These may come to play an increasingly important role for institutional data exploitation since a large number of respondents trust that anonymisation of educational data is possible but not necessarily sufficient to enable full internal exploitation of the educational data within an organisation.

- Competences: In the area of competences, participants mainly stressed the importance of self-directedness, critical reflection, analytic skills, and evaluation skills. On the other hand, few believe that students already possess these skills. This indicates to us a need to support students in developing these learning analytics competences. In conclusion of this section we can say, that the results suggest that there is little faith that learning analytics will lead to more independence of learners to control and manage their learning process. This identifies a clear need to guide students to more self-directedness and critical reflection if learning analytics should be applied more broadly in education. This interpretation is quite in contrast with some suggestions made with respect to empowerment of learners through providing graphical reflection of the learning process and further access to additional information regarding their learning progress [4].

5. LIMITATIONS
We are aware of several limitations to both the questionnaire and the presented results. As has been mentioned in the introduction, there is a dominance of responses from the Higher Education sector that makes the study only partly representative for other educational domains. Another limitation is the virtual absence of students (undergraduate or secondary) although we did receive a tiny fraction of responses from lifelong learners. This makes the results of the survey biased toward a top down perspective on learning analytics. In an open environment, these shortcomings may still be interpreted as representative for the expressed interest and the pervasiveness of the topic in particular constituencies. Or, they may simply be due to the limited reach social networks like facebook, linkedin, twitter, etc. have with respect to the distribution of the questionnaire.

Furthermore, the survey only represents a select few Western cultures. We need to be aware that substantial differences exist in educational cultures and that learning is always local. It would perhaps provide for interesting future research to compare these results with other dominant education cultures. Furthermore, for a more comprehensive study, the dissemination of the questionnaire should be supported by public bodies like school foundations and governmental institutions as it can be expected that the used dissemination channels reached a more technically interested group of the stakeholders.

One hopefully time-restricted limitation is the low awareness of learning analytics among the target survey group and learners especially. With the rise of useful and popular learning analytics applications, we hope that this limitation will ease away over time and thus yield more concrete insights into the field of attitudes in learning analytics. As such, this survey can only be taken as an indicative insight of innovators and early adopters.

To address these limitations and to gain more valid insights into expectations from and understanding of learning analytics we intend to do further research and plan to target the K-12 and adult education sector more specifically. Additionally, investigating the student perspective more intensely might reveal interesting contrasts to the above reports.

6. FUTURE RESEARCH
As announced in the beginning, this short survey was conceived as a first step toward more refined research into perceptions and potential for learning analytics. After collecting a more representative and selective dataset, we plan to apply more advanced analysis methods like the Group Concept Mapping [10] to further analyse different stakeholder groups and to identify consensus about particular issues in learning analytics among them. Group Concept Mapping will allow identifying thematic groups within learning analytics and it allows making a clear distinction between different aspects of learning analytics.

Finally, we want to announce that the underlying datasets of this article and the planned extended version of this dataset will be made publicly available at the dspace.ou.nl repository (at http://dspace.ou.nl/handle/1820/3850). In that way, we would like to encourage the learning analytics community to gain additional insights from our dataset for the fast evolving of the learning analytics research topic.

7. ACKNOWLEDGMENTS
We would like to thank all participants in the survey and those who further disseminated it to their networks. Further, we want to thank the NeLLL funding body and the AlterEgo project that sponsored part of the authors’ efforts.

8. REFERENCES


