RESEARCH ARTICLE

How do different stakeholders utilise the same data? The case of school leavers’ and graduates’ information systems in three European countries

Abstract

This paper analyses how three European countries produce and use data within a specific educational policy field, that of school leaving and graduation. It compares how stakeholders in England, Finland and the Netherlands know what happens to the leavers’ from schools and universities. Through gathering evidence about the methodological underpinnings of the school leavers’ and graduates’ information systems (SLGIS) and whose data needs they aim to satisfy, this research provides insight into the discrepancies of data production. Moreover, as stakeholders from the policy and the institutional level are represented to a different extent when SLGIS are set up, the utility of the resulting data in their work is of differing degree. The paper models the problems of how school leavers’ and graduates’ information systems are set up and utilised thus suggesting some of the possible solutions. Using this particular example to highlight the discrepancies of data production and utilisation in terms of the focus, the timing, the methodology and sampling as well as the processes of dissemination. To provide a wider relevance, the paper outlines key issues around the utility of research evidence.

Key words: school leavers’ survey; data utilisation in education; education policy; comparative education
Introduction: definitions and research design

There are many references in policy discussions to the use of evidence and building changes on such information. However, little is known about the process of data production and utilisation or the role of different stakeholders. This paper aims to show the discrepancies of producing and using research data through comparing three national examples. The rationale behind such exercise is to understand the processes within which ‘evidence’ is produced, the data-needs of the different stakeholders, and how stakeholders interact over the information collected. If ‘evidence-based policies’ have real importance it would be crucial to understand their mechanisms and caveats. According to several authors there is little research detailing the actual process of data utilisation within policy-making (see for example: Davies et al., 2000, Pollitt, 2006, Askim, 2007). Moreover, Pollitt paints a rather disappointing picture,

Grand statements about the importance of performance information for democracy sit alongside extensive if patchy evidence that ministers, legislators and citizens rarely make use of the volumes of performance information now thrust upon them. (Pollitt, 2006: 48)

To further the discussion of how evidence is produced and used within different settings, this paper gives a novel perspective. Through gathering information about both the production and the utilisation of data, and analysing the views of two levels of stakeholders, it provides a well-rounded picture. This paper investigates school leavers’ and graduates’ information systems (abbreviated SLGIS throughout) as an example of research evidence that is of potential interest to both the policy and the institutional level stakeholders. Moreover, comparing three national contexts highlights whether the differences in data production and use are due to the characteristics of a particular setting or can be seen as more general.
This paper has five main sections. Somewhat unusually, the introduction first details the definitions and research design used to enquire about the SLGIS for the second section to then outline the wider literature that contextualises this research. The third section gives an overview of the national characteristics of SLGIS so that the fourth section can follow with the discussion of the findings comparing the three national settings. The fifth section draws concluding remarks and suggestions on how to take the results of this research further.

**Definition of ‘SLGIS’**

This paper investigates information systems from different European countries aiming to uncover the trajectories of school leavers’ and graduates’ after the completion of formal education. To investigate the more comprehensive, longer standing research programmes three criteria are used: the SLGIS must a) collect and analyse school leavers’ and graduates’ data at the national level; b) provide evidence of post-educational outcomes from both secondary and tertiary education; and c) gain data either from more than one cohort of young people or contact the same cohort several times. This research excludes SLGIS with a regional scope, covering a specific educational institution, or only one educational phase but not the others. Information systems with the sole purpose of gathering data on young people not in education, employment or training (NEET) are not analysed here.

**Comparative context of ‘researching research’**

This paper takes a comparative view on a) how different nation states set up a particular information system on education and b) how the results of this data collection is then utilised by different stakeholders. SLGIS are a good example for such investigation, given that some form of school leavers’ and graduates’ data collection exists in multiple countries. The
international comparative research frame used throughout this paper allows for the analysis of the educational and wider societal context of the SLGIS (Bray and Thomas, 1995, Phillips and Schweisfurth, 2006, Hantrais, 2009).

This paper draws on two approaches to address the problem of ‘researching research’. To understand how SLGIS are constructed and used within a complex policy field, the conceptual framework builds on enumerology as framed by Bogdan and Ksander (1980) and ethnostatistics as outlined by Gephart (1988, 2006). Both of these approaches assume an ethnographical perspective on the ‘organizationally and institutionally sanctioned nature, meaning, and use of quantitative practices and products’ (Gephart, 2006: 426). They also emphasise the stakeholders’ perceptions of data and information rather than a ‘third person’ perspective which views data production from a particular theoretical standpoint. These concepts complement each other: whereas enumerology is ‘concerned with (lay) production and use of counts in policy and evaluation studies’, ethnostatistics provides a more specific focus on ‘how social scientists produce and use statistics in research’ (Gephart, 2006: 426). Through applying them together the perspectives of both the data producers and the data users are covered: enumerology refers to a wider range of stakeholders who are involved in the processes of setting the rules of the information systems, whereas ethnostatistics refer to a more specific group of stakeholders, that of data-producers and expert data-users who employ the data in statistical analysis.

The international comparative research framework requires constant comparison of the topics of research outlined here as well as understanding them within their societal context. The research questions are in line with those proposed by Bogdan and Ksander’s (1980) enumerology and incorporate some of the focus points of ethnostatistics outlined by
Gephart (1988, 2006). The lines of enquiry for how counting a particular phenomenon comes about, how the process is negotiated and how the resulting evidence is then used are as follows:

**Before ‘counting’:**
- In what context does a phenomenon such as school leaving and graduation become worth counting?
- How are the main aims and the procedures of counting settled by the different stakeholders?

**Whilst ‘counting’:**
- How are statistics on school leaving and graduation created in terms of their methodological setup, and the financial and institutional background?
- Have stakeholders initiated change in the SLGIS? If yes, how and what changed?

**After ‘counting’:**
- How are the outcomes of SLGIS disseminated?
- How and to what extent are the statistics convincing and influencing the different stakeholders?

**Research design and sampling**

The combination of different sources and processes of data collection helps to uncover first the bigger picture of European SLGIS and provides an in-depth understanding of the SLGIS in their context (Gorard and Taylor, 2004). In the first, cross-sectional phase the research this
paper is based on uses publicly available documents to compare the research design and methodology of SLGIS within Europe. This first, ‘quantifiable’ phase provides static outcomes uncovering the differences among cases (for details see Hordósy, 2014). The second, case study phase compares multiple national SLGIS and their utilisation process, and using a thematic analysis it explores and explains the differences found in the first stage (Bryman, 2012). For a discussion of the process of Europeanisation in SLGIS see Hordósy (Forthcoming).

‘Especially illuminating examples’ are sought for the second phase through non-probability purposive sampling to represent information systems with different methodological approaches from a variety of European regions (Hakim, 2000: 170, Bryman, 2012). To maximise variation with a minimum number of cases, three nation-states are chosen from a typology of four created along research design (repeated cross-sectional vs. longitudinal designs) and population covered (sample vs. census). First, the Netherlands is chosen as an example from mainland Europe with a SLGIS based on samples of subsequent cohorts queried in cross-sectional snapshots. Second, Finland is selected from amongst the Northern-European countries having a census-type sampling frame and a longitudinal design. And third, England is chosen to represent research programmes with longitudinal designs based on samples of several cohorts from amongst the UK Home-countries and Ireland.

In terms of methods, élite interviews are used to gather information not otherwise systematically available along with contextual details and insider interpretation (Gillham, 2005) and they are supplemented with documentary data analysis. The documentary data used here are predominantly official research reports and descriptions of design and methods...
of the SLGIS. Participants for the interviews are chosen through combining purposive sampling with snowball sampling to reach hard-to-find experts with in-depth knowledge of the SLGIS. This process results in interviews with practitioners and administrators who are involved in producing as well as using the school leavers’ and graduates’ information; in a number of cases the interviewees suggest a group-interview format. Thus a total of 60 people in 44 interviews are consulted in the three countries in personal or group interviews, in the selected research organisations, ministries, schools and universities. Other experts as researchers with academic roles and data-specialists are also consulted. Table 1 shows the distribution of interviews amongst the three countries and different groups of stakeholders. Interviews or interviewees are referred to by the group of stakeholder, a number or a letter, and the abbreviation of the country, i.e. ‘Research institute N, NL’ and ‘Ministry D, EN’; or ‘University 8, FI’ and ‘Expert interview 5, EN’.

[Insert Table 1] Number of interviews conducted in the different case study countries

Limitations

Note that the case studies of the SLGIS are mainly built on the interviews conducted in the three countries. Although many stakeholders are consulted, a case study using 13 to 16 interviews and several documents cannot aim to provide all viewpoints, or the entirety of the views of those interviewed. Nor can this research aspire to uncover the interrelated nature and the history of the connections of the different stakeholders within these national contexts. The case studies indicate the difference in data-needs comparing the political and the institutional levels. Furthermore, this research cannot employ the full potential of the ethnostatistical and enumerological approach, which would involve observing of the flow of school leavers’ and
graduates’ data within different settings.

Research context: data utilisation by different stakeholders

To situate this research in the wider context this section first looks at similar studies that compare the methodologies of researching school leavers and graduates. Only one of these investigates how the resulting evidence is used. The following sections therefore look at the topic of data utilisation of different stakeholders more broadly.

Comparative studies on SLGIS

There are a few accounts that set the context for this research via comparing school leavers’ and graduates’ information systems. A first such report by Mainguet (1999) compared the methodology of school leavers’ and graduates information systems. Their work reviewed the then existing research programmes on school-to-work transition with a view on the research methods used, the sources of information, and the topics covered. They also reviewed the international research programmes tapping into the topic of transitions. A further account of how the data on school transitions are collected was published in 2001; and it reviewed the data needs of policy-makers and researchers (Raffe, 2001). A more recent account to contrast how school leavers’ and graduates’ data are acquired and used in the different nation states was published in 2012. As Gaebel et al. (2012: 16) suggest in the study entitled Tracking Learners’ and Graduates’ Progression Paths (TRACKIT), ‘the tracking of students and graduates has so far received little attention, at least at European level’. Their work is concerned with the methodology and how the data about current and former students is used
at the policy level as well as within higher education institutions in different European countries. A generic model for graduate tracking is outlined by them:

(... a national body, a research initiative or a consortium of higher education institutions provides a standard survey scheme and supports institutions in implementing it. The standard questionnaire could usually be augmented by the individual institutions. While they have a major role to play in ensuring a high response rate, the data would be collected and evaluated centrally. An institution would receive its own data, but could also benchmark itself (or ask to be benchmarked, depending on the approach) against institutions of a similar kind. Aggregated data could be published, and used for research or other purposes. (Gaebel et al., 2012: 40)

The main contribution of this paper is to put more emphasis on the data utilisation aspect of the SLGIS. The following two sections look at what type of data tends to be used by the different stakeholders and what sort of processes they are usually employed in.

What data is used by policy makers and institutions?

Data for policy-making can derive from primary research or secondary research built on a systematic review or meta-analysis. The primary sources are categorised along the methodological divide into quantitative, qualitative and pluralistic approaches by Davies et al (2000). The quantitative evidence dominates the discussion about ‘what works’ in different policy settings, whereas qualitative approaches are said to ‘address issues such as why one intervention may be better than another, and they can contribute to an understanding of the context within which policies must be framed and implemented’ (Davies et al., 2000: 10). According to Fitz-Gibbon (2000: 72-73) the type of information to answer the crucial question of ‘what works’ in education are ‘observational data from surveys; evidence from
specific evaluation projects; and the findings from true experimental research’. Information classified as ‘qualitative’ are suggested to carry persuasive power along with ‘setting agenda, raising issues and providing ‘thick’ descriptions’ of important problems within the educational system (Fitz-Gibbon, 2000, Pawson, 2006).

The research on institutional data utilisation regards mainly the use of assessment information and data on internal progression of students (see for example: Kelly and Downey, 2011, Schildkamp and Kuiper, 2010, Wayman and Stringfield, 2006). Accountability is the main external reason for institutions to use data, as authorities place ‘tremendous faith in the power of data—especially standardized test data—to effect school improvement’ (Young, 2006: 521). The main internal driver is to understand the processes behind attainment and attempt to raise the outcomes (Kelly and Downey, 2011, Kerr et al., 2006, Schildkamp and Visscher, 2010).

How is data used by policy makers and institutions?
Regarding the application of data in the policy-process, Weiss (1979) lists six different models of why and how this occurs based on the observation of policy-making. The knowledge-driven, the problem-solving and the interactive models deal with a more specific ‘policy problem’, whereas the tactical, the political and the enlightenment models refer to a broader perspective. Further ‘classifications’ of research utilisation suggest that it could be instrumental, conceptual, symbolic/mobilising/persuasive, and having a wider influence (Estabrooks, 1999, Nutley, Walter and Davies, 2003, Squires et al., 2011, Weiss, 1988).

A possible classification of the processes in which schools use assessment data is suggested by Schildkamp et al. (2012). Note that these mainly concern internal affairs, they
do not relate to external processes such as being accountable for the policy level, or informing the wider public:

- Monitoring
- Instructional and curricular decisions
- Supporting conversations [with colleagues, students and parents]
- Professional development
- Reflecting on one’s own functioning
- Policy development and planning
- Strategically (Schildkamp et al., 2012: 232, originally from: Schildkamp and Kuiper, 2010)

There is less evidence related to the higher education level regarding how data are used. There are a number of studies on whether and how league tables and rankings have changed institutional practices; what happens to course evaluation and student satisfaction data within the institution; and the procedure of ‘informing choice’ in higher education (see for example: Davies, 2012, Hazelkorn, 2013, Kember, Leung and Kwan, 2002, Newton, 2000, Turner, 2005). However, there is little evidence detailing how any of these datasets are actually applied within the institutional decision making processes.

The link between the policy and the institutional levels

Data utilisation does not exist in isolation at one or the other level of the educational system, as the gathered information provides the link between the local and the central levels (Ozga, 2009). The utilisation of data is crucial due to the governance turn described by Ozga (2009: 150) as the ‘shift in strategy that is highly dependent on the appearance of deregulation, but that is equally marked by strong central steering through various policy technologies’. There
is therefore an increased need to understand what data is used and how educational ‘policies
are interpreted and ‘translated’ by diverse policy actors in the school environment, rather than
simply implemented’ (Braun, Maguire, and Ball, 2010: 549).

This paper analyses how one particular source of evidence, namely the school leavers’
and graduates’ information systems are set up; whether and how policy makers apply the
resulting evidence; and how the policy-processes of using SLGIS data compare between
different national settings. This paper also analyses how this particular set of data are
employed within educational institutions – if at all.

The interconnectedness and interrelatedness of stakeholders in producing, using, or
ignoring, perhaps even abusing evidence are analysed to gain a better understanding of what
helps and what hinders data utilisation at the different levels. The research literature on the
policy level and the institutional level data utilisation analyses how assessment, internal
progression, final exam or student satisfaction data are applied. These are strongly related to
the internal issues of the institutions with direct links between data and the institutional
practice. However, this paper concerns data from ‘beyond’ the school or the university,
therefore the link between the results gained through the SLGIS and the institutional level is
less clear.

Context: SLGIS in the case study countries

This section provides an overview of the SLGIS from the case study countries. For further
discussion and comparison for all European countries see Hordósy (2014). With respect to
the research design and sampling differences, this is the list of information systems analysed:

Cross sectional sample survey:

- School-leaver and graduate surveys (Netherlands)
- Further Education Learner Destinations (England)
- Aarresaari Network graduate research (Finland)

Cross-sectional ‘census’:

- Destinations of Leavers from Higher Education (England/UK)
- College destinations data (England)

Longitudinal sample survey:

- Youth Cohort Study (England/Wales) &
- Longitudinal Study of Young People in England (England)
- Destinations of Leavers from Higher Education, Longitudinal (England/UK)

Longitudinal ‘census’:

- Statistics Finland: Transition from school to further education and work (Finland)

The Dutch SLGIS based on data collected in a cross-sectional manner at all levels of the educational system has existed for more than 20 years. The system is built on separate surveys measuring school leaving and graduation 1.5 years after the former student has left
the institution and these are then combined into a national-level comprehensive report of the educational system. The data also serve as one component of the labour market forecast studies (See for example: ROA, 2010. Further information from: Research institute X, NL, Research institute Y, NL).

The two English data collections that give a national picture on the notion of school leaving are supplemented by research focusing on specific educational sectors. The Youth Cohort Study (YCS) conducted from 1985 till the end of 2000s and the Longitudinal Study of Young People in England (LSYPE) started in 2004 provide an overall picture of the transition process from compulsory education to further education, higher education and the labour market as well as young people’s life in general using data from one or more cohorts (Research institute P, EN). Separate data collections exist for further and higher education gathering evidence on the initial labour market outcomes of college leavers’ and university graduates’ (Expert interview 3, EN, Expert interview 4, EN). These data collections aim to get a snapshot of a cohort’s initial destinations using surveys. The Destinations of Leavers from Higher Education (DLHE) aims to collect information from at least 4/5th of the graduates half a year after they left university to gain comparable data on the whole higher education sector (See for example: National Statistics, 2008, ESDS, 2011, ESDS, 2012, Wild, 2010, Shury and David, 2013. Further information from: Research institute Q, EN).

Through combining different registers, ‘systematic collection of unit-level data’ that can be updated with the school leavers’ and graduates’ information is produced on virtually everyone in Finland (United Nations, 2007: 47). This data collection is a census of all school leavers’ from upper-secondary schools and all graduates’ from higher education. The data are reported one year after leaving and the methodology allows choosing any other time-frame as
well. Although this type of robust information is available on all school leavers and graduates in Finland, there are further data collections present (Research institute R, FI). One is the ‘career follow-up’ conducted by the Aarresaari Network (AN), the national career guidance organisation. This research initiated by the universities covers longer term labour market outcomes of graduates. The other two approaches analysed in this research are conducted either right after graduation, or approximately a year after leaving HE. These research programmes are to measure the perceived quality of education and the initial labour market outcomes. They are called ‘exit-polls’ and ‘initial destinations’ surveys (See for example: Statistics Finland, 2004, Statistics Finland, 2012. Further information from: Research institute S, FI, Research institute T, FI).

As the three national SLGIS seem to mirror the structure of the corresponding educational systems, the differences in the latter can explain some of the variation between the SLGIS. Finland’s educational structure is represented in the SLGIS as a comprehensive system of all levels and sectors. This system is supplemented at the HE level with further SLGIS. The Dutch schooling system streams pupils into separate sectors and levels mirrored by the SLGIS being built on separate sectorial accounts. Similarly, the diversity of the English educational system and the strong separation of the different educational stages and sectors are reflected in separate data collections. However, whereas there is a national level picture of the Dutch system, the diverging educational sectors and levels of the English system are not comparable with the available separate datasets.

Discrepancies around the production and utilisation of the data
This paper analyses the discrepancies of data utilisation in a specific educational policy context whilst comparing three national settings, that of England, Finland and the Netherlands. Discrepancies in this paper refer to the difference in how the national information systems are viewed by groups of stakeholders who operate at the policy and the institutional levels. The stakeholders whose interests regarding the aims and data-needs of SLGIS may be in conflict are chiefly: elected politicians, bureaucrats responsible for policy implementation, statisticians and researchers running the SLGIS, institutional leaders and managers, institutional career guidance professionals, academic researchers and other professionals using the SLGIS as secondary data. The following two sub-sections analyse first the process of data collection, then the process of data utilisation in terms of what aspects of them the discrepancies appear in.

**Discrepancies of data collection**

SLGIS have been predominantly instigated and funded by national governments and thus seem to reflect the data-needs of central administrations. However, improving further education and employment outcomes for school leavers and graduates tends to be seen as an institutional duty, meaning that schools and universities are expected to act upon their results. This could be seen problematic due to first, the nature of leavers’ and graduates’ outcomes being realised predominantly ‘after’ and ‘beyond’ the control of the institutions. Second, as the following discussion suggests, the institutional data needs on leavers and graduates are not congruent with those of the policy level. The main aspects where disagreement seems to arise regard the focus, the methodology, the timing, the coverage and sampling frame, and the context of school leaving and graduation. The following list shows the ‘ideal type’ of school leavers’ and graduates’ data derived and synthesised from the three national case studies; the next sub-sections provide further comparison and discussion of these differences. Note that as
with such comparative lists, there are of course specific national data needs dependent on the policy and institutional differences not represented here; some of these will be alluded to throughout the discussion.

*Policy level data needs on school leaving and graduation*

- Shorter time frame
- Focus on initial labour market outcomes
- Trend data, Cross-sectional studies
- ‘Census’ of all leavers and graduates
- Covering the institutional level
- Data on the national level

*Institutional level data needs on school leaving and graduation*

- Longer time frame
- Focus on longer term labour market outcomes, satisfaction with school/university, skills and competences needed at the workplace
- Longer term cross-sectional studies or longitudinal studies
- ‘Samples’ to provide space for asking school leavers and graduates,
- Covering information at the course level/institutional level
- Socio-economic and regional context of outcomes
Main focus of SLGIS

The main focus of governments when instigating SLGIS is reported to be the initial labour market outcomes. Governments primarily interested in evaluating the preparation for employment provided by an education system would presumably want to wait until after the initial turbulence of leaving education and settling into a career. Employment status and earnings six months after graduation can be very different from employment and earnings 3-5 years later. However, the longer the gap between leaving education and the collection of school leavers’ and graduates’ information, the greater the risk that the ‘employability’ evidence is perceived to be out of date by the time the data reaches the users. There are also further practical considerations regarding higher sample sizes, and the dangers of losing contact with individuals once they leave formal education (Ministry D, EN, Ministry E, EN).

Timing of SLGIS

Whereas in the English case the different data systems of the separate institutional sectors operate with a shorter time-frame, and they take time points around the first few years after leaving or graduation, the Finnish SLGIS at the tertiary level take a wider range of time-points beyond having the Statistics Finland longitudinal data. In the Netherlands all the information systems have the very same timing to enhance comparability across the educational system. A researcher interview from the Netherlands suggested that contacting former students 1.5 years after they have left means that they ‘are in a more stable water’ either in their subsequent educational institution or at the labour market (Research institute X, NL).

Both in Finland and in England the institutional interviewees expressed their frustration regarding the timing of the SLGIS. Whereas institutions preferred the collection of
data 3-5 years after graduation, governments have insisted on a short time period (between six months and a year after graduation). Institutions argue that this time-frame is not enough to judge educational outcomes and therefore the validity of the data is questionable. As suggested regarding the English case:

It’s a trade-off, really: 6 months is long enough for people to have gone and got a job and still the contact details held by the institution are good enough to get into contact with these people, but 6 months isn’t long enough for them to have necessarily got the job that recognises the job that they were capable of getting following the qualification that they’ve got… (University 3, EN)

As the utility of the SLGIS partly derives from the data showing trends over time, there has been little change in the research designs of SLGIS. This means that although the processes of using the data have probably changed, subsequently arising data-needs had to be satisfied with the very same data or setting up further SLGIS. The comparison of the cases suggests that adding new SLGIS to the existing systems is less confrontational and can satisfy more data-needs, as it happened with the newer Finnish higher education information systems. Additional, in-depth SLGIS might provide background data that contextualises the national picture. This is of course dependent on available funding. Also, it is unclear to what extent newer information systems can become important along the previously existing ones, and what happens if they tell contrasting stories from the nationally acknowledged figures.

Methodology of SLGIS

Policy-makers seem to be interested mainly in the immediate outcomes of school leavers’ and graduates’, which is also apparent in the tendency of using trend-data rather than longitudinal accounts. For instance the Statistics Finland data are available for a longer time-
frame on virtually all school leavers’ and graduates’. However, at the time of the fieldwork in 2012 the university funding is still planned to be based on cross-sectional accounts of this information, comparing outcomes a year after graduation. One possible explanation for using trend-data could be that it reduces complexity. Longitudinal information seems to be applied in planning national policies, whereas the trend-data are used more in monitoring the system and evaluating educational institutions. Setting up prospective information systems are subject to one fundamental problem: trying to predict the future of which they will collect data about. A data-collection that is aimed to gather information for a longer time-frame has to be set up without appropriate available evidence as to what might be relevant or irrelevant later on.

In all case study countries there has been some push towards more economical data utilisation. This either means reducing costs by changing providers, or minor aspects of the research methods, or by cutting funding for further research entirely. Especially the English case suggests that if it is not the policy-makers immediate interest to run the SLGIS, the continuation of research funding is not self-evident (Research institute O, EN). Mainly due to the time-frame of trend-data, this approach risks the utility and the value of the SLGIS. Although the SLGIS capture the notion of career outcomes after schooling and university, the data that policy-makers are really interested in tells the story in short. Such approach provides some headline data but not a lot of in-depth explanations.

Coverage and sampling frame of SLGIS

The information systems analysed in this research are either built on sample surveys or the combination of registry datasets. Conducting surveys on the one hand provides tailored information and allows collecting attitudinal data from school leavers’ and graduates’
regarding their education and their current employment (Expert interview 9, FI). The main issues are reaching sufficient sample-sizes for institutional comparisons and that funding for such research might be vulnerable and subject to change (Research institute O, EN). Combining different registers on the other hand is an economical utilisation of the already existing data and is thus a relatively cheap account (Research institute R, FI). The main problems here regard a) data-protection and whether the exercise can be carried out within the national legal frameworks and b) to what extent the existing registers are suited to answer questions on school leaving and graduation (Expert interview 8, FI). Policy-makers and civil servants seem to be pushing for SLGIS that are based on a census (Ministry D, EN, Ministry F, FI). This, regardless of the research design means reducing the breadth and depth of the information gathered. From amongst the cases it is only the Netherlands where the amount of information covered and the target sample sizes seem to be in balance and the majority of the stakeholders are content with the SLGIS (Ministry A, NL, Ministry C, NL). However, in this case the sample-size reached raises some issues for the institutional and programme level utilisation (University 1, NL).

Context of SLGIS

An individual’s labour market outcomes could be interpreted in light of their socioeconomic background, as suggested by Raffe (2008). However, as the economic aspects of school leaving and graduation tend to be the focus, SLGIS analysed in this research are rarely contextualised in terms of social inequalities. For example, in the case of England, the YCS and the LSYPE data collections provide vast amounts of information on the respondent’s background – this is utilised mainly at the national level. The DLHE outcomes – used both by institutions and the policy level – could be connected to the socioeconomic background data collected when the prospective students apply to university. However, there
are problems with the quality and the coverage of this data to provide a national picture (Expert interview 3, EN).

In Finland due to the individual identification-numbers and the possibilities of combining administrative data between generations it would be feasible to analyse social mobility and the impact of parents’ education on the labour market outcomes of the children. However, due to limited resources, this kind of analysis is not currently published as a part of the official statistics. As suggested by researchers involved:

A: But we have also in the registers data on your parents, we can combine parents to...
based on the population register data, which can be combined to student and degree data, so we can see the parents’ educational status and so on.
I: Is this done, or is this a possibility?
A: We don’t publish it at the moment.
B: No, but we have done some of the statistics. We don’t have resources to all. (Research institute R, FI)

The Dutch SLGIS provide background data about the students and a longer-term view on school leaving and graduation. In the Dutch system adding further demographic data to the HBO and WO-level\(^1\) survey outcomes would be feasible. However, the majority of the stakeholders do not seem to use this in their work. It is only the research organisations who seem to be more interested in the equality issues flagged up in the SLGIS (Research institute X, NL).

Regarding the organisations that conduct SLGIS, there is substantial diversity between and within countries. SLGIS are run either through government departments, statistical agencies, separate research organisations, or the educational institutions
themselves. Ownership of the data production seem to parallel an interesting difference regarding the underpinning idea of the SLGIS. Whereas research organisations and to some extent, government departments seem to be more inclined to consider the sociological aspects of school leaving and graduation, the statistical agencies and the educational institutions focus on the economic returns to education.

Two focus points are of particular importance to provide context to the SLGIS: first, some feedback from former students on their education and second, some longer term information on how careers unfold. It is especially institutional leaders and managers who are interested in gaining feedback on the relevance of education at the labour market along with the skills and competences that are required – some of the SLGIS cover these aspects (University 2, NL). For the work of career guidance professionals both a longer term and a broader view on career choices is desirable (University 3, EN, University 4, EN). However, precisely these two aspects seem to be of lesser importance in the SLGIS analysed in this research.

**Discrepancies of data utilisation**

To follow the outline of discrepancies when collecting school leavers’ and graduates’ data, this section discusses issues when utilising the outcomes. The following discussion details the levels of comparability within national information systems, and the processes of using the SLGIS outcomes. Building on these, the discrepancies of data utilisation are modelled. The section closes with a discussion of how change within SLGIS affects views on their utility.
Comparing at the national and the institutional level

An important question raised with respect to all national SLGIS is comparability. SLGIS that aim to provide a system level picture covering all educational sectors permit evaluation of the whole educational structure, thus potentially raising the policy makers’ responsibility in improving the outcomes. When the separate sectorial accounts are not comparable due to difference in design and methodology, the responsibility of achieving better outcomes could be pushed more on the institutions.

The question of comparability within and across sectors and levels of education arises in terms of what information the institutions themselves have, and to what extent they can benchmark their outcomes to others or to the national (sectorial) average. Both the Dutch and the Finnish information systems provide a national picture on school leaving and graduation. However, whereas the Dutch SLGIS provides detailed information at a given time-point, the Finnish statistical dataset provides limited information on longer term outcomes. The Dutch SLGIS is described by one of the researchers as follows:

(…) there is a part of the questionnaire, (…) which is precisely identical for all the different questionnaires. (…) So that you can compare higher vocational education, secondary vocational education, measured at the same period, measured the same group in the sense how long that they are in the labour market and measured precisely by the same questions. That gives you tremendous possibilities to compare it. (Research institute X, NL)

The English data collections of the different educational sectors and levels do not offer a comprehensive picture; it is only the YCS and the LSYPE that serve these purposes to a limited extent. Note that at the time of fieldwork for this research in early 2012 the data linking of different administrative datasets across all educational sectors is not feasible yet;
such datasets are published on the Department for Education websites at the end of 2013 (GOV.UK, 2013). Whereas in the Netherlands and Finland the same data are used in the policy processes and within institutions, in England these two procedures are more divorced for the secondary sector.

From the perspective of the schools and universities, data on others is useful when benchmarking themselves to institutions or to a national average, and using this information in their institutional decision making procedures, as well as when ‘informing choice’ through their marketing practices. In terms of benchmarking, at the secondary level the data rarely allows for such comparisons. Even if schools have comparable information as in the case of Finland, they do not seem to apply the data in this way. The datasets available for tertiary level, however, allow claims about being above or below a national, sectorial, or study programme average in all three case study countries. This information on graduate outcomes can feed into the institutional strategy planning as seen in the case of the Netherlands (University 1, NL) or help initiating change as an example from an English university shows:

I quite unashamedly use it [the DLHE] within the institution, look, this is the trend-data, over this period the last 3 years this is happening in department X, I don’t know what’s going on, everybody else is doing really well, what’s going on here? So I use it a bit of a ‘pull your finger out’ and get yourself sorted out. […] it’s an influencer, but it’s a very useful influencer. (University 5, EN)

Comparing this way means by definition having half below and half above the median; admittedly, the results of SLGIS are used as marketing tools only when they favour the institution. Especially in the case of not having marketable results, the majority of the
schools and universities use other sources of information on how their study programmes connect to working life through contacts with the labour market. This procedure is not comparable across institutions and it is impossible to trace the sources or assess the validity of the information.

Processes of data utilisation

One of the aims of this research is to uncover the different processes in which the potential stakeholders apply the SLGIS outcomes regarding school leaving and graduation, similar to those provided by Schildkamp and colleagues (2010) for the school level use of performance data.

Policy level data utilisation of SLGIS

- Monitoring the educational system: the data is one of the ‘actors’ in this process
- Planning national and sectorial policies: used as a key information especially in crises
- Evaluating national policy-making
- Tool for evaluating/auditing educational institutions
- Informing students’ choice through SLGIS results

Institutional level data utilisation of SLGIS

- Monitoring institutional progress: opinion of former students about their education and their success in gaining employment
- Institutional decision making: more tentative aspect for longer term strategy
- Institutional planning of finances: help to target resource (especially career guidance)
- Complying with evaluation requirements: SLGIS as an indicator in evaluation or audit
Informing future students (marketing) and the general public: marketed when the data tells a positive story mainly; national comparisons publish all

Informing current students (career guidance): labour market outcomes of prior students as possible routes

Keeping in touch with former students: alumni services

Other stakeholders’ data utilisation:

Using the SLGIS as secondary data

Modelling the discrepancies in data utilisation

In some cases the groups of stakeholders make use of the data in a similar manner, as outlined previously; however, as shown here, they might require different amounts and types of data for their work. The previous sections alluded to some of the issues with the SLGIS serving stakeholders who might need them for very different purposes, using them in diverging processes, requiring dissimilar types and amounts of data. Figure 1 provides an overview of the main processes of data utilisation, pointing to the discrepancies along the four important aspects of how the SLGIS are set up.

[Insert Figure 1] Modelling discrepancies of data utilisation regarding SLGIS

Beyond the discrepancies that exist due to the set-up of the SLGIS, further discrepancies occur when SLGIS are employed in a manner they were not originally designed for. Two processes are identified where one group of stakeholders starts to utilise the SLGIS outcomes in a novel way and the other group of stakeholders follows suit. The first such process is informing students. Initially institutional advisors applied the SLGIS in career
guidance, for example the DLHE was set up to support their work. Due to increased focus on school leaving and graduation, the national policy level took the very same datasets and launched initiatives of informing student choice at an earlier career stage; as explained in this English ministerial interview:

‘[if] people are sufficiently well informed about the importance of filling that schools gap and they know where best to get that education, then they will go there.’ (Ministry 5, EN).

Note that these processes of informing the two groups of students have very different time-frames and implications. When using SLGIS to inform current students on possible careers, the data tell the story of graduates who finished only a few years before them. When showing the same SLGIS outcomes to secondary school students choosing further or higher education, the time-gap is much bigger. By the time this future student moves to the labour market, the reality of employment could be substantially different from what they have been ‘promised’.

A second process where interaction between stakeholders can be observed is evaluation of education. The policy level started to use the SLGIS outcomes to evaluate the sector and the institutional level had to react to this procedure. Along with initiating change to better their SLGIS results, institutions started to gather sociological and contextual evidence especially if their SLGIS data suggested problematic labour market outcomes. An exemplar is how the use of DLHE has changed:

(…) when [the DLHE] started it was very much designed by the careers people as a tool for (…) evaluating their own work, and also to provide them with information for the
next cohort through. Whereas these days, and I think the big switch was round about 2000 to be honest, there is much more of an emphasis on... or there was that stage what became the new emphasis was performance measurement. It suddenly became about measuring the performance of the universities. (Research institute Q, EN)

A further area of interplay between needs of different stakeholders regards the extent to which the SLGIS are used as secondary research evidence by academics and other stakeholders. The comparison of the case studies suggests that if the collectors of the SLGIS are detached from the academic world as statisticians of separate organisations, there is a lower degree of secondary data utilisation by researchers. This could be due to diverging data-needs of stakeholders, fundamental problems deriving from utilising others’ data, insufficient information on the SLGIS, or the pricing of the SLGIS as secondary data.

Further discussion of change in the SLGIS

The long history of the SLGIS suggests that there must be some value in the data, assuming they would have been terminated otherwise. However, it is the long history and the relatively little change in the SLGIS that raises some suspicion that these information systems might be just the ‘usual thing’ to do. England provides an example of higher levels of change in its SLGIS. This could either be due to the changes in the educational and labour market structures the SLGIS gather data about (for discussions see for example Pring et al., 2009). Alternatively, they could be due to the relative ease of change: separate SLGIS for all school levels and sectors allow for modifications more easily than a comprehensive system. Whereas the separate English sectoral datasets are not a part of a broader information scheme, both the Dutch and the Finnish data collections are integrated into overarching structures of educational and labour market analyses, potentially adding further utility value. However,
this might also mean that the Dutch and the Finnish SLGIS are more problematic to alter. As explained regarding the change that occurred in providers of the academic higher education survey in the Netherlands:

There was much debate about the questionnaire that [they] have to change it completely. And then [a research organisation] said: please don’t! Because (...) we cannot make our monitor on student drop-outs. (Research institute Z, NL)

The relatively little change in the analysed SLGIS happened mainly after 2000 in all case studies and regarded the financial and institutional setup, the methodology, and to some extent, the focus of the SLGIS. These could be explained through the changing models of governance, the changing data-needs of governments and institutions, new possibilities opening up due to technological advancement, or the changing dynamics of the role of the state and the educational intuitions. Some of the change suggests similar trends in the analysed nation states: the push for institutional level data for evaluation, as well as the utilisation of more administrative data. This convergence might be explained by the European Union’s growing importance within education or, more specifically, the Bologna-process and its impact on accountability in higher education.

Conclusion

This paper analyses the case of school leavers’ and graduates’ data in three different European countries. It is looking at a specific type of research evidence through the frames of ethnostatistics and enumerology, providing important and useful insight into the issues around producing and using data.
As outlined in this paper, there are a number of discrepancies arising in the production and utilisation of national data between the policy level and institutional level users. First, in terms of data collection, the information systems analysed in this research are either built on sample surveys or the combination of registry datasets. Conducting surveys on the one hand provides tailored information and allows collecting attitudinal data from school leavers’ and graduates’ regarding their education and their current employment. The main issue with such surveys is, however, reaching sufficient sample-sizes for institutional comparisons and that funding for them might be vulnerable and subject to change. Combining different registers on the other hand is an economic utilisation of the already existing data and is thus a relatively cheap account. However, it bears several problems. Beyond data-protection issues and whether the exercise can be carried out within the national legal frameworks, a further issue is to what extent the existing registers are suited to answer questions on school leaving and graduation. Policy-makers and civil servants seem to be pushing for the data collection to be based on census. This, regardless of the research design means reducing the breadth and depth of the information gathered.

Second, in terms of utilising the resulting data, policy-makers seem to be mainly interested in the immediate outcomes of school leavers’ and graduates’, which is also apparent by the tendency of using trend-data rather than longitudinal accounts. One possible explanation for this might be that using trend-data reduces the level of complexity the data-user has to deal with. Longitudinal information seems to be applied in planning national policies, whereas the trend-data are used more in monitoring the system and evaluating educational institutions.
As more accountability measures are based on the SLGIS data, educational institutions have to follow the policy-level’s ‘lead’ regarding the time-frame and the focus of the data they pay attention to. As policy-makers decide on the SLGIS and the mentioned practicalities make it nearly impossible to prompt change, the situation is largely unresolved. It would be crucial for all parties to be more transparent and critical about the way they employ data in their work.

Furthermore, due to the lack of sociological and contextual information about the school setting there is little included in the explanation of outcomes in terms of the background and prior attainment of school leavers’ and graduates’ or the local economy they find jobs at. This means that policy analysts cannot easily see how well the system overcomes initial social stratification. Schools and universities that have more non-traditional or low socioeconomic-status (SES) students, or are situated in economically deprived areas, may appear not to be performing as well as they actually are.

Some of the examples discussed in this paper suggest that the policy level might not be inclined to include others’ information-needs. However, if a data system such as the SLGIS does not provide relevant and useable data for the institutional level, they cannot be expected to respond by changing their procedures to better the outcomes. Similarly, if the data collection is not considered valid evidence by academics, they will not use it as secondary data. This research can provide useful recommendations for national governments committed to setting up SLGIS that works well for a wider set of stakeholders, ‘who want to know what the answers are, even where these are inconvenient for a theory, administration or ideology’ (Gorard, 2006: 19). Without a real commitment to take at least some of those steps, the value to tax-payers on existing or future data systems such as the SLGIS is questionable.
References

Askim, Josten. 2007. "How do politicians use performance information? An analysis of the Norwegian local government experience." *International Review of Administrative Sciences* 73 (3): 453-472.

Bogdan, Robert and Margret Ksander. 1980. "Policy Data as a Social Process: A Qualitative Approach to Quantitative Data." *Human Organisation* 36 (4): 302-309.

Braun, Annette, Meg Maguire, and Stephen J. Ball. 2010. "Policy enactments in the UK secondary school: examining policy, practice and school positioning." *Journal of Education Policy* 45 (4): 547-560.

Bray, Mark, and Thomas R. Murray. 1995. "Levels of Comparison in Educational Studies: Different Insights from Different Literatures and the Value of Multilevel Analyses." *Harvard Educational Review* 65 (3): 472-490.

Bryman, Alan. 2012. *Social Research Methdos*. Oxford: Oxford University Press.

Davies, Huw T.O., Sandra M Nutley, and Peter C. Smith. 2000. "Introducing evidence-based policy and practice in public services." in *What works?: evidence-based policy and practice in public services*, edited by Huw T.O. Davies, Sandra M. Nutley and Peter C. Smith. Bristol: The Policy Press

Davies, Peter. 2012. "Can Governments Improve Higher Education Through ‘Informing Choice’?" *British Journal of Educational Studies* 63 (3): 261-276.
ESDS. 2011. “An ESDS guide: Guide to Longitudinal Study of Young People in England.”

_Economic and Social Data Service_. Available from

http://www.esds.ac.uk/longitudinal/access/lsype/L5545.asp.

ESDS. 2012. “SN 5545 - Longitudinal Study of Young People in England: Waves One to Seven, 2004-2010.” _Economic and Social Data Service_. Available from

http://www.esds.ac.uk/findingData/snDescription.asp?sn=5545.

Estabrooks, Carole A. 1999. "The Conceptual Structure of Research Utilization." _Research in Nursing & Health_ 22 (3): 203–216.

Fitz-Gibbon, Carol 2000. "Education: realising the potential " in _What works?: evidence-based policy and practice in public services_ , edited by Huw T.O. Davies, Sandra M. Nutley and Peter C. Smith. Bristol: The Policy Press.

Gaebel, Michael, Kristina Hauschildt, Kai Mühleck, and Hanne Smidt. 2012. “Tracking Learners’ and Graduates’ Progression Paths (TRACKIT).” _European University Association_. Available from http://www.eua.be/trackit.

Gephart, Robert P. 1988. _Ethnostatistics: Qualitative Foundations for Quantitative Research_, _Qualitative Research Methods, Volume 12_. London: SAGE Publications.

Gephart, Robert P. 2006. "Ethnostatistics and Organizational Research Methodologies: An Introduction." _Organizational Research Methods_ 9 (4): 417-431.

Gillham, Bill 2005. _Research Interviewing: The Range of Techniques_. New York: Open University Press.

Gorard, Stephen. 2006. “Does policy matter in education?” _International Journal of Research & Method in Education_ 29 (1): 5-21.
Gorard, Stephen, and Chris Taylor. 2004. *Combining methods in educational and social research*. London: Open University Press.

GOV.UK. 2013. “Collection - National pupil database (Website).” *Department for Education*. Available from [https://www.gov.uk/government/collections/national-pupil-database#history](https://www.gov.uk/government/collections/national-pupil-database#history).

Hakim, Catherine. 2000. *Research Design: Successful designs for social and economic research*. London: Routledge.

Hantrais, Linda. 2009. *International comparative research: theory, methods and practice*. Basingstoke: Palgrave Macmillan.

Hazelkorn, Ellen. 2013. "How Rankings are Reshaping Higher Education." in *Los Rankings Universitarios, mitos y realidades*, edited by Vicent Climent, Francesc Michavila, and María Ripollès. Anaya: Tecnos

Hordósy, Rita. 2014. “Who knows what school leavers and graduates are doing? Comparing information systems within Europe.” *Comparative Education*, 50 (4): 448-473.

Hordósy, Rita. Forthcoming. “Tracing Europeanisation: school leavers’ and graduates’ information systems as an example.” *Research in Comparative and International Education*.

Kelly, Anthony, and Christopher Downey. 2011. *Using Effectiveness Data for School Improvement: Developing and Utilising Metrics*. London: Routledge.

Kember, David, Doris Y. P. Leung, and K. P. Kwan. 2002. "Does the Use of Student Feedback Questionnaires Improve the Overall Quality of Teaching?" *Assessment & Evaluation in Higher Education* 27 (5): 411-425.
Kerr, Kerri A.; Julie A. Marsh, Gina Schuyler Ikemoto, Hilary Darilek, and Heather Barney. 2006. "Strategies to Promote Data Use for Instructional Improvement: Actions, Outcomes, and Lessons from Three Urban Districts." American Journal of Education 112 (4): 496-520.

Mainguet, Christine. 1999. "The Transition from the Educational System to Working Life. Use of National Statistics". CEDEFOP – European Centre for the Development of Vocational Training. Available from http://www.cedefop.europa.eu/EN/Files/6002_en.pdf.

National Statistics. 2008. Youth Cohort Study & Longitudinal Study of Young People in England: The Activities and Experiences of 16 year olds: England. London: National Statistics.

Newton, Jethro. 2000. "Feeding the Beast or Improving Quality?: Academics' perceptions of quality assurance and quality monitoring." Quality in Higher Education 6 (2): 153-163.

Nutley, Sandra M., Isabel Walter, and Huw T.O. Davies. 2003. "From Knowing to Doing : A Framework for Understanding the Evidence-into-Practice Agenda." Evaluation 9 (2): 125-148.

Ozga, Jennifer. 2009. "Governing education through data in England: from regulation to self-evaluation." Journal of Education Policy 24 (2): 149-162.

Pawson, Ray. 2006. Evidence-based policy: a realist perspective London: Sage Publications.

Phillips, David, and Michele Schweisfurth. 2006. Comparative and international education: an introduction to theory, method and practice. London: Continuum.
Pollitt, Christopher. 2006. "Performance Information for Democracy – The Missing Link?" *Evaluation: The International Journal of Theory, Research and Practice* 12 (1): 38-55.

Pring, Richard, Geoffrey Hayward, Ann Hodgson, Jill Johnson, Ewart Keep, Alis Oancea, Gareth Rees, Ken Spours, and Stephanie Wilde. 2009. *Education for All. The Future of Education and Training for 14-19 Year-Olds*. London: Routledge

Raffe, David, ed. 2001. “Comparative data on Education-to-Work transitions. Report of an International Workshop.” Paris, 21-23 June 2000. Available from [www.mzes.uni-mannheim.de/projekte/catewe/workshop/Report.doc](http://www.mzes.uni-mannheim.de/projekte/catewe/workshop/Report.doc).

Raffe, David. 2008. "The concept of transition system." *Journal of Education and Work* 21 (4): 277-296.

ROA. 2010. “Research Centre for Education and the Labour Market | ROA.” *Maastricht University*. Available from [http://www.roa.unimaas.nl/projects/roaflyer_A4_schoolleavers.pdf](http://www.roa.unimaas.nl/projects/roaflyer_A4_schoolleavers.pdf).

Schildkamp, Kim, Lyset T.M. Rekers-Mombarg, and Truus J. Harms. 2012. "Student group differences in examination results and utilization for policy and school development." *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice* 23 (2): 229-255.

Schildkamp, Kim, and Wilmad Kuiper. 2010. "Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors " *Teaching and Teacher Education* 26 (3): 482–496.

Schildkamp, Kim, and Adrie Visscher. 2010. "The utilisation of a school self-evaluation instrument." *Educational Studies* 36 (4): 371-389.
Shury, Jan, and David Vivian. 2013. “Destinations of Leavers from Higher Education 08/09 survey. Prepared for HESA By IFF Research.” IFF Research. Available from http://www.hesa.ac.uk/includes/C08019_resources/IFF_Research_technical_report.pdf?v=1.2.

Statistics Finland. 2004. “Use of Registers and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland.” Statistics Finland. Available from http://tilastokeskus.fi/tup/julkaisut/kasikirjoja_45_en.pdf.

Statistics Finland. 2012. Overview of the system of education statistics in Statistics Finland (manuscript). Helsinki: Statistics Finland.

Squires, Janet E., Carole A. Estabrooks, Petter Gustavsson, and Lars Wallin. 2011. "Individual determinants of research utilization by nurses: a systematic review update." Implementation Science 6 (1): 1-20.

Turner, David. 2005. "Benchmarking in universities: league tables revisited." Oxford Review of Education 31 (3): 353-371.

United Nations. 2007. “Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics.” United Nations Publications. Available from unstats.un.org/unsd/dnss/docViewer.aspx?docID=2764.

Wayman, Jeffrey C., and Sam Stringfield. 2006. "Data Use for School Improvement: School Practices and Research Perspectives." American Journal of Education 112 (4): 463-468.

Weiss, Carol H. 1988. "Have we Learned Anything New about the Use of Evaluation?" American Journal of Evaluation no. 19 (1):21-34.
Weiss, Carol H. 1979. "The many meanings of research utilisation." *Public Administration Review* 39 (5): 426-431.

Wild, C. Jane. 2010. “Review of Destinations of Leavers from Higher Education (Early DLHE) Survey.” *Higher Education Statistics Agency*. Available from [http://www.hesa.ac.uk/index.php/content/view/1892/128/](http://www.hesa.ac.uk/index.php/content/view/1892/128/).

Young, Viki M. 2006. "Teachers’ Use of Data: Loose Coupling, Agenda Setting, and Team Norms." *American Journal of Education* 112 (4): 521-548.
Notes

1 WO: Wetenschappelijk onderwijs - academic higher education; HBO: Hoger beroepsonderwijs - professional higher education