Review of skills taxonomies

Report prepared for the Skills and Productivity Board

May 2022

Frontier Economics

Danail Popov
Sarah Snelson
Thomas Baily

Disclaimer:

This report represents independent research that was commissioned by the Skills and Productivity Board to inform their analysis approach. The view and opinions expressed in this report are those of the authors, and do not necessarily represent the official views, policy or position of the SPB or any agency of HM Government

Contents

Glossary

ASHE – Annual Survey of Hours and Earnings

CAD – Computer Aided Design

CASCOT – Computer Assisted Structured Coding Tool

DG EMPL – Directorate-General of Employment, Social Affairs and Inclusion

DOT – Dictionary of Occupational Titles

ESCO – European Skills, Competences, Qualifications and Occupations

ESDC – Employment and Social Development Canada

ESF – Essential Skills Framework

ESS – Employer Skills Survey

ICT – Information and Communication Technology

ISCO – International Standard Classification of Occupations

IT – Information Technology

JSR – Job-Skills Repository

LFS – Labour Force Survey

LMI – Labour Market Information

Nesta – National Endowment for Science, Technology and the Arts

NQF – National Qualifications Framework

O*NET – Occupational Information Network

OECD – Organisation for Economic Co-operation and Development

ONS – Office for National Statistics

SFIA – Skills Framework for the Information Age

SFW – Singapore Skills Frameworks

SOC – Standard Occupational Classification

SPB – Skills and Productivity Board

SQF – Structured Query Language

SSA – Sector Subject Area

SST – Singapore Skills Taxonomy

Executive summary

Context

Productivity in the UK has flatlined for the past decade and there continue to be significant regional differences in productivity performance. This is a major concern for policy makers as productivity is a key determinant of living standards. Major government initiatives such as the Industrial Strategy and, more recently, the Plan for Growth[1] have sought to tackle the issue.

It is widely accepted in the theoretical and empirical literature that education and skills are key drivers of productivity and their contribution to productivity growth is significant – previous studies suggest that improvements in skills directly accounted for a fifth of the UK's productivity growth before the financial crisis.[2]

The skills, knowledge and attributes required by the UK labour market are evolving and it is important for the workforce to be well equipped to meet the demands of work now and in the future. Skills gaps and shortages are a feature of the labour market and underlying demand and supply trends can exacerbate current problems. As society ages, the demand for care occupations (where significant vacancy rates are already being reported[3]) is expected to increase further. Other occupations are likely to become obsolete due to automation, while the drive towards net zero is likely to create new occupations or change the skill requirements of existing occupations.

The significant risk of rising unemployment and a mismatch between the skills required in the past and those required in the future make it increasingly important to ensure that training and qualification programmes are well targeted.

The Skills and Productivity Board (SPB) was created to provide independent, evidence-based advice to ministers at the Department for Education on matters relating to skills and their contribution to productivity. In its first year, it is tasked with answering the following three questions:

- Q1: Which areas of the economy face the most significant skills mismatches or present growing areas of skills need?
- Q2: Can the SPB identify the changing skills needs of several priority areas within the economy over the next 5-10 years?
- Q3: How can skills and the skills system promote productivity growth in areas of the country that are poorer performing economically?

The SPB has an immediate interest in using a skills taxonomy – a system for naming, classifying and grouping skills – to better articulate and identify current and likely skills

1 HMT (2021) Build Back Better - our plan for growth (publishing.service.gov.uk)
2 BIS (2015) UK skills and productivity in an international context
3 House of Commons Library (2020) The health and social care workforce gap

gaps across the economy (Q1) to enable government policy to be more effective at targeting them.

Frontier Economics was commissioned to support the SPB in its work to identify how existing skills taxonomies can be harnessed to support its work. Our work has two main objectives:

- To review the alternative skills taxonomies used in different jurisdictions/contexts and assess how well they can support the work of the SPB; and
- To provide recommendations for how taxonomies can be used to address the key questions that the SPB has been tasked to answer.

Approach

Our work primarily involved desk-based research on existing skills taxonomies and the academic and grey literature which draws on these taxonomies. We worked closely with the SPB secretariat and engaged with the SPB Board at key points to discuss emerging findings.

As a starting point, we examined the strengths and weaknesses of the US Occupational Information Network O*NET – a skills taxonomy which provides measures of skills, abilities, work activities, training and job characteristics for almost 1,000 different occupations and which is the main source of occupational competency information in the USA. O*NET is widely used in the USA and abroad and was an obvious starting choice for the SPB given its comprehensiveness (it covers the entire labour market). Furthermore, O*NET is updated regularly and is open source, which makes it easily accessible.

We assessed what (if anything) other taxonomies can offer over and above O*NET and how they perform in key areas such as the inclusion of a more granular layer of technology skills, capturing new or emerging occupations, the ability to link skills requirements to existing qualifications and identifying specific skills in shortage. The alternative taxonomies considered include:

- The European Skills, Competences, Qualifications and Occupations (ESCO);
- The Skills Framework for the Information Age;
- The Singapore Skills Taxonomy; and
- A skills taxonomy developed by Nesta.

Findings

No skills taxonomy is perfect for all purposes. O*NET prioritises analysis of broad skills, with attributes summarised in a smaller number of measures than in other taxonomies. These can be further grouped using higher-level groupings contained within the taxonomy or other divisions such as ‘data-people-things’. The SPB’s work to date has focused on skills at this broad level and our view is that O*NET is suitable for this purpose and remains the most appropriate taxonomy to use.

Taxonomies such as those developed by Nesta and ESCO present skills at a very granular level. If such detail were relevant in future applications of the SPB, the data from these taxonomies could be used to supplement the information available from O*NET. Intermediate-level groupings could also be relevant, which would require analytical effort to develop the right approach to summarising and aggregating the data.

Beyond the taxonomies considered, we thought about how different data sources can be combined with skills taxonomies to add value. For example, the Employer Skills Survey (ESS) or online vacancy data can be used to measure skills shortages in more detail or to build rich datasets that enable productivity returns to be estimated. A key consideration here will be to find ways to align these data sources with existing taxonomies.

There is scope for improvement in other areas of interest too, but not without expending significant resource. Establishing a detailed qualification-skills mapping, for example, would be very valuable but would require significant manual effort, and there is little existing work to draw on. Another area which will require further consideration relates to emerging occupations and skills – while it would be possible to identify these using real-time data, this has limited value in the absence of other labour market indicators.

Recommendations

Our work highlighted some analyses that would enhance the body of knowledge and address the gaps associated with the SPB's use of O*NET in its current form. These include:

1. **Using Nesta or ESCO to derive intermediate-level measures of skill requirements in technology or other areas.** These measures would be more granular than the elements used in O*NET but more aggregate than the individual skills used in Nesta or ESCO. To derive the intermediate-level measures, it would be necessary to aggregate individual skills in Nesta and ESCO in a way that is consistent with the broader categories contained in O*NET.
2. **Doing an approximate mapping between skills in O*NET and qualifications in terms of Sector Subject Area and National Qualifications Framework level.** This approach could identify the types of qualifications prevalent in an occupation and link them to the prominent skills within the array of requirements. This type of analysis would be more challenging in the context of general skills and qualifications, but it could be used to get a sense of the group of skills associated with a qualification.
3. **Monitoring updates to taxonomies to ascertain where new occupations are added.** This would be relevant as part of a wider brief to keep abreast of developments in the analytical sphere. While awareness of newly added occupations or skills is important in understanding the evolving landscape, there is little evidence

that it can be leveraged in relation to these new areas unless primary data collection is also undertaken.

4. **Combining O*NET with data from the ESS and online vacancy data to inform views on shortages within occupations.** This could shed light on the nature of shortages in terms of the specific skills needed rather than in terms of numbers of people. The ESS would provide a broad view of this, while the Nesta data would start at a very granular level and potentially need further analysis to arrive at more general findings.
5. **Considering whether extensions to previous research (such as that conducted by Dickerson & Morris[4]) could provide insights about the returns to other skills.** This would offer a pragmatic route to understanding whether more detailed findings are available regarding returns to skills than the 'data-people-things' paradigm.

The options above have been highlighted based on delivering tangible improvement without involving significant additional resource. Any prioritisation of them would depend on the direction of the SPB's work going forward and the resources available. That said, options 3. and 4. (using ESS) can be considered to involve minimal resource and therefore be relevant in any case. Options 1., 2. and 5. are more resource-intensive and need a clearer justification.

We note that online job vacancy data could also provide powerful granular insights in many of these areas. However, as this would have significant resource requirements and methodological pitfalls, careful judgement would be needed as to whether it was justified.

4 https://cver.lse.ac.uk/textonly/cver/pubs/cverdp020.pdf

1. Introduction

UK skills challenges and the Skills and Productivity Board (SPB)

The UK, like much of the global economy, is facing a range of important structural changes, such as disruptive technologies, an ageing population and environmental challenges, with implications for skills demand and supply. Challenges include lagging productivity growth, skills mismatches, under-investment and gaps in basic skills. In addition, disruptive events such as Covid-19 and the UK's exit from the EU will further affect skills needs as well as the stock of available labour. It is therefore crucial that the education and training system can ensure that the courses and qualifications on offer are aligned to the needs of employers and the economy going forward.

Established by the Department for Education (DFE)[5] in 2020, the SPB's overall aim is to improve the nation's productivity by highlighting longer-term skills challenges and opportunities to address them. The SPB's advice will draw from a wide evidence base including academic research, its own quantitative and qualitative analysis, stakeholder input and employer submissions. The SPB has been asked to prioritise the following questions in the first 12 months:

- Q1: Which areas of the economy face the most significant skills mismatches or present growing areas of skills need?
- Q2: Can the SPB identify the changing skills needs of several priority areas within the economy over the next 5-10 years?
- Q3: How can skills and the skills system promote productivity growth in areas of the country that are poorer performing economically?

Skills taxonomies

A skills taxonomy is a system for naming, classifying and grouping skills. Taxonomies can cover a broad range of human attributes, including skills, knowledge and abilities. They also typically specify the skills requirements of different occupations.

It is unusual to observe directly how skills are used in the economy. Much of our knowledge about changing patterns of skills comes from labour market indicators (LMIs) such as administrative data on unemployment and vacancies or survey data on pay, which together identify the parts of the economy that are growing or contracting. These data are typically

[5] DfE is a UK government department responsible for education and skills in England. Policy is devolved elsewhere in the UK.

presented at the occupation rather than the skill level and are often used to make statements about whether an occupation has shortages (e.g. of nurses, HGV drivers, etc.).

A skills taxonomy can be used to translate data on occupational trends into trends relating to skills. This allows work to focus on the specific skills in shortage or those that will become more important over time. The key to using a taxonomy in this way is that the occupational data in the taxonomy can be linked to UK LMIs via UK Standard Occupational Classification (SOC).

A skills taxonomy can also help to provide meaningful groupings of skills for policy purposes. Taxonomies typically follow a hierarchical structure, enabling skills to be considered at different levels of granularity. The required levels vary by application: for example, a high-level application comparing requirements for communication and programming skills versus a more detailed application exploring specific programming languages. We find that taxonomies vary in terms of the level of granularity at which the analysis is best suited. For example, the O*NET taxonomy is presented at an aggregate level and drilling down into granular detail is more challenging compared to other taxonomies which are at a very granular level and do not easily produce summary aggregate measures.[6]

Skills taxonomies have many potential uses, ranging from the very general to the very specific. These include:

- Practical 'front-line' applications such as career guidance (what job will suit a particular skill set or what skills are needed for a particular job);
- Curriculum and training development;
- Providing an overview of skills mismatches; and
- Understanding historical trends in changing job requirements and projecting future changes.

The desired level of granularity of the taxonomy differs in these use cases. There is potential to use the hierarchical 'tree' structure which brings in these multiple levels of granularity and allows different questions to be tackled, all within the same unified framework. There are also many different dimensions on which skills can be categorised, such as transversal versus occupation-specific levels or conceptual divisions such as cognitive/interpersonal/physical.[7] Which of these dimensions should be prioritised will then determine the most desirable structure of the hierarchy.

---

[6] For example, O*NET defines a 'computer programming' skill in general terms, whereas the European Skills, Competences, Qualifications and Occupations (ESCO) provides a long list of specific programming languages.

[7] Also termed 'data-people-things'.

An ideal taxonomy for the SPB

Skills taxonomies are designed to answer a range of different questions and no existing taxonomy suits all purposes. An ideal taxonomy for the current SPB brief would do the following:

- It would provide a comprehensive overview of the requirements of a job across all dimensions. As a consequence, rather than identifying the most obvious and specific requirements of a job, it would set the broader requirements, providing a more holistic understanding of skills.
- It would allow job requirements to be expressed at different levels of granularity, as the appropriate level is not the same across different policy questions or applications. This would be best met through the levels nesting within each other so that the results were consistent and transparent between the levels. Ideally, it would also reflect more nuanced skill requirements, such as the level at which the specific skill is required and whether the skill required is ubiquitous, optional or related only to a subset of jobs.
- It would link skills to qualifications and other labour market data. This would show what skills a qualification leads to and the corresponding financial returns. More generally, it would capture wider information on the stock of skills, including on-the-job training.
- It would be updated on a regular basis to 1) capture the evolving skills requirements in existing occupations, 2) capture the skills requirements for new occupations and 3) enable users to track the changing importance of skills over time.

Existing taxonomies have strengths and weaknesses and none meets all the ‘ideal’ attributes listed above. Further, building a full bespoke taxonomy afresh which satisfied all the above would involve prohibitive cost and a long-time horizon. Recognising these limitations, this review focuses on how best to make use of existing taxonomies and data sources to meet the current needs of the SPB.

Focus on O*NET and uses of alternative taxonomies

As previously noted, the SPB was created recently and given an ambitious brief for its first year of operation. The tight timeframes and the need to tackle challenging questions using the available data necessitated a pragmatic decision around which taxonomy to use in the immediate term. The O*NET taxonomy[8] was chosen because of its comprehensiveness (it covers the entire labour market), because it is regularly updated

[8] O*NET provides measures of skills, abilities, work activities, training and job characteristics for each occupation (in the US system) and is the main source of occupational competency information in the USA. To construct O*NET, data are gathered from self-reported assessments by workers based on standardised questionnaire surveys in combination with professional assessments by job evaluation analysts.

and easily accessible, and because it provides robust and meaningful descriptors of job requirements. The taxonomy is a well-established and reliable source of information which has taken over two decades to compile. It is widely used in the USA and elsewhere, including by the Organisation for Economic Co-operation and Development (OECD), academics, decision makers and society at large. The significant investment required to produce and update the taxonomy and its established usage contribute substantially to its credibility. O*NET also has the advantage of being available in an accessible form, with prior work to make it compatible with UK occupation-level data.[9]

O*NET is well placed to address some of the key questions around current skills shortages and future skills needs that the SPB has been tasked to answer. The taxonomy has already been used in similar research – for example, the OECD's Skills for Jobs programme[10] uses O*NET in combination with LMIs to derive skills shortage measures.

As well as being referenced in many academic studies,[11] O*NET has been used by the Industrial Strategy Council[12] (in collaboration with McKinsey Global Insight), to project skills requirements up to 2030, a task closely aligned with the SPB's work.

Despite its advantages, it must be recognised that no taxonomy is perfect and O*NET cannot meet all the SPB's requirements. It is therefore appropriate to consider whether other taxonomies might be more suitable as substitutes in the longer term or whether multiple taxonomies can be combined to meet specific requirements.

This report gives an overview of the main taxonomies available. It considers each taxonomy's suitability as an outright substitute for or supplement to O*NET and examines more focused use cases to address the limitations of O*NET in terms of the SPB's prospective work. This leads us to highlight some options for improving the evidence base.

The remainder of this report is arranged as follows:

- Chapter 2 provides a description of the main taxonomies;
- Chapter 3 first sets out the rationale for using the O*NET taxonomy, followed by a discussion of its limitations and an assessment of other taxonomies to supplement it;

---

9 In particular, the mapping to UK SOC and possible linkage to labour market indicators make it very appealing.

10 https://www.oecdskillsforjobsdatabase.org/#FR/_

11 See for example: Autor, David H and Michael J Handel (2013), 'Putting Tasks to the Test: Human Capital, Job Tasks, and Wages', Journal of Labor Economics, 31(2), S59-S97 or Abraham, Katherine G and James R Spletzer (2009), 'New Evidence on the Returns to Job Skills', American Economic Review Papers and Proceedings, 99(2), 52-57.

12 https://industrialstrategycouncil.org/uk-skills-mismatch-2030-research-paper

- Chapter 4 sets out a number of ‘use cases’ for other taxonomies, where we consider their ability to address the limitations of O*NET; and
- Chapter 5 gives a summary of the overall findings and makes some recommendations for using skills taxonomies in the future.

2. Assessment of skills taxonomies

This chapter describes the various taxonomies currently available and considers how they could be used to meet the SPB's needs. A brief overview of the taxonomies is provided below (with a fuller summary provided in Annex A). This is followed by a discussion of O*NET and consideration of each of the other taxonomies.

Overview of skills taxonomies

O*NET is a US-based system which measures, for each occupation, the importance of different skills and abilities. This is done through a combination of the input of job evaluation experts and surveys of job incumbents. The system measures job requirements in terms of 177 different elements, covering around 1,000 occupations. O*NET was first published in 1998 and is well established in research uses.

The European Skills, Competences, Qualifications and Occupations (ESCO) is a European Commission project, first published in 2017. ESCO has separate 'pillars' for categorising and linking occupations, skills and qualifications. Like O*NET, it draws on job evaluation expert input. However, a key difference is that skills are measured at a very high level of granularity, with around 13,500 distinct skills appearing in a multi-level hierarchy. In practical terms, this lends it a granular emphasis.

Nesta skills taxonomy is derived using 'graph clustering' analysis of online job adverts in the UK, with skills that appear in the same adverts being placed in the same cluster. Like ESCO, the Nesta taxonomy has many skills (10,500), which are organised in a multi-level hierarchy, also giving this a granular emphasis.

The Skills Framework for the Information Age (SFIA) is a competency framework which describes the skills requirements of digital occupations. It is periodically updated by sector experts but only covers digital skills. A detailed mapping for Saudi Arabia has been developed, which links skills requirements to occupation codes.

The Singapore Skills Taxonomy (SST)[13] is derived using neural network analysis of skills framework documentation. This is similar to the Nesta UK taxonomy in structure but has the advantage of not relying on vacancy data, thus avoiding concerns around representativeness or completeness.

Canada Skills and Competencies Taxonomy is a system under development which draws on both O*NET and preceding national skills frameworks, illustrating a hybrid approach. The detailed data have not yet been released.

[13] Note that, while SST and Employment and Social Development Canada (ESDC) are discussed in the summary of skills taxonomies (see Annex A), they do not have available data that can be readily incorporated into UK occupational classifications and, therefore, are not considered further in this assessment.

Benefits and limitations of O*NET

The SPB has been tasked with exploring a range of skills questions and, depending on the question, which different skills taxonomies or data sources may be the most suitable. In relation to this, O*NET has several key advantages which make its use a pragmatic choice for the SPB in answering its first-year questions. However, other users or uses may require a different or additional taxonomy. It is therefore appropriate to set out both the benefits and limitations of O*NET, to help with understanding where alternatives may be of use.

The key advantages of O*NET are as follows:

- *Comprehensive coverage*. O*NET provides detail for most occupations[14] and measures attributes that capture a broad range of job requirements, including technical skills, physical and cognitive abilities and personal qualities. It also describes the contextual features of a job.
- *Input from job evaluation experts*. This is the most objective and robust measure of job requirements, and the taxonomy reflects considerable input as it is populated by job evaluation experts.
- *Ease of practical application.* The O*NET occupation codes have already been mapped to UK SOC codes and data are available in a format which makes it straightforward to combine them with UK labour market information.[15]
- *Level of granularity*. O*NET skills requirements are presented at a relatively aggregate level, which is close to what is required for much of the SPB's current work.

Given these practical benefits, O*NET provides a suitable taxonomy for assessing skills shortages or future jobs at a broad level. However, this does not mean it is necessarily equipped to deal with other questions that may be relevant. The taxonomy has some general features which would ideally be improved:

- *Applicability to the UK.* O*NET is a US-based system designed to measure skills requirements in the US job market. In theory, UK and US job roles may differ, reflecting technology usage, commercial patterns, industrial profile, legislation, regulation and the economy. At a practical level, the O*NET-SOC occupational hierarchy does not map directly to UK SOC, and some granularity is lost, with some occupations lacking a match or relying heavily on averaging.[16]

---

14 O*NET does not evaluate military occupations, and newly added occupations can take time for detail to be collated.

15 For example https://www.lmiforall.org.uk/

16 For example, writers and translators are in separate O*NET occupation codes, with very different ratings for foreign language importance. However, they appear in the same UK SOC code, with an averaged foreign language score that is appropriate for neither group.

- *Timeliness.* The input from job evaluation experts and surveys is resource-intensive, so a pragmatic decision is made around the trade-off between update frequency, data quality and sample size. Consequently, the data in O*NET can be up to five years old. This may be a limitation for very fast-evolving jobs but is unlikely to be a material consideration as larger occupations are updated more regularly than smaller ones to limit impact on the whole labour market view. It should also be noted that O*NET incorporates modern forms of data collection, such as natural language processing and machine learning, and uses web scraping to maintain the Tools and Technology module.[17]

The more fundamental aspects of O*NET which may limit its usefulness in addressing specific policy areas of interest include:

- *Technology skills.* O*NET is limited in this regard as it only includes a very general 'computers and electronics' knowledge element and 'interacting with computers' work activity, coupled with an exhaustive list of the software used in an occupation, but without any reflection of the intensity of use. There are also technology components which appear in large numbers of task ratings, but these are not comparable across occupations or easy to identify. More generally, O*NET lacks granular detail on skills which may be relevant in more detailed applications.
- *Difficulty mapping to UK qualifications.* O*NET does not provide any link between skills and UK qualifications. Ideally, a taxonomy would show what skills a qualification brings, thus linking skills to UK training output. However, the 'level anchors' within O*NET are not readily comparable across elements, and they do not correspond to UK qualification levels. While O*NET also contains an overall 'job zone' field describing whether a college degree, university degree, etc. is required, it has no direct link to the associated skill.[18]
- *Ability to incorporate new occupations, skills and technologies.* O*NET's process for adding new occupations involves experts in sectors of interest submitting proposed additions to the US Department of Labor, with approval depending on the job being sufficiently different from others. This can result in delays in identifying relevant new roles and skillsets. While there is a 'bright outlook' flag which describes the growth prospects of an occupation, this reflects US labour market analysis of established occupations rather than emerging occupations. O*NET also contains a 'hot technologies' flag, but this merely states the top 200 most prevalent software packages identified at a previous point in time and is neither forward-looking nor growth-based.

[17] For more detail on these developments, refer to https://www.onetcenter.org/reports/EmergingTasksNLP.html

[18] This is because it is not always the case that the highest-rated skill in an occupation will be the degree subject.

- *Capturing specific skills in shortage.* While O*NET, in combination with other sources, can be used to explore which occupations are in shortage, and by construction the associated skills, it cannot show what the skills in shortage are within an occupation. For example, the labour market data may indicate a general shortage of plumbers but will not indicate whether there is a particular skill in shortage such as fitting pipes or maintaining boilers. It therefore will not shed light on skills requirements changing at this level (e.g. boiler maintenance becoming more important over time).

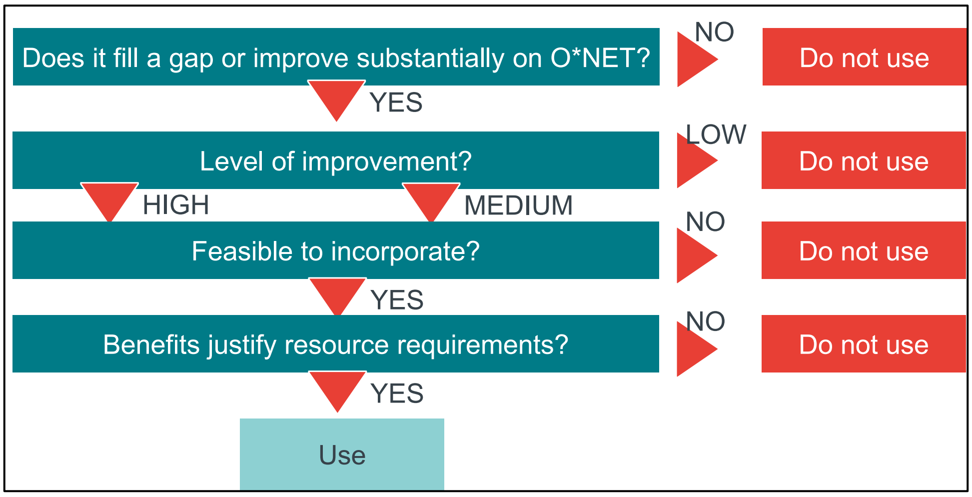
Consideration of alternative taxonomies

While we established that O*NET is a good option for assessing the broad-based skills questions considered thus far by the SPB, we also want to understand whether other taxonomies may be suitable for use in the SPB's work going forward either as an outright replacement for O*NET or used in conjunction with O*NET.

We take a sequential approach to this by considering whether a given taxonomy can address any of the fundamental gaps discussed in the previous section and/or offers a material level of improvement compared to O*NET. We also consider how feasible it is to incorporate alternative taxonomies into existing work and, finally, consider possible improvements against their resource cost.

The logic of this approach is that it can identify clear improvements incremental to O*NET rather than weighing up strengths and weaknesses across a broad range of attributes which leave a mixed picture. The diagram below represents this logic, which is applied in respect of each of the O*NET limitations discussed above.

Figure 1 - Criteria for assessing other taxonomies



In the following sections, we consider whether alternative taxonomies could offer an outright substitute for O*NET in the context of the work of the SPB. We then consider the criteria above for each taxonomy and for each of the following four questions:

- Can the taxonomy provide a more detailed picture of technology skills to replace or complement those in O*NET?
- Does the taxonomy enable a mapping of skills to UK qualifications?
- Does the taxonomy offer the potential to incorporate new occupations, skills and technologies?
- Does the taxonomy do a better job than is possible with O*NET in capturing *specific* skills in shortage?

Our high-level assessment of the most relevant taxonomies is shown in **Figure 2**. The general finding from this assessment is that there are specific areas such as ‘technology skills’ and ‘specific skills shortages’ where the use of other taxonomies can help the SPB tackle its brief more effectively. We provide further detail on this in the sections that follow.

Figure 2 - Summary of Assessment

		Overall substitute for O*NET	Technology skills	Mapping skills to qualifications	Incorporating new/emerging occupations	Specific skills in shortage
ESCO	Level of improvement	No – possible complementarities but requires significant investment.	Medium	Low	Low	Low
	Feasibility		Yes	N/a	N/a	N/a
	Resource		Medium	N/a	N/a	N/a
Nesta	Level of improvement	No – but there are complementarities.	Medium	Low	Low	High
	Feasibility		Yes	N/a	N/a	Yes
	Resource		Low	N/a	N/a	Medium
SFIA	Level of improvement	No – it focuses on ICT skills only.	Low	N/a	N/a	N/a
	Feasibility		N/a	N/a	N/a	N/a
	Resource		N/a	N/a	N/a	N/a

Source: Frontier Economics

ESCO

Our overall assessment is that ESCO is not a suitable substitute for O*NET for the purposes of SPB. Using it could add value in the area of technology skills but this would have significant resource implications.

		Overall substitute for O*NET	Technology skills	Mapping skills to qualifications	Incorporating new/emerging occupations	Specific skills in shortage
ESCO	Level of improvement	No – possible complementarities but requires significant investment.	Medium	Low	Low	Low
	Feasibility		Yes	N/a	N/a	N/a
	Resource		Medium	N/a	N/a	N/a

ESCO presents skills at a more granular level than O*NET, so the key question is whether a high-level or granular system is most desirable. While a granular system may be suitable for front-line careers advice or CV uses, the SPB's focus is likely to be on more general uses where this level of detail is less important.

Presentation of ESCO at a more aggregate level would require further consolidation and reduction to summarise skills at a more aggregate level, which would require a further developed methodology. The resulting descriptors at this level, e.g. counts of skills in each category required for an occupation, may be less valuable than the intensity and level indicators reported in O*NET. In this regard, O*NET offers a more suitable framework.

ESCO provides a facility to interface with many aspects of labour market information, which would potentially provide a powerful integrated tool concerning job mobility or online CVs. However, for this value to be realised it would require UK participation in the relevant schemes to generate the relevant data. These data would need to be mapped to UK standards. The costs of adopting ESCO in a way that would fully leverage its benefits are considerable, and this is well outside the scope of the SPB's remit. Therefore, while this aspect of ESCO could be considered an area of relative strength, it is not immediately relevant for the purposes of the SPB.

Technology skills

ESCO includes extensive detail on the specific technology skills required for each occupation, a key gap in O*NET.[19] For example, in ESCO, an information and communication technology (ICT) network administrator would be identified as needing to know ICT security legislation.

However, ESCO is set up to flag many potentially relevant skills as being essential or optional for an occupation. Cross-cutting or general technology skills may not be considered sufficiently essential for an occupation and may not be captured by ESCO.[20] In contrast, O*NET is designed to give precise numerical scores for a much smaller and consistent set of skills.

It would be possible to supplement the O*NET occupation-skill profiles by linking ESCO occupation codes to UK SOC, thus providing a list of technology skills required for each UK SOC code. Combining ESCO data with O*NET data needs technical knowledge of how to extract the ESCO data, manual identification of relevant skills within ESCO and development of a methodology for aggregating technology skills to broader groups.

Link to UK qualifications

To date, there has been little mapping of the qualifications pillar within ESCO to skills or occupations, and the procedure has been considered too resource-intensive to carry out manually, with concerns around validity when using an automated approach. If such mappings were to become available, they would still be of limited use as the European qualifications would need to be mapped to their UK equivalents, which could be as resource-intensive as and less accurate than doing the mapping directly. There is also a question about how useful the European qualification stocks data would be, considering their completeness and their applicability to the UK.

New and emerging occupations

ESCO is continuously updated. A recent article[21] describes work by the ESCO data science team using artificial intelligence to automatically maintain the occupation pillar and facilitate cluster analysis using vacancy and qualification metadata. More generally, the update process draws on LMIs and changes to national occupational classifications and involves consultation with sector experts, member states and the European Commission. It is not obvious that this is any quicker than O*NET and, again, new

[19] Technology skills are updated with the rest of the content in ESCO, drawing on a combination of desk research, targeted feedback from sector experts and expert groups. Newly added content includes 'e-learning architect' (verifying online delivery of learning), and 'develop food scanner devices'.

[20] For example, no technology skills are listed in relation to secondary school history teachers, although some technical capability would be expected.

[21] https://ec.europa.eu/esco/portal/news/e532488b-e1f1-4e32-a1b8-ff3f4689f599

occupation data would be of little use without accompanying data prepared on an equivalent occupational basis.

Specific skills in demand/shortage

ESCO only measures job requirements and therefore does not capture any direct information about which skills are in demand within an occupation or are driving any shortage.

Nesta

Our overall assessment is that the taxonomy developed by Nesta could add significant value in two areas: technology skills and specific skills in shortage. The resource implications of using this taxonomy in addition to O*NET are relatively low.

		Overall substitute for O*NET	Technology skills	Mapping skills to qualifications	Incorporating new/emerging occupations	Specific skills in shortage
Nesta	Level of improvement	No – but there are complementarities.	Medium	Low	Low	High
	Feasibility		Yes	N/a	N/a	Yes
	Resource		Low	N/a	N/a	Medium

The Nesta taxonomy, derived from online vacancy data, has several advantages in terms of ease of periodic updating without a need for extensive manual expert input and its direct UK focus. There are also plans to integrate it with ESCO. Regardless of whether these developments come to fruition, O*NET will continue to have strength in focusing on broad skills (and measuring them on a continuous 7- or 5-point Likert scale[22]), whereas the Nesta hierarchy places more emphasis on the technical skills used in different occupations. The taxonomies are therefore geared to answering different questions. The focus on occupational skills and the exclusion of some transversal skills in the Nesta taxonomy can also be considered problematic given the central role of these skills in facilitating movement between jobs.

A more general issue around using online vacancy data is representativeness, as online adverts may not be common in all occupations and the skills profile for advertised roles may be different to the average skills requirement. It is also worth noting that many jobs are filled internally by word of mouth and are therefore never advertised. As a result, jobs that appear as online vacancies may be difficult to fill through internal or informal channels and differ in important respects.

[22] https://www.britannica.com/topic/Likert-Scale

Technology skills

A wide range of detailed technology skills are available in the Nesta taxonomy. As they are mapped to occupation, this can provide a skills intensity matrix at varying levels of granularity. At the most granular level, for example, 45% of data scientist job adverts require Python and 25% of job adverts require SQL. At a higher level of aggregation, 85% of adverts require data engineering.

Again, the issue of salience is relevant as vacancies may not mention skills that are relevant but are either assumed to be present, and therefore not explicitly requested, or are not at the front of the mind of the recruiter. This could lead to understatement of the importance of general skills or technology skills in occupations where they are not the immediate focus. It could also understate skills that are a prerequisite to such an extent that they are not specified in the job advert. As a result, the improvement offered by Nesta is focused on salient technology skills, with gaps remaining in relation to more general technology skills that are less likely to be mentioned in job adverts.

It would be feasible to incorporate data on salient technology skills into work using O*NET on the assumption that the forthcoming update will provide detailed tabular data linked to SOC codes, thus enabling direct use of the occupation-level requirements. The resource requirement associated with this would be low, assuming that the flat file data are available in the forthcoming update.

Link to qualifications

The Nesta taxonomy does not include qualifications, and there are currently no plans to add them to the taxonomy.[23]

New and emerging occupations

For the Nesta taxonomy to complement O*NET in this area, periodic updating and maintenance of the Nesta work, drawing on new online vacancy data, would be required. Finding new occupations empirically will have a time lag because of the time it takes for new occupations to accumulate enough data points to emerge. This is not necessarily quicker than the expert-led process used in O*NET. In addition, common to any route for exploring new occupations, analysis is limited by the lack of wider labour market information at the relevant occupational level, as the relevant SOC codes would not yet exist. This would therefore be of limited use without novel real-time data linked to job titles, which would be costly to create and thin in coverage.

[23] Although online job adverts allow for a mapping between skills and qualifications; this is an involved task with methodological and conceptual challenges. The feasibility of this is discussed in the following chapter.

Specific skills in demand/shortage

This may be an area of strength for the Nesta taxonomy compared to others as online vacancy data can reveal the specific skills demanded by employers and give some indication of what is driving any occupational shortage. This would show in real time the evolving skills requirements – for example, if adverts for software engineers emphasise specific types of software skills or adverts for nurses emphasise certain aspects of care.

However, salience bias is a potential limitation: a high frequency of mentions could reflect an acute shortage of the skill or could reflect it being a popular descriptive term in job adverts. Similarly, a skill may be considered a prerequisite and be assumed as a given to such an extent that it is not explicitly mentioned in a job advert.

Another limitation of a taxonomy based on online vacancy data is that vacancies do not necessarily result in external recruitment. Employees may be recruited and trained internally to meet demand, which is likely to be the case if the external labour market is perceived as unlikely to contain the desired skills. For example, a growth in demand for online teaching skills could result in a drive for the internal training of existing teachers without a sharp redeployment in terms of hiring decisions. Another potential issue is that the likelihood of jobs being advertised online varies by a number of factors including sector and skill level, which could lead to exclusion bias.

Noting the limitations set out above, it is feasible to combine information on specific skills shortages with O*NET. Linking the skills profiles to occupations should not be difficult, but significant effort is likely to be required in interpreting the results given the above limitations.

SFIA

As the focus of the SFIA is exclusively on ICT skills, we confine the assessment to this area. The SFIA is a good example of open and expert-led frameworks being developed autonomously. SFIA provides a rich set of descriptors of technology skills which, in theory, would supplement O*NET well. The relatively frequent update cycle and the fact that it is expert-curated are also advantageous. On the whole, however, given its very narrow focus, it offers the SPB relatively little improvement over and above what is available in O*NET, given its current brief.

		Overall substitute for O*NET	Technology skills	Mapping skills to qualifications	Incorporating new/emerging occupations	Specific skills in shortage
SFIA	Level of improvement	No – it focuses on ICT skills only.	Low	N/a	N/a	N/a
	Feasibility		N/a	N/a	N/a	N/a
	Resource		N/a	N/a	N/a	N/a

3. Use cases for extending analysis

The discussion of O*NET in the previous sections highlights some key limitations in the context of the SPB's future work. There may be useful ways in which O*NET can be supplemented to help the SPB address its brief.

Some of the SPB questions can be addressed through a combination of alternative taxonomies or data sources. In the simplest case, this can be done by adding variables that describe an occupation's skill requirement in greater detail than in O*NET.

In each of the use cases which we consider below, we describe the analytical question and summarise the difficulty associated with using O*NET in this context. We then assess how different datasets in turn could be used to address this gap, paying close attention to the analytical and resource requirements of doing so. We note that the use cases described here are not intended to be exhaustive, and we do not anticipate what work the SPB may undertake in the future.

Technology skills

As previously noted, one of the strengths of O*NET is that it is well suited to analysing skills at a broad level. However, in some cases, the taxonomy may not go into enough detail to answer pertinent policy questions. This is a general feature which could affect many areas of skills, but here we focus on technology skills because of their wide usage and increasing importance across many different occupations and sectors.[24]

O*NET provides importance ratings for a 'computers and electronics' knowledge element, a technical skills 'programming' element and an 'interacting with computers' work activity element. O*NET also provides information on the specific types of software used in occupations; although the list is relatively exhaustive, it does not reflect the of use or level of ability required in the application of these programmes. It would be useful to have a more tangible and intermediate layer of technology skills showing what types (and intensities) of computer skills are needed in different jobs.

In what follows, we describe four possible options for providing this tangible intermediate layer of detail on technology skills and then compare their strengths and weaknesses to reach a conclusion about which option looks most promising.

Option 1: use further detail available in O*NET task ratings

Around 18,000 different tasks are specified in O*NET. The task ratings are occupation-specific but are then aggregated to derive skill ratings. Some tasks include some technology aspects from which relevant insights can be drawn. Tasks are grouped into

24 Similar observations apply when considering granularity in other contexts.

intermediate layers of ‘intermediate work activity’ and ‘detailed work activity’, and ultimately into the ‘work activity’ elements presented in the main indicators.

A manual classification could be developed to identify and group technology-related skills and derive a more detailed classification than is currently available . This would involve identifying which groups of tasks in O*NET are technology related and then summarising these in the required intermediate-level measures. Work of this nature has been undertaken by McKinsey[25] to group tasks into a ‘workplace skills’ classification that is based on O*NET task ratings.

The advantage of this approach is that task ratings are consistent with the wider O*NET framework and cover all occupations. However, task ratings are not complete and do not capture many appropriate skills such as data visualisation or skills related to specific types of software. The level of resource required to implement this approach would be high. The approach requires flagging and classifying tasks and forming relevant groupings from the bottom up.

Option 2: supplement with a more granular layer from ESCO

ESCO has a detailed list of skill requirements covering all occupations, and these can be grouped in different levels of hierarchy. ESCO has more detail on digital skills than O*NET, with a whole branch of the knowledge tree focused on ICT and a branch of the skill tree focused on working with computers. Subject to extracting the data, these can be mapped easily to UK SOC codes. The level of resource required would be medium. The data would need to be aggregated to some intermediate level of granularity, which would require rules around how these summary statistics are measured and standardised, e.g. the count of skills specified within a group.

As noted earlier, ESCO is less good at capturing the skills required at lower levels of intensity. The binary classification it uses does not capture skills with a level of requirement below a given threshold. For example, ESCO may say that an architect needs to know computer aided design (CAD) but not that a restaurant manager needs to be proficient in using point-of-sale technology.

Option 3: supplement with a more granular layer from Nesta

In practical terms, this would be similar to supplementing O*NET with ESCO data, as it would involve merging additional data with occupation codes and summarising them at an intermediate level of detail based on groupings provided in the taxonomy.

There are, however, important conceptual differences. Nesta has the advantage of continuous measures of skill, and this may better capture skills intensity in occupations where the technology skill requirement is lower. For example, in the Nesta taxonomy,

[25] https://industrialstrategycouncil.org/uk-skills-mismatch-2030-research-paper

situations where 10% of adverts require a skill and one where 2% of adverts require it can be distinguished from each other. In contrast, some less-relevant skills may not appear at all in the binary classification used by ESCO, and where they do appear it is not possible to understand the intensity with which they are required. On the other hand, online job advert data have limitations in terms of representativeness (not all types of jobs are as well covered) and salience (not all relevant or important skills are listed in a job advert).

Option 4: supplement with a more granular layer from SFIA

It would be possible to use the occupation-skills mapping developed for Saudi Arabia and link it to the UK SOC code. This would give objective (expert-curated) job requirements as set out by sector professionals for a range of ICT occupations. However, SFIA only covers core digital occupations, which significantly limits its use for the SPB. The level of resource required to perform this integration would be medium, requiring textual analysis to link the occupational fields.

Summary and recommendation

If the SPB has a requirement to consider ICT occupations in isolation, SFIA would provide the most objective and systematic assessment, based on expert industry judgement.

More generally, both ESCO and Nesta would be viable options to bring in additional detail on technology skills. In both cases, this would involve aggregating the data to the required level of detail and then linking the output to O*NET using the UK SOC code.

If the Nesta taxonomy were published in sufficient detail (i.e. with underlying data tables), this might be the most pragmatic option for the SPB. The Nesta taxonomy has the advantage of splitting technology skills into clear groups, whereas ESCO has multiple categories to consider within different parts of the taxonomy, requiring more manual effort to identify the relevant parts. If concerns around representativeness can be assuaged, the taxonomy developed by Nesta would represent an attractive option as it provides continuous rather than binary measures.

Qualifications

The mapping between qualifications and skills is important as it would allow the output of training providers to be understood in skills terms and show how skills needs might be met by the education and training system. It would also enable a greater understanding of the stock of skills, as evidenced by the stock of qualifications. However, it is important to acknowledge that not all skills are provided through formal (qualification-focused) training, so evidence on wider skills formation would remain desirable.

Ideally, we would like to know which qualifications deliver which skills, i.e. the causal contribution of training. In practice, however, what is observable is only the co-occurrence of skills and qualifications. For example, a doctor is required to have good communication skills but these may be accumulated as part of on-the-job experience and prior education rather than arising specifically from studying medicine.

In practical terms, establishing links to qualifications is difficult as none of the taxonomies provide any mapping to UK qualifications. While ESCO has a qualifications pillar which, in principle, would provide a mapping to skills that could be leveraged in a UK context, at present there is very little mapping to draw on. We consider two possible options for integrating qualifications information into the SPB's work.

Option 1: top-down assignment using O*NET

UK qualifications are defined in terms of levels using the National Qualifications Framework (NQF), which is aligned with the European Qualifications Framework. Qualifications are also assigned to a Sector Subject Area (SSA). This is a 2-level hierarchy, with the upper level containing 15 categories, below which sit 50 sub-groups. For example, 1 is 'Health, Public Services and Care', below which sits 1.1. 'Medicine and Dentistry'. Information on qualification level and subject is captured in this form in the Labour Force Survey. This information gives an indication of the stock of qualification types by occupation.[26]

The qualification profile from the Labour Force Survey (LFS) can be linked at the occupation level to the skills profile in O*NET to show which 'bundles' of skills and qualifications are found together. A relatively manual 'rule-of-thumb' approach would focus on obvious correspondences between qualification subjects and skills, for example to find that medicine is the main skill of a doctor, as well as the subject of their highest qualification. It is likely that this would find some very strong linkages in areas where there is strong progression from gaining the qualification to working in the corresponding occupation. However, in many cases, it will be difficult to find such a link – if, for example, a qualification leads to many possible destinations or if the LFS-qualification and O*NET skills definitions in combination do not describe the correspondence well. This would leave many qualifications without a distinct skills profile.

A more ambitious empirical approach would involve estimating the skills profile as a function of the qualification mix, i.e. measuring which types of qualification are associated with uplifts in certain types of skill.[27] With an exhaustive empirical approach such as this, further care would be needed in interpreting the results, particularly in forming causal

[26] Note that this is in terms of qualifications held rather than the actual requirement. Employees' qualification stock will not always align with requirements, for example due to holding superfluous qualifications that are not required in the current role.

[27] For example, this could be a series of regressions with a skill as the dependent variable and various qualifications as the explanatory variables. This approach might require data reduction to avoid spurious results.

hypotheses. Again, the granularity of any mapping would depend on how well the LFS and O*NET profiles combine.

Option 2: explore co-occurrence of skills and qualifications in online vacancy data

This would involve analysing online vacancy microdata and seeking to identify which skills occur in conjunction with a qualification. For example, a project management qualification may appear in job adverts which specify a mix of resource management skills, from which it could potentially be inferred that these skills are associated with the qualification. This would give a skills-qualification matrix which shows, conditional on a qualification being held, the average requirements for different skills. The challenge would then be to attribute the increased (or reduced) requirements for particular skills to the qualification under consideration. This largely follows the empirical approach discussed above, but with the difference that the online vacancy data would be more granular than the LFS in terms of describing qualifications. While this may be beneficial, it would also present difficulties in terms of having many more qualifications to consider and scope for 'overfitted' regressions to produce spurious results. Again, causal interpretation of the link between qualifications and skills would be difficult.

The benefit of this approach is that it could give nuanced and granular insight and is flexible in allowing skills to be categorised at different levels of detail. However, there are questions about the validity of attributing the co-occurrence of a skill with a qualification to the qualification and, more generally, around representativeness when using vacancy data (discussed previously). It should also be noted that qualification requirements are not complete, again for salience reasons. This is particularly likely for lower-level qualifications, which may not gain much traction under this approach. In addition, the resource cost would be high due to the cost of purchasing data, undertaking primary processing, developing a robust empirical methodology and structuring the outputs.

Recommendation

Overall, the most pragmatic approach would be to link UK qualifications to O*NET elements and level ratings using LFS data, which is made straightforward by the limited number of categories under consideration. This would give a high-level mapping which would be adequate for many purposes. Although it would not tackle causal attribution, it would show the skills mix associated with different types of qualification.

Including new and emerging occupations

It was noted earlier that O*NET suffers some lags in terms of adding new occupations, relying on sector specialists to identify them followed by an approval process with the US Department of Labor and then time lags for the detailed profiles to be built up using job evaluation experts or surveys of incumbents. For example, although 'data scientist' now appears as an occupation in the taxonomy, it does not yet appear in the latest release in terms of skills ratings. We therefore considered whether other taxonomies might provide better evidence.

An overarching consideration, however, is that even if taxonomies or data sources can identify new occupations, there is very little more that can be done with this insight in terms of identifying skills mismatches as other labour market data will not capture them at an occupational level. So this use case is restricted to considering the ability of taxonomies to identify new occupations rather than providing much additional data.

Option 1: ESCO

The update cycle draws on ESCO's data science team, LMIs, and national vacancy and qualification metadata, with clearance involving the approval of member states. For comparison with O*NET, ESCO does contain a detailed skills profile for a data scientist. However, comparison of the occupation codes in version 1.0 (2017) and version 1.0.8 (2020) finds no difference in occupation codes. It would be advisable to check version 1.1 (the next material update, in development) for any new occupations. Checking for new occupations in ESCO by comparisons between versions is also feasible.

Option 2: Nesta

An updated version of the Nesta taxonomy is due to be published in the autumn of 2021. It is expected to include data published at the occupation level (SOC code) rather than job title level. If so, any data at the SOC code level would not give further insight into new occupations as it is not possible to distinguish the old and new occupations within the code. This option is therefore not considered further at this point.

Option 3: Online job vacancy data

Job vacancy platforms (or scraped data) can provide job titles in free text. This could be used to manually identify new roles that are being advertised. This information can be selected to specifically reflect the UK context. However, this is a significant task requiring a developed analytical procedure, acquisition and manipulation of raw data, and validation of findings (i.e. that new job titles are sufficiently distinct to warrant being defined as separate occupations). The fact that established taxonomies take some time to add new occupations and involve significant industry expertise reflects the complexity of the task and resource required to generate robust insights. This approach may not give results any sooner than other taxonomies and, therefore, there would seem to be little value in attempting this task independently.

Recommendation

There is limited value to be gained from analysing new occupations within the existing taxonomies in the context of the SPB's work. Analysis using O*NET requires complementary information from other sources to be available in order to address the questions the SPB is seeking to answer. This means that until new occupations are also reflected in other sources of labour market information there is little to be gained from capturing them within the skills taxonomy.

Despite this, insight could be provided by following emerging occupations. A pragmatic solution may be to periodically review updates to existing taxonomies to identify any changes. The forthcoming ESCO v1.1 therefore warrants attention.

Skills in shortage within an occupation

Moving somewhat beyond skills taxonomies and the overall skills shortage approach, we consider the wider issue: taxonomies such as ESCO and O*NET set out the generic requirements of a job but they do not shed any light on how the supply and demand of skills in an occupation compare, i.e. if there are any skills mismatches *within* that occupation. Existing analysis of skills mismatches (OECD Skills for Jobs and SPB methodology) leverages labour market mismatches at an occupational level and then uses the job requirements to describe this in terms of skills mismatch. This is reasonable as the majority of labour market information is defined in terms of occupation. However, it does not identify the actual skills that are driving any shortage. For example, the occupation-shortage approach would state that there are not enough plumbers and, by extension, there is a shortage of building and technical skills (skills that are rated as being important for plumbers).[28] However, while there may be a surplus in terms of people working in the occupation, there may be an acute lack of particular skills, which means there is still a shortage (e.g. related to smart heating systems). It is therefore instructive to consider how the evidence in this area can be improved. We consider three potential options.

Option 1: high-level analysis using Employer Skills Survey (ESS)

The ESS provides evidence on what skills are in shortage within an occupation. It includes information on areas such as the skills lacking in the workforce or skills in hard-to-fill vacancies. It is therefore tailored and highly relevant to this question.

The ESS is defined at a very aggregate level. Skills are described in varying levels of detail, containing up to 25 fields (e.g. ‘Managing or motivating other staff’, ‘Advanced or specialist IT skills’) and some higher-level groupings. In many cases, these align relatively well with elements in O*NET and would therefore allow some understanding of skills shortages situated within a systematic framework for understanding job requirements. However, an important limitation is that the occupations are very crudely defined, i.e. only at the 1-digit SOC level. While it is possible that the microdata could be

[28] Or to take an example in the opposite direction, there may be ample supplies of a skill, but it is underutilised, as those with the skill work in other occupations where pay or conditions are better.

worked further (in combination with Standard Industrial Classification (SIC) codes), the sample sizes are unlikely to enable much more robust insight at any level of granularity.

Nevertheless, a more systematic application of ESS could identify some common patterns affecting groups of occupations. It could also reveal more common trends in skills shortages, although there could be pitfalls in aggregating in this way.

Option 2: use Nesta skills frequency tables

The detailed occupation-skill matrices used by Nesta (or available from other online job sources) can identify skills which are considered important at a much more granular level. For example, if we saw CAD software being specified heavily in adverts for architects, this could indicate that it is an acutely needed skill in that occupation.

Clearly, a skill may be specified in an advert simply because it is a general requirement without indicating any distinct shortage. For example, CAD may just be considered a core part of the current architect skillset rather than being indicative of any systematic mismatch between employers' requirements and the skillsets of potential applicants. The solution, therefore, would be to benchmark the skill request frequency against an objective measure of the job requirements.[29] If the skill request frequency was higher than what would be expected based on skills requirements, this would be indicative of a shortage *within the occupation*.

An example of this sort of research is undertaken by Burning Glass, which, for baseline skills, benchmarks the advert frequency against O*NET importance ratings using data for the US.[30] This finds, for example, that project management skills are in shortage within design and engineering professions. A similar analysis could be undertaken with UK data. However, a constraint is that the benchmarking can only be done as far as the granularity of the skills requirements data permits, which in the case of benchmarking against O*NET is a limitation. Therefore, the approach would tell us whether there is a shortage of project management or communication skills among architects but not whether there is a shortage of CAD skills.

Option 3: bespoke analysis of online vacancy data

For more granular insight, an option would be to use raw online vacancy data. This would enable trends in skills intensity to be understood. For example, if we observed that among data scientists the proportion of adverts specifying Python was growing whereas SQL was falling, this could indicate that Python was growing relatively in demand. Observing this as a trend is more robust than simply observing a high frequency of requests in a current snapshot. However, it may simply point to skills that are in high

29 This could draw on objective evaluator-based measures such as O*NET, although this would be limited in terms of how granular such an assessment could be.

30 https://www.burning-glass.com/research-project/baseline-skills/ Note that, as these variables have different distributions, statistical techniques are needed to transform them to make them comparable.

churn, featuring in many vacancies that are subsequently met, meaning that this approach lacks external objective validation.

However, while this approach could identify shortages at a granular level in terms of ‘trending’ skills, it would not identify more long-term shortages.

Recommendation

Both the ESS and Nesta options are pragmatic ways of using available data to understand shortages of skills *within an occupation* at a broad skill level. A key advantage of using the Nesta data is that they enable analysis of specific occupations, whereas the ESS groups occupations at a very high level, which may limit its value. However, the benefit of using direct evidence via surveys rather than indirect measures to describe skills shortages should be clear.

Understanding shortages at a more granular level of skills would require using raw online job advert data, which would involve high resource requirements and is unlikely to be a desirable course of action.

Productivity

Here we consider how skills taxonomies and data sources can be used to aid understanding of productivity and the returns to skills. This shows which skills are attracting a premium, which has implications for tailoring the output of training or upskilling workers. It should be noted that both options discussed below use wage as a proxy for productivity. We consider two alternative options.

Option 1: earnings regression using taxonomy at an aggregate level

The concept of this approach follows Dickerson and Morris (2019),[31] who used successive vintages of O*NET to measure changes over time in an occupation’s skills requirements. Together with pay data from LFS and Annual Survey of Hours and Earnings (ASHE), this enables a regression of earnings, a proxy for productivity, on skill mix. The authors find a positive return to cognitive skills and a negative return to physical skills. Felstead et al. (2007)[32] reached similar findings using successive waves of the ‘Skills at Work’ survey. Abraham and Spletzer (2009) used a similar approach to regress wages on skill requirements based on a single vintage of O*NET.

[31] https://cver.lse.ac.uk/textonly/cver/pubs/cverdp020.pdf

[32] https://www.cardiff.ac.uk/__data/assets/pdf_file/0009/118683/1.-Skills-at-Work-in-Britain-mini-report.pdf

Although the analysis can be performed using administrative or survey data, as homogenous occupation-level skills profiles are used,[33] this limits the amount of variation that can be used to identify skills premia. As a result, authors have typically focused on analysing only three skills: cognitive, physical and interpersonal skills.

It would be desirable to understand skills premia at a more granular level. There is scope to attempt similar analysis using an augmented range of skills measures or to unpick results further into subsets of occupations.

Option 2: earnings regression using online vacancy microdata

Using vacancy microdata instead of skills requirements that are homogenous within an occupation greatly enhances the amount of variation available for identifying the effects of different skills. This would allow a much more granular analysis to be conducted. In theory, it could go as far as estimating the wage uplift associated with an architect role which requires CAD skills.

While the greater variation is advantageous, it also presents methodological difficulties. More granularity can bring the risk of spurious results, such as negative returns. Other difficulties include whether more skills in an advert mean a more skilled job as opposed to a more detailed advert, whether the returns relate only to specific occupations, as well as questions of representativeness in terms of which vacancies (or skills requirements) are advertised online.

Recommendations

A pragmatic next step would be to use occupation-level skills profiles in conjunction with labour market datasets such as LFS or ASHE to explore whether more granular skill results are feasible. A sensible starting point would be to expand gradually on some of the very high-level skills groupings that have been tested to date (e.g. cognitive, physical, interpersonal) and establish whether these give plausible results.

[33] This is in comparison to analysing the skills and labour market returns to individuals, whose pay and characteristics will vary from the occupational average, giving more data with which to recover the returns to skills.

4. Conclusion and recommendations

Conclusions

A skills taxonomy groups skills while outlining the skill requirements of an occupation. As most LMIs are presented at the occupation level, this enables us to translate labour market data into skills terms. Thus, instead of looking at the occupations in shortage or the occupations that will become more important, we can view this through a skills lens. The key to using a taxonomy in this way is that it can be linked to UK LMIs via UK SOCs. The skills taxonomies we have considered can all be mapped to UK SOCs in some way, although UK SOCs are often broader, necessitating some aggregation.

The hierarchical structure of a taxonomy gives a view of skills at different levels of granularity; the required level will vary by application. Inevitably, taxonomies vary in terms of how well they perform at these levels of granularity: a taxonomy which prioritises comparisons in broader terms may have less detail in terms of narrow skills, and a taxonomy rooted in detail may not be as comparable when analysed at a broader level.

O*NET can be considered to prioritise analysis of broad skills, with attributes summarised in a smaller number of measures than in other taxonomies. These can be further grouped using the higher-level groupings already presented in the taxonomy or other divisions such as ‘data-people-things’. The SPB’s work to date has focused on skills at this broad level and, therefore, O*NET is suitable for this purpose.

However, no skills taxonomy is perfect for all purposes. Taxonomies such as Nesta and ESCO present skills at a very granular level. If such detail were relevant in future applications of the SPB, the data from these taxonomies could be used relatively easily to supplement the evidence available from O*NET. Intermediate-level groupings may also be relevant, which would require analytical effort to develop the right approach to summarising and aggregating the data. The relative merits of evaluator-based assessments of requirements versus online advert frequency would need to be carefully considered as they measure quite different things.

Many extensions to the analysis point to different data sources, such as the ESS or online vacancy data, to measure skills shortage in more detail or building rich datasets that enable the productivity returns to be estimated. A key consideration is looking for ways to align these data sources with existing taxonomies.

In other respects, there is less scope for improvement without expending significant resource. For example, a qualification-skills mapping would require significant manual effort, with little existing work to draw on directly.[34] In addition, while it would be possible to identify

[34] We note that the Singapore Skills Frameworks map to levels in terms of the Workforce Skills Qualification system. However, it would be difficult to leverage this information without both mapping these levels to UK levels and the Singaporean skills into another taxonomy.

emerging skills and occupations using real-time data, this has limited value in the absence of other LMIs.

Recommendations

Our work has highlighted some analyses which would enhance the body of knowledge and address the gaps associated with the SPB's use of O*NET in its current form. These include:

1. **Using Nesta or ESCO to derive intermediate-level measures of skill requirements in technology or other areas.** These measures would be more granular than the elements used in O*NET but more aggregate than the individual skills used in Nesta or ESCO. To derive the intermediate-level measures, it would be necessary to aggregate individual skills in Nesta and ESCO in a way that is consistent with the broader categories contained in O*NET.
2. **Doing an approximate mapping between skills in O*NET and qualifications in terms of SSA and NQF level.** This approach could identify the types of qualifications prevalent in an occupation and link them to the prominent skills within the array of requirements. This type of analysis would be more challenging in the context of general skills and qualifications, but it could be used to get a sense of the group of skills associated with a qualification.
3. **Monitoring updates to taxonomies to ascertain where new occupations are added.** This would be relevant as part of a wider brief to keep abreast of developments in the analytical sphere. While awareness of newly added occupations or skills is important in understanding the evolving landscape, there is little evidence that it can be leveraged in relation to these new areas unless primary data collection is also undertaken.
4. **Combining O*NET with data from the ESS and online vacancy data to inform views on shortages within occupations.** This could shed light on the nature of shortages in terms of the specific skills needed rather than in terms of numbers of people. The ESS would provide a broad view of this, while the Nesta data would start at a very granular level and potentially need further analysis to arrive at more general findings.
5. **Considering whether extensions to previous research (such as that conducted by Dickerson & Morris[35]) could provide insights about the returns to other skills.** This would offer a pragmatic route to understanding whether more detailed findings are available regarding returns to skills than the 'data-people-things' paradigm.

35 https://cver.lse.ac.uk/textonly/cver/pubs/cverdp020.pdf

The options above have been highlighted based on delivering tangible improvement for the use cases without involving excessive resource. Any prioritisation of them would obviously depend on the direction of the SPB's work going forward and the resources available. That said, options 3. and 4. (using ESS) can be considered to involve minimal resource and therefore be relevant in any case. Options 1., 2. and 5. are more resource-intensive and need a clearer justification.

We note that online job vacancy data could also provide powerful granular insights in many of these areas. However, as this would have significant resource requirements and methodological pitfalls, careful judgement would be needed as to whether it was justified. That said, there would be economies of scale in using such data on several fronts at the same time.

Annex A – Overview of skills taxonomies

O*NET

The Occupational Information Network (O*NET)[36] system has been developed by the US Bureau of Labor Statistics over many years. It was first published in 1998 and has been continuously updated. It replaced the previous Dictionary of Occupational Titles (DOT) system, which was first published in 1939.

O*NET provides information for each occupation.[37] There are around 1,000 different occupation codes in the US SOC.[38] For each occupation, O*NET includes 177 elements covering skills, knowledge, abilities, work activities and work style. Each element is given a rating for each occupation in terms of importance and level.[39] O*NET also includes various additional measures relating to training, experience, tools and technologies, and detailed tasks. These are described in the box below.

O*NET gives a rich assessment of the requirements of a job across all these dimensions. It goes well beyond listing just the skills that are prominent for an occupation by expressing requirements for all skills, including those that are not considered to be core. For example, rather than listing the plumbing skills that would be required of a plumber, O*NET describes the requirements across the whole range of areas covered.

Overview of O*NET descriptors and elements

O*NET contains the following descriptors, which are groupings of more detailed ‘elements’:[40]

- **Skill** (35 elements), grouped into: basic skills (10 elements), complex problem solving (1 element), resource management (4 elements), social skills (6 elements), systems skills (3 elements), technical (11 elements).
- **Knowledge**. These are organised sets of principles and facts across general domains (33 elements).
- **Ability**. These are ‘enduring attributes of the individual’. There is a question as to how far these can be improved by training. There are 52 elements, grouped into:

36 https://www.onetcenter.org/

37 This is recorded in a separate O*NET-SOC hierarchy, which is a slightly more detailed version of the US SOC.

38 The current version of O*NET (v26) lists 1,016 different occupation codes and has detailed information on 873 of these.

39 Level ratings are not given for the ‘work styles’ component. In practice, we find that level and importance ratings are highly correlated, with only a small number of cases where the level is higher/lower than expected, given the importance rating.

40 The term ‘element’ is used to refer to the individual skill, knowledge, ability fields that are then grouped together. For example, ‘near vision’ is an element within sensory abilities.

cognitive (21 elements), psychomotor (10 elements), physical (9 elements), sensory (12 elements).

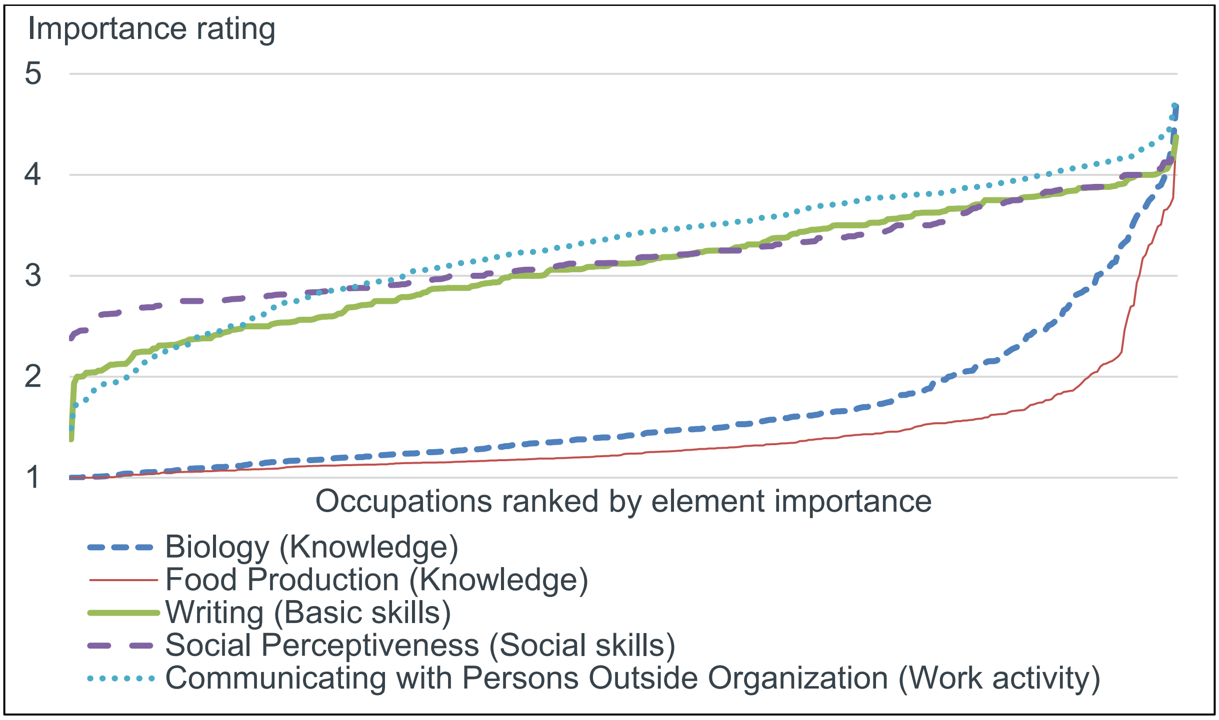
- **Work style**. These are personal qualities such as leadership or initiative (16 elements). Note, these are not given level ratings.
- **Work activity**. This summarises broad groups of tasks. There are 41 elements, grouped into: information input (5 elements), mental processes (10 elements), work output (9 elements), interacting with others (17 elements).

O*NET is updated periodically and draws on the input of both job evaluation experts and surveys of job incumbents.[41] O*NET data are open source and published in a flat Excel file format, which is advantageous from a practical perspective.

It should be noted that, while elements are scored across all occupations, some elements show low dispersion, being similarly demanded across occupations, whereas others are specific to certain occupations. This is illustrated in **Figure 3**, which shows the distribution of importance scores of a selection of elements across occupations. Each line shows the importance ratings of an element within occupations ranked in ascending order. Social perceptiveness and writing both have a flat profile, with most occupations having importance ratings above 3, indicating that the requirement is widespread. However, in the cases of biology and food production, only a small proportion of occupations have importance ratings above 3, indicating that the requirement is more specialised. This is more the case for the 'knowledge' and 'technical skills' groups.

[41] Skill and ability ratings are provided exclusively by job evaluation experts, whereas knowledge and other fields draw more extensively from job incumbents. Note that ratings by job incumbents are less reliable than the input of evaluation experts, although internal research by the SPB has not found any biases.

Figure 3 - General and specific skills in O*NET using an illustrative selection



Source: Frontier Analysis using O*NET data

The systematic grading of general skills across all occupations is unique to O*NET and something that sets it apart from other taxonomies. This is particularly useful in the context of understanding transferable skills, transitions to different occupations, and forecasting the demand for skills and how provision might be targeted accordingly. However, O*NET presents occupation-specific skills in considerably less detail than the other taxonomies explored for this work.

Applications of O*NET

- **Practical applications**. O*NET contains various portals that allow the user to inspect the profiles of different jobs in detail in an easy-to-use format. This includes 'My Next Move', which provides easy-to-use search options and career overviews for students and job seekers, and O*NET OnLine, which provides comprehensive occupational descriptions and data for use by career counsellors, workforce development offices, human resources professionals and researchers.
- **Skills shortage analysis**. The OECD Skills for Jobs programme estimates which skills are relatively in more shortage in different countries, to highlight key trends and patterns. The approach uses various labour market indicators (for example, wage growth, employment growth, hours worked growth, unemployment rate, change in underqualification) to develop a composite indicator of shortage at the occupational level. This is mapped to the O*NET skills matrix, and for each skill they then assess whether the occupations using it are in shortage, giving a sur-

plus/deficit score for each skill. This can then be used to compare skills mismatches across countries or industries. The SPB's work for Q1 has some similarities with this approach.

- **Labour market projections**. Various studies have combined the O*NET skills profiles with forecasts of labour demand by occupation to estimate which skills will become more important in the future. For example, a McKinsey report for the Industrial Strategy Council[42] uses cross-country analysis of consumption, ageing and technology trends to derive occupational demand projections. Using demographic projections and an adapted O*NET framework (with 'work skills' derived by manually aggregating task ratings), workers are moved into the most similar growth occupations. The supply and demand by skill can then be used to measure mismatches. Nesta (2017)[43] assesses future skills by using industry experts' evaluation of the prospects of a limited number of occupations, and then using machine learning to extrapolate from these assessments to other occupations based on similarity of skills profile in O*NET.
- **Returns to skills**. A recent paper by Dickerson and Morris[44] groups O*NET skills in a 'data-people-things' paradigm and uses successive waves of O*NET to explore how job requirements have changed over time. This is combined with ASHE and LFS data on pay to build an occupation-level panel dataset that controls for various characteristics and allows for estimation of the returns to cognitive, physical and interpersonal skills.

O*NET also contains other fields which indicate how much experience and education are required in an occupation, as well as lists of specific tools or types of software used.

O*NET has been used in a broad range of applications, as discussed in the box above. O*NET has also been used in other countries, such as Australia, New Zealand, China and Czech Republic. O*NET has been mapped to UK SOC codes and is integrated in the 'LMI For All' platform, where it is available in an Application Protocol Interface (API).[45]

European Skills, Competences, Qualifications and Occupations (ESCO)

ESCO is a major project coordinated by the European Commission's Directorate-General for Employment, Social Affairs and Inclusion (DG EMPL), the first version of which was published in July 2017. It provides a multilingual classification of skills, qualifications and occupations to promote labour mobility in the European Union.

---

42 https://industrialstrategycouncil.org/sites/default/files/UK%20Skills%20Mismatch%202030%20-%20Research%20Paper.pdf Previous research by McKinsey identifies automatable jobs in O*NET by assessing whether individual tasks can be automated.

43 The future of skills employment in 2030

44 https://cver.lse.ac.uk/textonly/cver/pubs/cverdp020.pdf

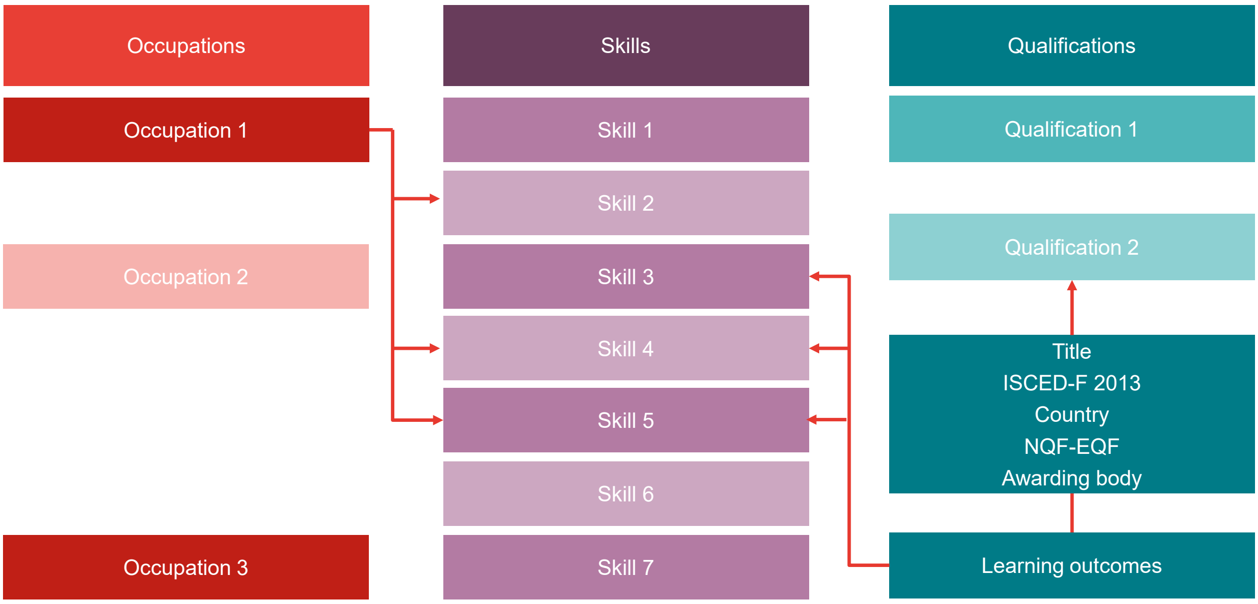
45 https://api.lmiforall.org.uk/

ESCO offers a comprehensive system for classifying skills, qualifications and occupations, systematically showing the relationships between the concepts. This provides an interface for different aspects of the labour market system, including employment figures, monitoring online vacancies and accreditation of qualifications. Intended uses include practical front-line applications (careers advice and training) and broad analysis to understand labour market trends.

The ESCO data are presented in a rich data format which can be interrogated by API. The data are not available in flat Excel files, so some expertise is needed to use ESCO practically.

ESCO comprises a skills pillar (skills, knowledge and competencies), an occupation pillar and a qualification pillar, which are linked as shown below:

Figure 4 - ESCO - links between occupation, skills and qualification pillars



Source: Adapted from ESCO Service Platform model

Occupation pillar

ESCO provides a hierarchical breakdown of occupations in successive levels of detail. The top four levels correspond to the International Standard Classification of Occupations (ISCO-08). ESCO can therefore be linked to UK data on occupation in a straightforward manner using existing mappings[46] published by the Office for National Statistics, in a similar manner as has been done for O*NET.[47]

Skills pillar

Each occupation specifies several skills, rated as either ‘essential’ or ‘optional’, i.e. a binary rating.[48] The skills pillar is a large hierarchy consisting of very granular skills grouped into many successive levels. For example:

Knowledge > ICT[49] > software and applications development > Python

Skills are also classified in terms of generality: whether a skill is occupation-specific, sector-specific, cross-sector or transversal. In the case of transversal skills, the binary rating in ESCO (compared to the continuous ratings provided in O*NET) may be overly simplistic if the skill is required to some degree across many occupations.

There are separate classifications for ‘attitude and values’, ‘knowledge’, ‘language’ and ‘skills’; each of these has its own hierarchy. Examples are shown in the text boxes for skills and knowledge, listing the various top-level components and then picking out and expanding a highlighted (emboldened) selection. For example, the knowledge hierarchy box starts with the top layer of aggregation before focusing on groups within the ICT category and then goes further into specific items within software and applications development. Overall, ESCO includes around 13,500 items in the skills pillar, and is therefore very detailed, with the bulk of this due to detailed coverage of occupation-specific skills.

46 https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2020/classifyingthestandardoccupationalclassification2020soc2020totheinternationalstandardclassificationofoccupationsisco08

47 This uses the Computer Assisted Structured Coding Tool (CASCOT) programme developed by the Warwick Institute for Employment Research. In some cases there are problems with the mapping, such as a small number of occupations not having a match, or an occupation straddling multiple codes. This is an inevitable problem if seeking to combine with other occupational classifications, and not a particular difficulty with ESCO.

48 Technically ‘ternary’ rather than ‘binary’, as a skill requirement has three settings: ‘essential’, ‘optional’ and ‘not required’.

49 Information and Communication Technology (ICT).

Knowledge hierarchy: zooming in on ICT

Agriculture, forestry, fisheries and veterinary

Arts and humanities

Business, administration and law

Education

Engineering, manufacturing and construction

Generic programmes and qualifications

Health and welfare

Information and communication technology

- computer use
- database and network design and administration
- information and communication technologies (ICTs) not further defined
- information and communication technologies not elsewhere classified
- **software and applications development and analysis**
    - AJAX
    - Ajax Framework
    - Android (mobile operating systems)
    - Ansible
    - …Python

Skills hierarchy: zooming in on ICT

S – skills
S1 – communication, collaboration and creativity
S2 – information skills
S2.0 – information skills
S2.1 – conducting studies, investigations and examinations
S2.2 – documenting and recording information

- compile statistical data for insurance purposes
- complete evaluation forms of calls
- fill out forms
- lay out digital written content
- maintain data entry requirements
- observe human behaviour
- output electronic files

S2.3 – managing information
S2.4 – processing information
S2.5 – measuring physical properties
S2.6 – calculating and estimating
S2.7 – analysing and evaluating information and data
S2.8 – monitoring, inspecting and testing
S2.9 – monitoring developments in area of expertise
S2.4.0 – processing information
S2.4.1 – gathering information from physical or electronic sources
S2.4.2 – entering and transforming information
S3 – assisting and caring
S4 – management skills
S5 – working with computers
S6 – handling and moving
S7 – constructing
S8 – working with machinery and specialised equipment

The presentation of skills at occupation level is only provided at the most granular level possible. For example, for a data scientist, ‘data mining’ is specified as essential knowledge but requirements are not defined at the more aggregate intermediate category it sits in (‘database and network design and administration’). In many cases, an intermediate level of aggregation would be of more use than the most granular level.[50] However, ESCO contains only the most granular skills, with no detail at any more

[50] In many cases, a granular skill will not even span multiple occupations, so it is hard to undertake any occupational skills comparison at that level.

aggregate level. As a result, a method would be needed to derive skills indicators at the more intermediate level.[51]

For example, a score for ‘S2 Information Skills’ could be calculated by counting the number of skills within that group that are required in an occupation. However, some groupings will be larger and will therefore appear more important if using a simple ‘adding up’ approach. This would need further thought to arrive at a system which allows skills requirements to be compared across occupations for categories of different sizes, such as the relative importance of ‘information skills’ and ‘working with computers’.[52] We are not aware of any work to date that derives skill requirements at these intermediate levels.

Qualifications pillar

ESCO also includes a qualifications pillar which is based on the European Qualifications Framework. Qualifications can be linked to both skills and occupations. The qualifications pillar is also intended to facilitate integration with the Europass online CV system, Open Badges, and other systems of digital credentials. For example, a Diploma in Maritime Studies – Sea Fishing has a unit on Vessel Construction and Stability, which corresponds to the ESCO skill ‘ensure watertight integrity’. In theory, these data sources would provide information on the stock of skills, as measured by qualifications.

It is for member states to populate the qualifications pillar, but so far little data have been added. It would also require mapping the qualifications to skills and occupation. A feasibility study[53] for the European Commission found that manual mapping would be resource-intensive and not justify the cost. A hybrid approach mixing human and machine learning with topic and concept matching was piloted, but the results were not very reliable. There is therefore little value so far in drawing on the ESCO qualifications pillar to supplement O*NET.

Nesta

The 2018 Nesta skills taxonomy[54] is a data-driven approach. It is estimated using 41 million online job adverts scraped from the web.[55] Raw text data are scraped and analysed using natural language processing to match synonymous skills terms. Then, using ‘graph clustering’, skills are grouped together based on appearing together in the

---

51 Aggregating up is more challenging in the case of ESCO than O*NET, because the ‘list’ approach means that some groupings of skills will be much larger, necessitating some approach to standardisation for comparability.

52 Potential approaches to normalise could include counting the number of skills required within a broader category and then measuring either in a ranking of occupations or relative to the maximum number specified.

53 Luomi-Messerer, Andersen, Wilson and Blakemore (2019), available at: https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8181&furtherPubs=yes

54 Please refer to https://escoe-website.s3.amazonaws.com/wp-content/uploads/2020/07/13161304/ESCoE-DP-2018-13.pdf for more detail on this project

55 Using data from Burning Glass covering the years 2012-2017.

same job advert. As an empirical approach with job advert as the unit of analysis, the partitioning largely follows occupational lines. For example, construction skills will be placed in a separate cluster to information technology (IT) skills, as IT jobs and construction jobs tend to specify different groups of skills. This method produces a tree-like hierarchy with successive stages at which different groups split off and are subdivided further. Part of the Nesta hierarchy is shown below, with three layers of subdivision, although each of the elements to the right will itself be a grouping. It should be noted that this approach is less suited to analysing general/transversal skills as, by definition, they will not be in any one occupational cluster.

Note that for the clustering algorithm to work, transversal skills need to be removed, as they would otherwise dominate the subsequent hierarchies. This results in 66 skills being removed, including e.g. communication, organisation, teamwork, writing, planning, research, English and problem solving. Many less ubiquitous broad skills are retained on the basis that 'some broad skills serve as a glue connecting more niche skill clusters'. So whereas O*NET covers general skills with some specialist areas presented at a high level, the Nesta hierarchy largely focuses on occupation-specific skills. The exclusion of transversal skills is somewhat problematic, as they are clearly important in facilitating movement between jobs.

The taxonomy also provides a matrix which sets out for each occupation the proportion of job adverts which specify a particular skill, thus providing a skills-occupation mapping. This says, for example, that in occupation X, 50% of adverts require skill A and 35% require skill B. This gives numerical indicators of skills intensity in a continuous manner, in the same way as the importance ratings in O*NET.[56] However, it should be noted that job importance ratings and advert frequencies are different, both conceptually and statistically. Online vacancy hierarchies focus on salient skills: the skills considered important to employers when posting a vacancy. However, job adverts may focus on 'top-of-mind' skills, with the assumption that other more general or transversal skills will be taken as given or covered elsewhere in qualification or experience requirements.[57] This is useful for understanding specific skills in shortage (e.g. if there is a need for architects to have skills in computer aided design) but not for understanding skills requirements more generally. Online job advert-based data are therefore particularly relevant for analysing the pattern of specific tangible skills within an occupation (including direct measures of changes over time) rather than more general peripheral skills that are not prioritised in a job advert.

[56] Note that currently published output includes the interactive skills tree and top skills listings for the top 200 job titles in a pdf report. Nesta's intention is to make the subsequent version of the taxonomy fully open source, providing the required information in a fuller and more digestible form.

[57] For example, an advert for a research physicist might not specify high numerical skills if these are already considered a pre-requisite within the field and do not warrant specific mention.

Practical usage of Nesta taxonomy: Regional analysis

Recent research[58] by Nesta uses the taxonomy to analyse regional skills mismatches.

Job adverts are assigned to SOC codes by Burning Glass. These SOC codes are then mapped to skills profiles using the Nesta taxonomy. A measure of skills supply comes by linking census and Annual Population Survey worker numbers by SOC code to the occupational skills profile from the Nesta taxonomy.[59] Similarly, skills demand is measured using Office for National Statistics vacancy data, again linked to the skills profile. This enables a comparison of skills shares within the stock of workers (supply) and vacancies (demand). A skill is in shortage if it is over-represented in vacancies relative to supply. This is analogous to the Q1 analysis by SPB, but it uses different shortage indicators, and the Nesta taxonomy focuses on occupational skill sets whereas the O*NET focuses more on intensity of transversal skills.

Skills Framework for the Information Age (SFIA)

The SFIA is a detailed competency framework for describing skill requirements in digital occupations. It includes 102 digital skills arranged in a three-layer hierarchy. For example ‘Network planning’ is a skill within the ‘Technical strategy and planning’ subcategory of the ‘Strategy and architecture’ category. The skills cover a range of specialist areas, such as data visualisation, testing and financial management. They are described in seven levels of competence.[60] This is intended to facilitate practical training and career development within a firm, such as how the requirements progress up the levels. It is also designed to support education and training providers and the design of curricula and accreditation provision.

The framework is updated every three to four years in a collaborative exercise which brings together industry experts. It is therefore relatively timely and focused on a specific area.

Although SFIA is not directly linked to occupation codes, SFIA undertook a bespoke mapping to the Saudi Arabia occupational hierarchy, which is based on the International Standard Classification of Occupations, and which in turn can be linked to UK SOC codes. This includes competency levels for each individual occupation. The coding would say, for example, that a data scientist needs analysis and visualisation skills at level 5.

58 https://productivityinsightsnetwork.co.uk/app/uploads/2019/08/Nesta_regional_skill_mismatch_reportv2.pdf

59 This implicitly assumes that an individual’s skills fully match those of their job.

60 The levels of competence do not correspond to UK qualification levels. The 7 levels are: 1 – follow; 2 – assist; 3 – apply; 4 – enable; 5 – ensure/advise; 6 – initiate/influence; 7 – set strategy, inspire, mobilise. Each level is described in terms of autonomy, influence, complexity, knowledge and business skills.

Some work would be needed to understand how the levels used in SFIA correspond to levels used in other systems such as O*NET or UK qualifications.

SFIA has been mapped to UK qualifications for specific bodies, such as Government Digital Service (on digital apprenticeships), and the Cabinet Office, for work on the Digital, Data and Technology Profession (DDaT) capability framework.[61] This has produced a range of job-related skills profiles and competency frameworks.

Other international taxonomies

Below we consider taxonomies developed in Singapore and Canada. As neither of these is published in an open format skills requirements by occupation, these are presented as examples rather than as full options for use.

Singapore Skills Taxonomy (SST)

The SST is a hierarchical classification which clusters skills based on their similarities in terms of importance in the same set of occupations. The taxonomy was developed in collaboration with Nesta and uses the same hierarchical clustering approach as the Nesta UK taxonomy. However, a key difference is the data used to generate the taxonomy: whereas the Nesta UK taxonomy uses online job adverts, the SST uses the detailed text set out in Singapore Skills Frameworks (SFws).

SFws are detailed frameworks which set out the job descriptions, competencies, work functions, tasks and skill requirements of different occupations within a sector. At the time of writing, frameworks have been developed for 34 different sectors, such as the built environment, retail and healthcare. An example from the ‘Energy and Chemicals’ SFw is shown below. Technical skill levels correspond to Singaporean qualification levels.

Figure 5 - Singapore Skills Framework example – requirements for the ‘Operations Specialist’ role

[61] This is a project to specify the skills needs of different digital jobs with the UK Civil Service.

TECHNICAL SKILLS AND COMPETENCIES		GENERIC SKILLS AND COMPETENCIES (TOP 5)	
Change Management	Level 4	Problem Solving	Intermediate
Continuous Improvement Management	Level 3	Decision Making	Intermediate
Crisis Management	Level 3	Teamwork	Intermediate
Data and Statistical Analytics	Level 3	Communication	Intermediate
Emergency Response Management	Level 4	Interpersonal Skills	Intermediate
Engineering Drawing Interpretation and Management	Level 4		
Engineering Management of Change	Level 3		

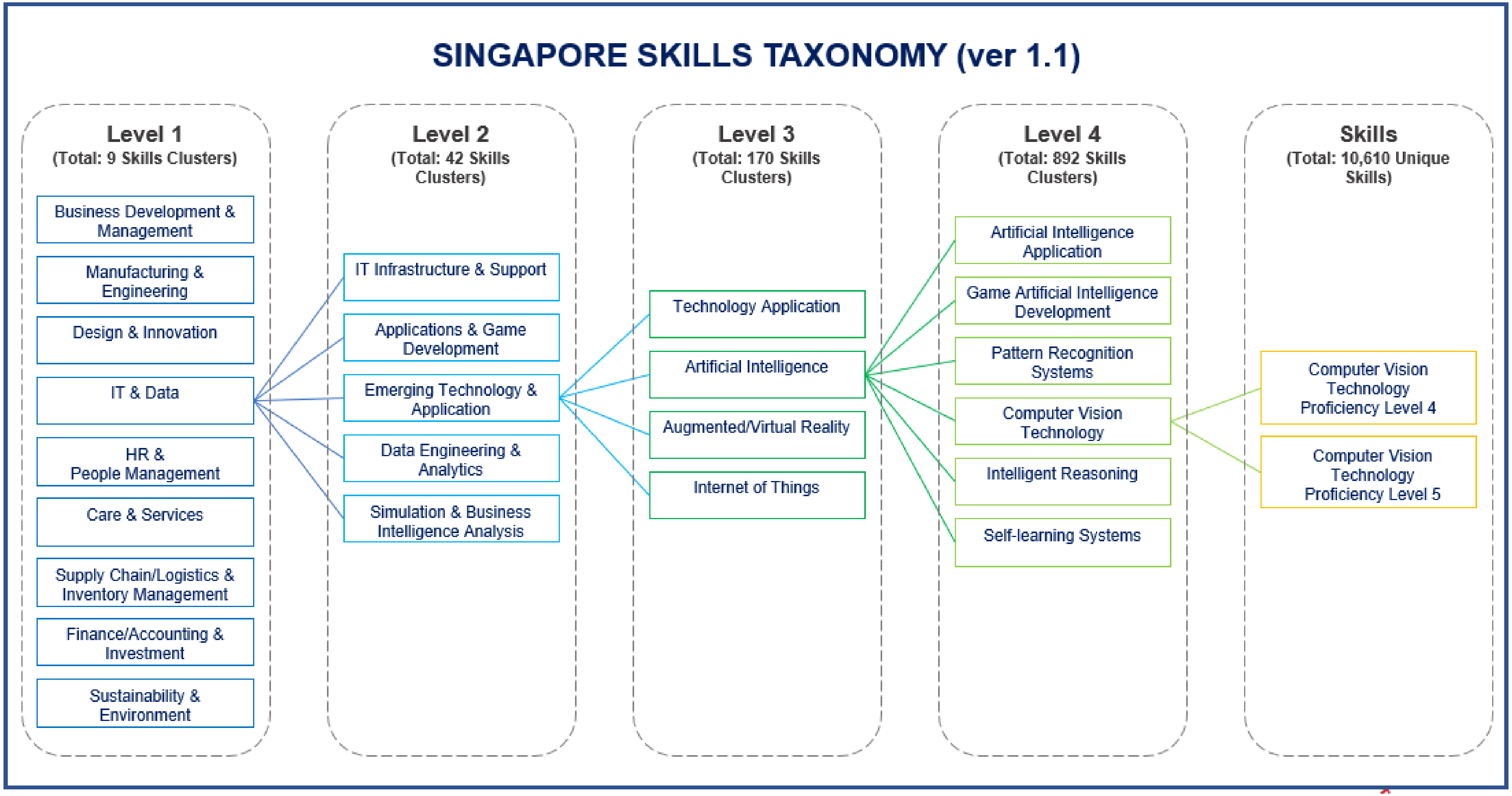
Source: Skills Future SG

The SFws are developed in collaboration with industry stakeholders and serve practical uses such as career advice and training. The use of industry experts means this hierarchy does not face the same generality/representativeness caveats which would apply in relation to a taxonomy based on online job adverts.

In total, derivation of the taxonomy incorporates 10,000 different skills, 34 sectors and 1,692 job roles.[62] The resulting hierarchy has five layers and parallels the Nesta UK taxonomy in terms of the focus on classifying technical skills and the partition being defined in terms of occupational skillsets. Only 16 generic skills are included. These are categorised into basic/intermediate/advanced ratings. Such an approach presents problems when considering the importance of transversal skills in the context of facilitating job changes. As with the ESCO and UK Nesta taxonomies, the SST is also rooted in a very granular level of detail, and effort would be needed to adapt it to give higher-level categories.

62 As a comparison, the UK SOC includes 30,000 job roles/titles in its index, which would make application in a UK context difficult. The taxonomy utilises rich skills framework data that are created in the Singaporean context. Such necessary data does not exist for the UK and would be more difficult to compile, given the size and diversity of the UK economy.

Figure 6 - Summary of the Singapore Skills Taxonomy



Source: Skills Future SG

The SST and SFw sit in a wider Jobs-Skills Repository system which is envisaged to bring in job adverts, CVs and training output and link to occupation and industry classifications. This system can be considered exemplary in systematically integrating these various aspects of labour market information into a unified framework. However, considerable resource is likely involved in generating and maintaining the frameworks. Given the comprehensive and detailed nature of this system, significant manual effort would be needed to adapt it to UK purposes and (as with ESCO) investment in adapting the wider labour market information system to this framework. This would seem to be outside the scope of work considered by the SPB.

Canada Skills and Competencies Taxonomy

Employment and Social Development Canada (ESDC) has developed a Skills and Competencies Taxonomy. The taxonomy uses several Canada labour market sources, the American O*NET system as well as a variety of national and international competency-based frameworks.[63] Descriptors are drawn from across these different sources.

In many cases, the taxonomy uses elements from O*NET, but with descriptors placed together in a grouping which corresponds to a category developed by the ESDC. For example, 'Oral Communication' is a skill from the preceding Canadian skills framework but, for consistency with O*NET, it has been split into three descriptors: 'Active Listening', 'Oral Comprehension', and 'Oral Expression'.[64] The skills field is split into the following categories: foundational, analytical, technical, resource management and interpersonal skills. In other cases, new descriptors have been added from preceding parts of the Canadian skills frameworks.[65]

This approach illustrates how a taxonomy from another country can pragmatically be adapted and repurposed.

In terms of practical usage, the taxonomy will be used in the Occupational and Skills Information System (OaSIS) which will measure these descriptors for around 900 occupations and provide a similar tool with broad applications as O*NET in the USA. The taxonomy to date therefore only provides the groupings of elements; the detailed data on skills requirements by occupation will come with OaSIS.

---

63 Frameworks drawn on include Career Handbook; Skills and Knowledge Checklist; Classification of Instructional Programs Canada (CIP); Essential Skills Framework (ESF). The ESF is the preceding framework that the Skills and Competency Framework replaces.

64 The expanded skills section is shown in full at https://noc.esdc.gc.ca/SkillsTaxonomy/Skills

65 For example, the Knowledge field uses descriptors from the Classification of Instructional Programs Canada modified with information from the O*NET system and the National Occupational Classification's Skills and Knowledge Checklist.

Annex B – Full O*NET taxonomy

ONET_descriptor	ONET_group	Element_Name
Ability	Cognitive abilities	Oral Comprehension
Ability	Cognitive abilities	Written Comprehension
Ability	Cognitive abilities	Oral Expression
Ability	Cognitive abilities	Written Expression
Ability	Cognitive abilities	Fluency of Ideas
Ability	Cognitive abilities	Originality
Ability	Cognitive abilities	Problem Sensitivity
Ability	Cognitive abilities	Deductive Reasoning
Ability	Cognitive abilities	Inductive Reasoning
Ability	Cognitive abilities	Information Ordering
Ability	Cognitive abilities	Category Flexibility
Ability	Cognitive abilities	Mathematical Reasoning
Ability	Cognitive abilities	Number Facility
Ability	Cognitive abilities	Memorization
Ability	Cognitive abilities	Speed of Closure
Ability	Cognitive abilities	Flexibility of Closure
Ability	Cognitive abilities	Perceptual Speed
Ability	Cognitive abilities	Spatial Orientation
Ability	Cognitive abilities	Visualization
Ability	Cognitive abilities	Selective Attention
Ability	Cognitive abilities	Time Sharing
Ability	Psychomotor abilities	Arm-Hand Steadiness
Ability	Psychomotor abilities	Manual Dexterity
Ability	Psychomotor abilities	Finger Dexterity
Ability	Psychomotor abilities	Control Precision
Ability	Psychomotor abilities	Multilimb Coordination
Ability	Psychomotor abilities	Response Orientation
Ability	Psychomotor abilities	Rate Control

Ability	Psychomotor abilities	Reaction Time
Ability	Psychomotor abilities	Wrist-Finger Speed
Ability	Psychomotor abilities	Speed of Limb Movement
Ability	Physical abilities	Static Strength
Ability	Physical abilities	Explosive Strength
Ability	Physical abilities	Dynamic Strength
Ability	Physical abilities	Trunk Strength
Ability	Physical abilities	Stamina
Ability	Physical abilities	Extent Flexibility
Ability	Physical abilities	Dynamic Flexibility
Ability	Physical abilities	Gross Body Coordination
Ability	Physical abilities	Gross Body Equilibrium
Ability	Sensory abilities	Near Vision
Ability	Sensory abilities	Far Vision
Ability	Sensory abilities	Visual Color Discrimination
Ability	Sensory abilities	Night Vision
Ability	Sensory abilities	Peripheral Vision
Ability	Sensory abilities	Depth Perception
Ability	Sensory abilities	Glare Sensitivity
Ability	Sensory abilities	Hearing Sensitivity
Ability	Sensory abilities	Auditory Attention
Ability	Sensory abilities	Sound Localization
Ability	Sensory abilities	Speech Recognition
Ability	Sensory abilities	Speech Clarity
Work style	Work styles	Achievement/Effort
Work style	Work styles	Persistence
Work style	Work styles	Initiative
Work style	Work styles	Leadership
Work style	Work styles	Cooperation
Work style	Work styles	Concern for Others
Work style	Work styles	Social Orientation

Work style	Work styles	Self Control
Work style	Work styles	Stress Tolerance
Work style	Work styles	Adaptability/Flexibility
Work style	Work styles	Dependability
Work style	Work styles	Attention to Detail
Work style	Work styles	Integrity
Work style	Work styles	Independence
Work style	Work styles	Innovation
Work style	Work styles	Analytical Thinking
Skill	Basic skills	Reading Comprehension
Skill	Basic skills	Active Listening
Skill	Basic skills	Writing
Skill	Basic skills	Speaking
Skill	Basic skills	Mathematics Skill
Skill	Basic skills	Science
Skill	Basic skills	Critical Thinking
Skill	Basic skills	Active Learning
Skill	Basic skills	Learning Strategies
Skill	Basic skills	Monitoring
Skill	Social skills	Social Perceptiveness
Skill	Social skills	Coordination
Skill	Social skills	Persuasion
Skill	Social skills	Negotiation
Skill	Social skills	Instructing
Skill	Social skills	Service Orientation
Skill	Complex problem solving skills	Complex Problem Solving
Skill	Technical skills	Operations Analysis
Skill	Technical skills	Technology Design
Skill	Technical skills	Equipment Selection
Skill	Technical skills	Installation

Skill	Technical skills	Programming
Skill	Technical skills	Operation Monitoring
Skill	Technical skills	Operation and Control
Skill	Technical skills	Equipment Maintenance
Skill	Technical skills	Troubleshooting
Skill	Technical skills	Repairing
Skill	Technical skills	Quality Control Analysis
Skill	Systems skills	Judgment and Decision Making
Skill	Systems skills	Systems Analysis
Skill	Systems skills	Systems Evaluation
Skill	Resource management skills	Time Management
Skill	Resource management skills	Management of Financial Resources
Skill	Resource management skills	Management of Material Resources
Skill	Resource management skills	Management of Personnel Resources
Knowledge	Knowledge	Administration and Management
Knowledge	Knowledge	Clerical
Knowledge	Knowledge	Economics and Accounting
Knowledge	Knowledge	Sales and Marketing
Knowledge	Knowledge	Customer and Personal Service
Knowledge	Knowledge	Personnel and Human Resources
Knowledge	Knowledge	Transportation
Knowledge	Knowledge	Production and Processing
Knowledge	Knowledge	Food Production
Knowledge	Knowledge	Computers and Electronics
Knowledge	Knowledge	Engineering and Technology
Knowledge	Knowledge	Design
Knowledge	Knowledge	Building and Construction
Knowledge	Knowledge	Mechanical
Knowledge	Knowledge	Mathematics
Knowledge	Knowledge	Physics

Knowledge	Knowledge	Chemistry
Knowledge	Knowledge	Biology
Knowledge	Knowledge	Psychology
Knowledge	Knowledge	Sociology and Anthropology
Knowledge	Knowledge	Geography
Knowledge	Knowledge	Medicine and Dentistry
Knowledge	Knowledge	Therapy and Counseling
Knowledge	Knowledge	Education and Training
Knowledge	Knowledge	English Language
Knowledge	Knowledge	Foreign Language
Knowledge	Knowledge	Fine Arts
Knowledge	Knowledge	History and Archeology
Knowledge	Knowledge	Philosophy and Theology
Knowledge	Knowledge	Public Safety and Security
Knowledge	Knowledge	Law and Government
Knowledge	Knowledge	Telecommunications
Knowledge	Knowledge	Communications and Media
Work activity	Information input	Getting Information
Work activity	Information input	Monitoring Processes, Materials, or Surroundings
Work activity	Information input	Identifying Objects, Actions, and Events
Work activity	Information input	Inspecting Equipment, Structures, or Material
Work activity	Information input	Estimating the Quantifiable Characteristics of Products, Events, or Information
Work activity	Mental processes	Judging the Qualities of Things, Services, or People
Work activity	Mental processes	Processing Information
Work activity	Mental processes	Evaluating Information to Determine Compliance with Standards
Work activity	Mental processes	Analyzing Data or Information
Work activity	Mental processes	Making Decisions and Solving Problems
Work activity	Mental processes	Thinking Creatively

Work activity	Mental processes	Updating and Using Relevant Knowledge
Work activity	Mental processes	Developing Objectives and Strategies
Work activity	Mental processes	Scheduling Work and Activities
Work activity	Mental processes	Organizing, Planning, and Prioritizing Work
Work activity	Work output	Performing General Physical Activities
Work activity	Work output	Handling and Moving Objects
Work activity	Work output	Controlling Machines and Processes
Work activity	Work output	Operating Vehicles, Mechanized Devices, or Equipment
Work activity	Work output	Interacting With Computers
Work activity	Work output	Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment
Work activity	Work output	Repairing and Maintaining Mechanical Equipment
Work activity	Work output	Repairing and Maintaining Electronic Equipment
Work activity	Work output	Documenting/Recording Information
Work activity	Interacting with others	Interpreting the Meaning of Information for Others
Work activity	Interacting with others	Communicating with Supervisors, Peers, or Subordinates
Work activity	Interacting with others	Communicating with Persons Outside Organization
Work activity	Interacting with others	Establishing and Maintaining Interpersonal Relationships
Work activity	Interacting with others	Assisting and Caring for Others
Work activity	Interacting with others	Selling or Influencing Others
Work activity	Interacting with others	Resolving Conflicts and Negotiating with Others
Work activity	Interacting with others	Performing for or Working Directly with the Public
Work activity	Interacting with others	Coordinating the Work and Activities of Others
Work activity	Interacting with others	Developing and Building Teams
Work activity	Interacting with others	Training and Teaching Others
Work activity	Interacting with others	Guiding, Directing, and Motivating Subordinates
Work activity	Interacting with others	Coaching and Developing Others
Work activity	Interacting with others	Providing Consultation and Advice to Others

Work activity	Interacting with others	Performing Administrative Activities
Work activity	Interacting with others	Staffing Organizational Units
Work activity	Interacting with others	Monitoring and Controlling Resources