Introduction

In response to the PIP and ESA Assessment inquiry, this submission summarises the results of a four year ESRC research project into incapacity benefit assessment. The submission focuses on ESA assessments (the WCA), which were the focus of the study.

While at the end of the report I consider a number of the Committee’s questions that relate to establishing the ‘genuineness’ of claimants, this submission particularly focuses on the Committee’s question, “What examples of best practice in assessing eligibility for benefits are available internationally, and how transferrable are they to ESA and/or PIP?” My project included a comparative study of assessments in nine countries: four Anglo-Saxon countries (the USA, Canada, Australia and New Zealand) and five European countries that international experts identified as likely best practice (the Netherlands, Germany, Norway, Sweden, and Denmark). The review was based on over 150 documents and 40 expert interviews, and has been reported in detail in two peer-reviewed papers (Geiger et al., in press, Geiger, 2017).

Further detail is available on everything that I discuss below, and there are also other elements of the project that look at opinions about the WCA among the general public and elite groups, all of which I would be happy to send on to / discuss with the Committee. For transparency, I should add that I worked on secondment as an expert advisor on ESA and WCA policy at the DWP in 2015-16; the views here are my own, are should not be taken to reflect the DWP, the ESRC, the University of Kent, or any other body.

International best practice

Why the WCA fails to assess work capability

In my view – and that of many others who understand the WCA process – one of the key problems of the WCA is that it fails to assess work capability. There are two problems. Firstly, there is no transparent evidence that the functional descriptors match the actual demands of jobs in contemporary Britain, and many people believe that they are too harsh. Secondly, and more fundamentally, the WCA is inaccurate if claimants have two or more types of impairment, which probably includes at least half of all disabled people.\(^1\) This is because the scores for each type of impairment are added together to determine if people are put in the ESA WRAG. However, the combined score has no relationship whatsoever to whether someone with this particular functional ‘profile’ – that is, this particular combination of functional impairments – will have a chance of working.

---

\(^1\) The claim that around half of disabled people have two or more impairments comes from a new analysis of the Health Survey for England 2014. People were asked if their longstanding health condition affected them in any of nine different domains (vision, hearing, mobility, dexterity, learning or understanding or concentrating, memory, mental health, stamina or breathing or fatigue, and socially or behaviourally). Of those aged 18-64 reporting limitations in at least one domain, 53% reported limitations in multiple domains. These domains are not exactly the same as the WCA categories, but if anything it seems likely that the greater number of categories in the WCA would lead to even higher figures there. While being slightly further away from the WCA descriptors, a DWP study has similarly shown that two-thirds of claimants have multiple health conditions, rather than just a single condition; Sissons, P., Barnes, H., and Stevens, H. (2011). *Routes onto Employment and Support Allowance*. Research Report No 774. London: DWP (Department for Work and Pensions).
Strikingly, I was fortunate to be able to conduct a single focus group with Maximus WCA assessors in one part of the country, and they highlighted this as their biggest issue:

**Assessor B:** “There’s some that you see who’ve maybe got ten, fifteen things, and they all impact slightly. And therefore in one descriptor they don’t reach the points, so you think actually they could walk for ten minutes but they really cannot do more than that. They might be going to the toilet ten times a day but not actually having incontinence. And all these things in one person, as well as lots of mental health issues. But they still go out themselves, they still can speak to people cause they’re polite people, but they’re not actually physically able to get out much, but the evidence is there that they probably could do 200 metres. So again it’s the descriptors they don’t match, but when you match all of them together, they couldn’t go to work for whatever, sixteen hours a week.”

**Assessor C:** “Just because somebody might be able to walk 100 metres one day, they might not be able to manage it for five or six days a week because of their mental health. And I think there’s an overlap [between mental and physical health] but I don’t think that the WCA necessarily covers that, so you could say from a whole picture that they are probably too unwell to work. You know, some of the ones who have sort of a variety of moderate severity conditions but nothing that would put them into a Support Group or take them over the threshold, I think that there are some gaps. And I can think of a few assessments where after they’d left I’ve thought, ‘It’s such a shame because there wasn’t quite a way for them to…’ Do you know what I mean?”

The change to ESA in April 2017 makes matters even worse. Previously, people with multiple less severe impairments that scored 15 points at the WCA could receive £102/wk in the WRAG – noticeably more than the £73 on JSA, and almost as much as the £109/wk in the Support Group. Now people in the WRAG receive the same benefit as those on JSA. This means that multiple less severe impairments do not ‘add up’ to a more severe one (even in an unsatisfactory way): the only thing that is considered is whether your most severe impairment reaches a certain threshold. There is no inherent reason why multiple impairments should be ignored; indeed, the PIP assessment adds up points from different types of impairment. In contrast, the WCA has always dealt poorly with multiple types of impairment, and now simply does not deal with them at all.

However, without understanding how other countries conduct their assessments, it is not clear how a revised WCA could do better capture work capability. On the basis of my comparative study, I argue that there are three different types of direct work capability assessment: expert assessments, demonstrated assessments, and structured assessments.

**International best practice: Expert assessments**

The first form of directly assessing work capability is the most common: to ask a professional to use their expertise to judge whether an individual is capable of work. However, there are longstanding concerns about the consistency and validity of such discretionary assessments.

One issue is around correctly understanding the demands of work. Commonly assessments are done by independent doctors (Germany) or allied health professionals (Australia, Canada), but the training of these professionals is around diagnosing/treating ill-health rather than occupational health. Nor do we have a clear idea of what assessors consider to be the general demands of the workplace – insurance physicians tend not to mention job requirements explicitly when making individual decisions about work capability.² One solution is to adopt a new professional category with more

---
relevant expertise and more explicit reporting requirements, such as the Dutch professional category of ‘labour market experts’. It is this type of expert assessment that has been proposed by some disability activists in the UK, in one of the few relatively detailed proposals for replacing the WCA.52

A further key issue is to ensure consistency in this type of discretionary assessment. A recent systematic review found that expert assessments of work ability “show high variability and often low reliability”.3 They suggest that this can be partly combated through standardisation, which can be seen in, for example, the standardised inputs that are prepared for rehabilitation assessment meetings in Denmark, via a standard rehabilitation plan that is completed by the claimant in partnership with their caseworker. The expert-based elements of assessment in the Netherlands are perhaps the most structured, in which insurance physicians follow both interview protocols and disease-specific guidelines for assessing work-related functioning.4 Yet even with such standardisation, getting consistent work capability judgements from expert assessment is difficult.5

For both of these reasons, there can be a considerable gap between the formal definition of work capability being assessed, vs. the actual criteria used by assessors. Even today, experts I spoke to in Australia variously described their benchmark hours criterion as ‘arbitrary’ and ‘almost a fictitious construct’.

Overall, experts can assess work capability with some degree of legitimacy, and are used in many systems around the world. Nevertheless, there are some concerns over the validity and reliability of their judgements. These may be partially mitigated through appropriating training/expertise, and standardisation of inputs, decision protocols and reporting requirements.

**International best practice: Demonstrated assessments**

A fundamental challenge in work capability assessment is that many people’s functional capacities and ability to cope in different workplaces are inherently uncertain. Leading models of supported employment such as Individual Placement and Support therefore use an iterative learning process to assess an individual’s work capability: they try the most suitable work environment first, and see how the person manages. A similar principle can be applied to incapacity benefits assessment: work capability can be assessed based on the actual experiences of the individual in the labour market – hence this is a ‘demonstrated assessment’.

Perhaps its clearest statement can be seen in an Australian high-level strategy document, where the assessment was tasked with assessing claimants’ work capability over the next two years, but the author noted that for many claimants “there is little or no practical evidence on which to base this judgment”.6 It therefore recommended that most claimants should only be eligible for the disability

---


pension “when their ‘Continuing Inability to Work’ has been demonstrated”. Since the ensuing reforms, Australian claimants need to actively participate in a ‘program of support’ for 18 months before being eligible for the disability pension,7 at which point they are referred to an expert assessment. Similar reasoning can be seen in Denmark, where claimants are now only awarded a disability pension if an assessing multidisciplinary team is confident that the individual has no capacity for work.8 In practice the majority of claimants are required to go through a scheme called Resource Activation for one to five years.

Again, there are several issues that need to be considered in this approach. Firstly, because rehabilitation benefits are generally lower than disability pensions, critics have argued that this is simply a benefit cut for people who have no realistic chance of work. For example, in Denmark, there has been considerable media and political attention on those placed in work trials or Resource Activation who have very low levels of assessed work capacity, including includes a widely-reported case of someone with 30mins of work capacity at low speed, twice per week.9

Second, these assessments require investment in both assessment and rehabilitation. For assessment, there is still a need for considerable expertise in interpreting people’s past experiences and in deciding what future rehabilitation steps are still feasible (if any). For rehabilitation, this model only provides an accurate picture of work capability if people go through rehabilitation that maximises work capability. In practice, however, there are examples from almost every country where this does not happen. For example, despite a series of reforms in Australia, a recent Government consultation found that “providers and people with disability expressed widespread, almost universal, concern about [the assessments], including consistent feedback that they often refer people with disability to inappropriate services”.10

Finally, even though demonstrated assessments seem to overlap most strongly with assessments for employment support, this overlap is only partial. This is partly because the claimants’ relationship with the assessor may be one of distrust when being evaluated for financial support, but more trusting when their rehabilitation needs are being evaluated. Yet even if these tensions can be overcome, modern ability-based rehabilitation needs to be based on a holistic assessment of an individual, including inter alia their motivation, but motivation is not usually considered a legitimate influence on benefit eligibility. Conversely, benefit eligibility assessments examine people’s capacity to do jobs that they have no desire to do, which is unhelpful for the purposes of rehabilitation. It is therefore possible to combine these assessments in an inefficient way that increases the resources required for assessment, which was a key reason why Australian dual-purpose assessments were later abandoned.

Overall, there are challenges with demonstrated assessments – but countries like Denmark seem to have managed to implement them with sufficient investment in expertise and rehabilitation to be a success. In the UK, there have been many calls to improve the link of the WCA to getting people back to work, from all sides of the political spectrum.11 The question is whether the UK is in a position to

---


11 Almost every proposal around the WCA has argued that there should be a greater link with employment support, even if
invest in an expert- and rehabilitation-focused overhaul of the entire system at the present time, particularly given our historically weak systems of vocational rehabilitation.\footnote{A fierce critique of the history of vocational rehabilitation in Britain can be found in Grahame, R. 2002. The decline of rehabilitation services and its impact on disability benefits. \textit{Journal of the Royal Society of Medicine}, 95, 114-117. There have nevertheless some improvement since; Frank, A. 2016. Vocational Rehabilitation: Supporting Ill or Disabled Individuals in (to) Work: A UK Perspective. \textit{Healthcare}, 4, 46.}

**International best practice: Structured assessments**

The final type of directly assessing work capability is the ‘structured assessment’, exemplified by the Dutch system. The full set of claimants’ functional capacities are assessed, and these are then compared to the required functional profiles – that is, all the different capacities in combination that someone needs to be able to do that job – in 7,000 actually-existing jobs in the Netherlands in a database called CBBS.\footnote{A full list of sources is provided in my full paper, but key introductory sources on CBBS include Pennings, F. 2011. The New Dutch Disability Benefits Act: the link between income provision and participation in work. In: Devetzi, S. & Stendahl, S. (eds.) \textit{Too Sick to Work? Social security reforms in Europe for persons with reduced earnings capacity}. The Netherlands: Kluwer Law International, Schellart, A. J., Mulders, H., Steenbeek, R., Anema, J. R., Kroneman, H. & Besseling, J. 2011. Inter-doctor variations in the assessment of functional incapacities by insurance physicians. \textit{BMC Public Health}, 11, 864, Broersen, J. P. J., Mulders, H. P. G., Schellart, A. J. M. & van der Beek, A. J. 2012. The identification of job opportunities for severely disabled sick-listed employees. Ibid.12, 156-156.}

It covers 28 different functional domains against which claimants are assessed, allowing variation between regular demands and peak demands, as well as covering the required work pattern, education, experience and skills of the job. This provides an empirically-based assessment of jobs that the individual can do.

There are several issues that need to be considered in structured assessments. Firstly, like all good assessments, it works best if the assessor has considerable expertise in occupational health. The claimants’ functional profile is fed into the CBBS database, but a labour expert provides the final definitive judgement to ensure that obvious errors or data limitations do not lead to unfair decisions.\footnote{Tummers, L., Bekkers, V. & Steijn, B. 2009. Policy Alienation of Public Professionals. \textit{Public Management Review}, 11, 685-706, van Berkel, R. 2013. From Dutch disease to Dutch fitness? Two decades of disability crisis in the Netherlands. In: Lindsay, C. & Houston, D. (eds.) \textit{Disability Benefits, Welfare Reform and Employment Policy}.} Moreover, the Netherlands have been experimenting with personalised expert judgements as to possible job adjustments that would enable the person to work. In other words, while the database is a valuable aid to decision-making, it does not fully substitute for expertise. Secondly, while structured assessments can provide valid judgements of whether people should receive financial support, they are not necessarily helpful for helping people get back to work. They ignore psychosocial factors, do not start from the priorities of the individual in question, and do not consider what would help the individual to work. However, as the final WCA independent review pointed out, \textit{“determining benefit eligibility and supporting employment outcomes may not be compatible objectives”}.\footnote{p82-83 of Litchfield, P. 2014. An Independent Review of the Work Capability Assessment – year five. Presented to Parliament pursuant to Section 10 of the Welfare Reform Act 2007.}

Finally, collecting data on the requirements of jobs within a country can be expensive. Given the prohibitive cost of covering all jobs nationally, CBBS covers about 20\% of all of the possible occupational codes in the Netherlands, weighted towards ‘lower level jobs’ that are potentially available to all claimants. Nevertheless, it still requires a team of about 35 full-time specialists in the social insurance agency to make on-site observations of Dutch jobs. One alternative is to focus on the

The final WCA independent review pointed out, \textit{\textit{“determining benefit eligibility and supporting employment outcomes may not be compatible objectives”}.} Finally, collecting data on the requirements of jobs within a country can be expensive. Given the prohibitive cost of covering all jobs nationally, CBBS covers about 20\% of all of the possible occupational codes in the Netherlands, weighted towards ‘lower level jobs’ that are potentially available to all claimants. Nevertheless, it still requires a team of about 35 full-time specialists in the social insurance agency to make on-site observations of Dutch jobs. One alternative is to focus on the

\footnote{12 A fierce critique of the history of vocational rehabilitation in Britain can be found in Grahame, R. 2002. The decline of rehabilitation services and its impact on disability benefits. \textit{Journal of the Royal Society of Medicine}, 95, 114-117. There have nevertheless some improvement since; Frank, A. 2016. Vocational Rehabilitation: Supporting Ill or Disabled Individuals in (to) Work: A UK Perspective. \textit{Healthcare}, 4, 46.}


the exact form this takes varies. For example, compare the Labour disability manifesto’s commitment to \textit{“a personalised, holistic process which provides each individual with a tailored plan, building on their strengths and addressing barriers”} [https://labour.org.uk/wp-content/uploads/2017/10/manifesto-for-disabled-people-1.pdf] with the scheme outlined by one of Iain Duncan Smith MP’s former special advisors in Pickles, C., Holmes, E., Titley, H. & Dobson, B. 2016. Working welfare: a radically new approach to sickness and disability benefits. London: Reform.
functional requirements of a much smaller number of jobs. The Dutch SMBA assessment for youth
disability benefit provides functional profiles for 15 relatively light minimum wage jobs (e.g. 'parking
lot attendant', 'receptionist'), which are each meant to be representative of the requirements of wider
groups of jobs nationally.

Overall, despite some costs, the Dutch structured assessments seem to produce decisions that are
widely accepted as fair, and are consistently cited by international experts as best practice. In
providing a basis for ‘objective’, standardised assessments, they also seem to best fit the requirements
of the UK system.

How the UK system could learn from international evidence

The Committee has asked, ‘Should the options for reforming the Work Capability Assessment mooted
in the Government’s Improving Lives green paper be taken forward?’ However, I would argue that
the Green Paper provides little detail about the future of the disability assessment for financial
support, other than saying it “should still focus on the impact that an individual’s health condition has
on them” (¶135). There is much in the Green Paper on how the WCA makes assessment as they relate
to conditionality; this is a separate conversation, but I am happy to send on further material if this
should prove to be useful (I review the evidence and makes a series of recommendation in an
upcoming Demos report).

In the same Demos report, I consider how we could learn from this international evidence and ensure
that the WCA actually assesses work capability. I argue that the Government should:

1. Overhaul the WCA descriptors, so that they transparently reflect the British labour market. It
would be relatively straightforward to do this: the Government could collect data on the functional
requirements of British jobs – that is, the specific capabilities that people need to be able to do each
job. There are various choices open to the Government in implementing a structured assessment
(around who collects the data, how many jobs data should be collected for, and around whether this
should be a ‘real-world’ test), but any of these would be a major step forward on the WCA.

2. Overhaul the structure of the WCA, so that it looks at the combined impact of multiple impairments
on work capability. It is again straightforward to do this, if we follow the previous recommendation
and collect data on the functional requirements of work in Britain. Instead of just matching each type
of impairment to British jobs in isolation, the Government should measure the functional profiles
required in different jobs – that is, all the different capacities in combination that someone needs to be
able to do that job. That way, the functional profile of the claimant can be matched to the functional
profile that jobs require.

3. Make sure that the assumptions that the system makes about employers match the legal
requirements placed on employers. It would be possible for assessors to consider whether reasonable
adjustments would make a particular type of job possible for a particular claimant. The risk, however,
is that the resulting decisions over-estimate what most employers will do, and are therefore unfair. As
such, I recommend that the Government links any considerations of workplace adjustments to what is
currently legally required and enforced in practice. Indeed, given that the Government has been
classifying more people as ‘fit for work’, they should also impose more significant burdens on
employers to make the changes necessary for this group of people to have a real opportunity to work.

Paradigm Shift: from Impairment and Functional Limitation to the Disability Approach. Washington, DC: World Bank,
disability in Europe - similarities and differences. Strasbourg: Council of Europe.
Other issues on the accuracy of the WCA (and PIP assessments)

The Committee has asked a number of other questions about the assessment of impairments at the WCA and PIP assessments:

- Is Department of Work and Pensions quality control for contractors sufficient and effective?
- Do contractor assessors possess sufficient expertise to carry out assessments for people with a wide range of health conditions?
- Why do claimants seek to overturn initial assessment outcomes for ESA and/or PIP?
- What accounts for the rate of overturned decisions at appeal for PIP and/or ESA?
- What changes could be made earlier in the process to ensure fewer claimants feel they need to appeal?

In my project, I considered the way the WCA tries to assess the ‘genuineness’ of claimants – an unusual way of framing the issue, but one that I find useful as a contrast to the challenge of assessing the work capability of someone with a given set of impairments. I have found there are substantial concerns about each of the strategies WCA assessors use to assess ‘genuineness’. Not only was medical evidence often unavailable, but claimants’ treatment history only provides indirect evidence about their impairments. Assessors therefore combined this evidence with their wider medical knowledge to decide if the reported impairments were ‘likely’, potentially leading to unfair decisions where people’s impairments were unusual, or where there were other reasons for a lack of treatment. Similarly, informal observations at the assessment (such as how people walked) provide highly unreliable clues on fluctuating conditions. Many key actors felt that appeal tribunals made better decisions about ‘genuineness’, not because they had more written medical evidence, but because they asked better questions to the claimant and weighed the evidence they had more fairly.

In the aforementioned Demos report, I therefore argue that the Government should:

1. Ensure that assessors’ reports of what claimants said can unquestionably be trusted. A number of claimants have reported that PIP assessors have fabricated some or all of their reports, sometimes supported by strong evidence such as secret recordings. While this remains anecdotal, and may or may not apply equally to the WCA, it is clearly very damaging to the legitimacy of DWP disability assessments in general. The Government should therefore audio record all assessments, and annually review a sample of these to ensure that recording is accurate. The claimant should also be able to see – and comment on – the first part of the assessment report during the assessment.

2. Improve the supply of useful medical evidence into the WCA. Almost every report on the WCA since it was introduced has argued that the supply of medical evidence must be improved, but achieving this in practice has been slow and difficult. To further improve this, the Government should reverse the current burden on assessors to justify only where they do request further medical evidence, instead requiring them to justify where they do not. These requests for information should be light-touch, using a secure electronic system.

3. Improve the accuracy and transparency of any decisions that contradict claimants’ own description of their lives. A legitimate system cannot be based simply on whatever claimants claim, yet nor can a legitimate system simply ignore claimants’ own description on the basis of unreliable evidence. The public are broadly sympathetic to claimants here: legitimacy is a balancing act, but we can get a much better balance than the current WCA. The Government should inter alia:
• Require assessors to ask claimants if they have an explanation for any evidence that seemingly contradicts their description of their impairments, rather than jumping to a decision that the claimant is wrong;

• Set a high evidence threshold for over-ruling claimants’ own description of their lives, and ensure that assessors consistently apply it;

• Allow claimants to go through a process of treatment to obtain medical evidence on their condition, and then go through another WCA without delay.

January 2018
References


