

Deskilling: A New Discourse and Some New Evidence

Doug Fraser *

Abstract

This article provides a brief introduction to a research program which has been under way since 2007 to examine, using data from a large-sample panel survey, whether jobs in Australia are becoming more or less skilful over time. It redefines the debate on deskilling which ran through the 1970s and 1980s by expanding the focus beyond simple job quality to issues of current policy interest, notably the contribution of skill to innovation and productivity. To map this kind of dynamism it is necessary to use a metric capable of capturing change over short periods. This is achieved by adding a third dimension, skill-intensity, to Spenner's classic definition of skill in terms of two loosely related constructs, worker autonomy/control and substantive job complexity. Data drawn from the first eight waves of HILDA are analysed to demonstrate that this metric is capable of capturing statistically significant change in the average skill content of jobs over much shorter periods than was possible with the metrics used in earlier decades. The broadly parallel aggregate trendlines for skill-intensity and autonomy/control suggest that these two dimensions are linked in the way Spenner suggests, even though major discrepancies appear between them in some industries and occupations.

Keywords

Autonomy; control; deskilling; employee discretion; job quality; informal learning; innovation; skill-intensity; skill polarisation; skill trends; skill utilisation; substantive task complexity; workplace learning.

Introduction

This article offers a summary introduction to a larger research program which has been in progress for the last four years, to find a better way of mapping and explaining changes in the skill content of jobs (Fraser 2009). The purposes of this initial exposition are confined to explaining the rationale behind the research and the metric which has been developed to this end, and demonstrating with real data that this metric not only is capable of being operationalised for practical purposes, but has the potential to support findings of a kind which were not possible with the quantitative methods in use in earlier decades.

Most of the really credible research so far on this topic has been qualitative. Indeed, qualitative research is essential for such purposes as exposing the mecha-

* The University of New South Wales

nisms of change in job content, identifying how real people experience and talk about such changes, and in particular revealing the distinctive complexities of individual workplace contexts, thus safeguarding against any premature assumption of simple or universal causal rules. Conversely, only quantitative research can persuasively demonstrate whether the phenomena identified in individual case studies apply across the workforce, and if so, in what circumstances. And though statistical research is necessarily reductionist, the helicopter view it provides by synthesising the detail from a large number of cases can also reveal patterns and effects which might not be initially apparent even in a large number of qualitative studies, owing to the very distinctiveness and complexity of individual cases which permit qualitative research to generate such rich data.

Consequently, good research requires a combination of both approaches, preferably alternating in such a way that each can inform the next stage in the other. The reason for focusing on quantitative evidence at this point is that it has always been the weaker of the two traditions in the deskilling debate. Moreover, the twenty years since the original deskilling controversy ran out of steam have brought advances both in purpose-designed surveys and in statistical series of more general application which allow more precise and relevant analyses than were feasible in that earlier period, and thus justify a reassessment of what statistical analysis can contribute to the understanding of skill trajectories at the individual, workplace, industry and national levels.

Whereas the earlier research into the changing skill content of work sought to confirm or refute the hypothesis of an alleged secular trend under capitalism (Form 1987; Attewell 1987), the present project originated in innovation research. Its ultimate purpose was and remains to assess the contribution of skill to an industry's or a nation's potential to innovate successfully. This different emphasis has two important consequences for framing the research. Firstly, it moves the discussion away from its original contestatory focus into one which keys in smoothly with many of the most important issues in mainstream economic strategy. Secondly, it shifts the core phenomenon of interest away from a 'secular' (i.e. consistent, cumulative and long-term) historical trend toward the kinds of dynamism and system effects on which modern innovation theory focuses. Central to the approach introduced in this article is a way of defining and estimating the skill content of jobs which retains some continuity with earlier deskilling literature but sheds more light on the dynamic aspects of skill change.

A third way in which this project differs from the majority of work in the quantitative tradition, especially in Australia, is that it concentrates less on the supply of skills (specifically, credentialled skills) than on how skills are used in the workplace. While by no means new, this emphasis links the project into a tradition of research which goes back to the landmark case studies conducted by the UK National Institute for Economic and Social Research in the 1980s and 90s (Maglen and Hopkins 1998), and has achieved some prominence in Australia in the last decade through the Skill Ecosystem Program (Buchanan 2006; Windsor 2006; Payne 2007), but which has remained very much a minority strand in both mainstream labour economics and vocational education and training (VET) research.

The rationale behind this approach is that credentialled skills, as they emerge from the training system, represent only potential productivity: actual productivity is achieved through the purposive and in most cases selective application of skill at the workplace, under circumstances and practices which can either enhance or diminish its productive potential. Moreover, this tradition assumes that skill is almost inevitably modified in the process of its practical application. Thus a skill once created, whether through training or informal learning, does not become a permanent, stable asset of the economy. Instead it is developed, instrumentalised, updated, contextualised and individualised as, and to the extent that, it is used; and if is not used, it decays sooner or later.

The research so far has identified a range of findings which, albeit mostly tentative, are of considerable interest in their own right. However, in the author's judgement it will be more effective, and kinder on readers, to set these out one at a time in future publications once the general rationale and logic of the study are in the public domain. The many obvious lacunae in the summaries of the research which appear in this paper should be understood in that context.¹ Taking into consideration these modest aims, it is only fair to warn readers in advance that the limited and selective findings set out here are not in themselves of epoch-making significance, nor can they be treated as conclusive in any real sense. The short run of data so far available for analysis, together with the complete absence of prior Australian data to place them in a historical context, mean that they cannot yet be linked confidently to any secular trends of the kind which the earlier deskilling research sought to reveal, even assuming such trends exist. However, they are necessary to set a context for the more detailed findings to be set out in later publications, and should be sufficient to show that provided this or comparable data collections can be maintained for a period of twenty years or longer, the approach has the potential to provide conclusive and meaningful answers to questions which were of burning interest in the earlier debate, but always just eluded resolution by statistical evidence.

Background: In Search of a Metric

The deskilling debate that began in the mid-1970s was driven by a large body of case study research, much of it of high quality, which documented deskilling in individual occupations or workplaces and gave useful insights into the mechanisms by which it occurred. Its weakness, which ultimately meant that it carried only limited conviction with those not disposed in advance to accept the hypothesis, was that by itself it could never prove that these cases were part of a dominant trend (Attewell 1987: 325, 337; Spenner 1993: 825–826). The attempts that were made to support or refute the hypothesis with quantitative evidence had little in the way of good, relevant data to go on, and hence relied on proxy indicators which were seldom consistent across studies, generally very indirect, and often verging on the eccentric. However, the more fundamental difficulty was that the two sides of the debate meant different things by skill.

In effect, the cases for and against deskilling involved not just different semantics but different *discourses* (Spenner 1990: 400–401; Vallas 1990: 380–383). The argument for deskilling saw the issue primarily in terms of a decline in *quality*

of work. Braverman's original thesis defined loss of skill effectively as the loss of workers' control over their work (Braverman 1974), while much of the later work in this tradition drew on Kohn and Schooler's theories of how work could fulfil psychological needs (Spenner 1988). The argument for increasing skill content was generally based on an increase in the number and/or level of competences involved in a given job — in essence, a growth in the *knowledge content* of jobs. The two discourses proceeded in parallel, overlapping substantially in reference but seldom directly engaging each other. Any convergence would have required agreement on a common definition of skill capable of supporting objective measurement. Arguably, the ultimate failure of the parties to agree on such a common definition was the main thing, other than simple exhaustion, which brought the debate to an end sometime around 1990.

One scholar who consistently sought such an evidence-based middle ground was Kenneth Spenner. By the early 1980s he had resorted to what he himself called a 'hypothesis and pragmatic approach' (Spenner 1985: 402). He argued that skill should be recognised as a multidimensional construct. Since each of the two discourses made sense in its own right, and each could claim a strong evidentiary basis, any combined metric for the skill content of jobs must take both into account. Thus, skill needed to be measured in at least two dimensions: the *substantive complexity* of the work involved, and the *autonomy/control* exercised by those who performed it.

Spenner never argued that these two dimensions exhausted the definition of skill, simply that they were the minimum required in any context for satisfactory measurement of the skill content of a job. He did not fully theorise the association between them, and by 1990, in the face of objections (e.g. Form 1987: 31), was declaring himself agnostic on whether autonomy/control needed to be defined as part of skill, just so long as it was treated as a relevant matter (Spenner 1990: 404). Perhaps understandably, therefore, he himself offered no suggestions as to how the two dimensions might be combined into a common scalar metric, and nobody since appears to have tried to develop one. Despite this, Spenner's ad-hoc definition has been generally accepted since as the classic approach to identifying changes in work skill.

One reason for the continuing acceptance of Spenner's approach to skill has been its practical success as the structuring principle behind the one program of purpose-designed mass surveys so far implemented for the specific purpose of tracking skill over time across a national economy through common but specific indicators. In the UK, five loosely coordinated surveys have so far been conducted between 1986 and 2006 at roughly 4–5 year intervals (Felstead, Gallie and Green 2002; Felstead, Gallie, Green and Zhou 2007), using large and representative national samples. This series now offers a clear picture of the skill trajectory of the UK economy, in growing detail with each new survey that goes into the field, over the two decades since the deskilling debate was a live one. Over that time it has provided persuasive answers to several of the major questions that remained unanswered in the original debate, and thus offers a model of how better-focused survey data can act as a bridge between qualitative

research and the kinds of statistical found object on which earlier quantitative research had to fall back.

Spenner's model is the starting-point for the new metric which was designed for the purposes of this research. Like the original, this expanded metric embraces the dimensions of *substantive complexity*, covering 'the level, scope and integration of mental, interpersonal, and manipulative tasks in a job' (Spenner 1983: 829) and *autonomy/control*, defined as 'the discretion or leeway available in a job to control the content, manner and speed with which tasks are done' (Spenner 1985: 135, 1990: 402–403). However, to pick up a dynamic aspect of skill change which is not fully captured by the original model, the new metric adds a third dimension, referred to as *skill-intensity*.

Skill-intensity describes the relationship between the skill demands of the job and the skill base of those workers who must perform it; in other words, the extent to which a job exercises, challenges and develops the skills of the person doing it. It applies independently of whether the job requirement or the employee's skill base is high or low in its own right. It is made up of three complementary sub-dimensions: *match*, *stretch* and *learning*.

- Match refers to the extent to which the worker uses her existing skills in her work. Leaving aside its implications for work quality, this question has clear economic implications because a skill which is not being used represents a dead investment, and because skills will decay, or at best lose their currency and relevance, if they remain unused for any length of time. Mavromaras, McGuinness and Fok (2007), using data from the first five waves of HILDA, estimated the productivity loss due to the former component alone across the Australian economy as equivalent to between 2.6 per cent and 5.2 per cent of Gross Domestic Product.²
- Stretch means a gap between the worker's existing skills and the demands of the job which challenges him to extend his skill base in order to do the job better or keep up with changes in its requirements. Too much stretch (i.e. underskilling) can simply mean the reverse of overskilling, in that productivity is lost because workers lack the skills to do their job effectively. However, a small margin of stretch is desirable to encourage both the job and the worker's skill base to evolve and adapt to new demands and opportunities.
- Learning complements stretch by making adaptation possible. As used here, the term embraces both the extent to which the worker wants and/or needs to learn new skills or refine her existing ones in order to do the job properly or better, and the extent to which the job provides adequate opportunities for that learning, either through training or through the knowledge acquired, individually or collectively, in the work process.

It is the combination rather than the presence of an individual element that makes a job skill-intensive. A good match between the worker's skill base and the demands of the job, without the dynamic element of stretch, can lead to complacency, path-dependence and ultimately competence traps. Stretch is productive only if the workforce has the capacity and opportunity for learning to

close up the skill gaps as they emerge. Learning is worthwhile only to the extent that the job takes productive advantage of it within a reasonable time. Thus an adequate indicator of skill-intensity needs to be a composite one, made up from multiple variables that capture the different aspects.

It should be clear by this point that skill-intensity is not a replacement, substitute or proxy for substantive complexity. It is a third dimension in its own right, complementing the other two. Whereas substantive complexity describes characteristics which attach to the job, more or less irrespective of who is doing it, skill-intensity describes a relationship between the worker and the job which is at least partly specific to each worker-job pair, and consequently captures more of the variety which actually exists in the labour market. Perhaps more importantly, while the substantive knowledge content of jobs will generally evolve or decline over relatively extended periods (except where sudden technological advances or new fashions in management make a whole set of skills obsolete), the two-way nature of the skill-intensity relationship opens the way for constant adjustment and adaptation, both in practice and in expectations (since unlike substantive complexity, it is in part subjective). It should therefore show meaningful change over much shorter periods, although the meaning will be different from that of a change in substantive complexity.

As a basis for research, the skill-intensity construct has a number of methodological advantages:

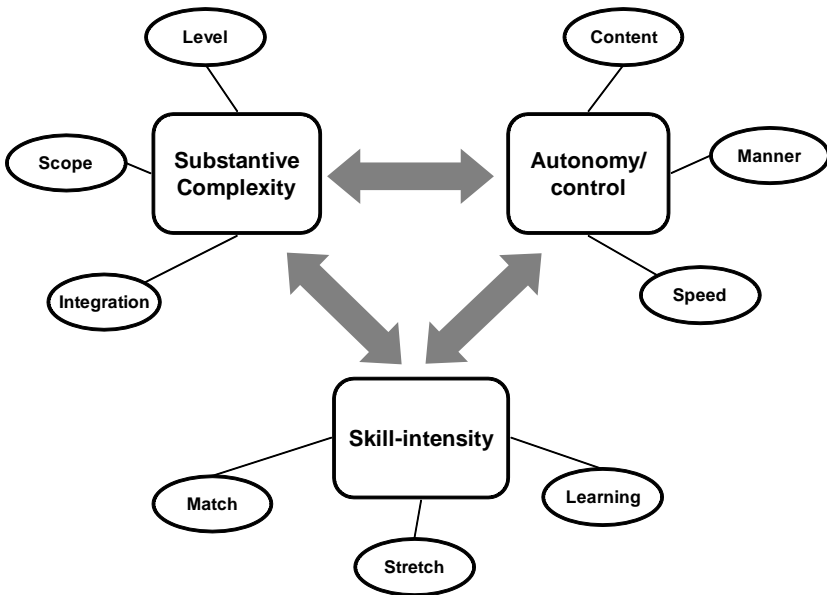
- It effectively captures the dynamism in the skilling system by treating the skill content of a job, or the skill base of an individual, not as a fixed quantity existing in its own right but as something that grows, decays and changes constantly in line with the demands of the job, the informal learning that takes place on the job, shifts in organisational culture, the level of motivation of the worker and the management philosophy and work practices that prevail in the workplace;
- Because it focuses on the match between the individual's expectations and perceptions of his own job, it permits valid measurement by self-report without the need for any third party to mediate between report and analysis, e.g. by applying external criteria or standardised categories. Thus it lends itself to capturing the full variety of individual experiences;
- It measures skill (or at least, this aspect of it) at the point where it is actually exercised, i.e. in an individual job, in a specific work context, and at the moment of its productive application;
- Because it is 'blind' to the actual technical or knowledge content of the job, it represents a neutral measure that can be used without bias across the full occupational hierarchy and all types of work.

There are, of course, offsetting disadvantages. By taking as given the skill base which each worker brings to her job, data on skill-intensity taken in isolation provide no information on the comparative skilfulness of the workforces employed in different occupations or locations and no indication of whether the actual complexity, knowledge content or requirement for relevant aptitudes or behaviours (dexterity, concentration, interpersonal skill) is greater in one type

of job than in another. By itself, therefore, it has little value for point-in-time comparisons between industries or occupations, let alone for the sort of international comparison required to categorise different economies as high-, low- or medium-skilled. Its real value lies in providing a sensitive basis for comparing the amount and direction of *change* occurring over a given time period in different industries or nations, or at different levels in the workforce hierarchy.

It follows that data on skill-intensity, though representing valuable information in their own right, do not provide enough evidence by themselves to say with confidence that an industry, occupation or economy is high- or low-skilled. Before any such claim can be made, or refuted, it is necessary to have data on at least one and preferably both of Spenner's other dimensions. The complete metric, represented graphically in Figure 1, involves (at the minimum) all three components of each of the three dimensions.

Figure 1: Dimensions and sub-dimensions of skill



However, the three dimensions overlap sufficiently to permit a reasonably accurate estimate of the situation or trend even if information is available on only two of them. For instance, any increase in substantive job complexity should logically create more situations where there is a range of choices for action, the consequences of a wrong choice are critical and hard to reverse, and so many factors influence the correct choice that only someone on the spot has enough knowledge of the circumstantial factors to make a correct decision. Such jobs cannot be exercised effectively unless the jobholders, individually or collectively, have a high level of input to the relevant decisions. Thus autonomy/control can often be taken as evidence of complexity, since highly complex jobs could not be performed unless they incorporated high levels of autonomy/control. In the

same way, the element of continuous learning in the definition of skill-intensity overlaps and may well co-vary in some cases with one of the best recognised indicators of substantive complexity, the amount of learning required to do a job properly (Felstead, Gallie and Green 2002: 27).

Methodology

To demonstrate the usefulness of the new metric it should be sufficient, as a beginning, to show that it can shed new light on the major questions which preoccupied the participants in the original deskilling debate, but which were never satisfactorily resolved at the time. So while the research project as a whole has moved into more detailed investigations which are possibly of more current practical interest, that part of the research to be covered by this introductory exposition addresses four main questions drawn from the earlier debate:

1. Are jobs in Australia, on average, becoming more or less skilful over time?
2. Assuming such change is occurring, does it reflect a generic change in the nature of work in Australia, or simply structural change resulting in a different proportional representation of intrinsically high- and low-skilled industries and occupations (i.e. a compositional effect)?
3. Is there a trend for jobs to concentrate at the high and low ends of the skill spectrum, with a 'vanishing middle' (i.e. a polarisation effect)?
4. Do skill-intensity and autonomy/control co-vary in a way that suggests they make up interdependent parts of a broader construct of skill?

A general-purpose social survey was chosen as the data source for this project, primarily for practical reasons: data were readily available and of known high quality, it used a well-designed representative national sample far bigger than could be achieved for a purpose-designed survey with the resources to hand, and five waves of back data were available for analysis right from the start of the research. Practical considerations aside, a comprehensive survey of this kind offers advantages for this type of research which would not easily be reproduced in a specially designed instrument, since the broad scope of the questionnaire creates the possibility to control the findings for a wide range of individual characteristics (age, gender, personality type, mental and physical health, previous employment experience, family circumstances, recent life events) and socio-economic factors which might confound the response. Even without such controls, a large, well-balanced sample may also reduce the likelihood that local or adventitious factors like these will significantly bias the aggregate results.

HILDA (Household, Income and Labour Dynamics in Australia)³ is a panel survey commissioned by the former Commonwealth Department of Social Security, now known as the Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA), and designed and managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). It has been conducted annually since 2001, and eight waves of data are now available for analysis, with funding guaranteed for a further four through to 2012. The sample covers around 12–13,000 households and 20,000 individuals, of whom

between 12,000 and 14,000 have provided individual-level data in each wave. Of these, around 7–8,000 are employed at the time of each survey and hence eligible to answer the skill-related questions.

HILDA offers most of the advantages listed in the previous paragraph. For example, it includes extensive data on income and a large range of variables on health conditions which might be linked to perceived work quality, in either direction, but which it would not normally be feasible to collect in a specialised survey on work skills, while the questionnaires for individual waves contain special sections on such topics as personality types and retirement intentions which lend themselves to cross-analysis with the data on skill, if only on a cross-sectional basis. Both the sample and the questionnaire have been designed to high professional standards and the data are subject to intensive quality control, with errors detected in individual waves being retrospectively corrected in subsequent releases. As a panel survey it offers the crucial advantage of permitting cohorts and individuals to be tracked over time, making it possible at least in principle to answer such questions as whether individuals' learning in a job is confined to the first year or two or continues over longer periods, and whether their expectations of skill-job match adjust over time to the reality as they experience it.

Conversely, it avoids many of the drawbacks one might expect from a general-purpose survey, notably the risk that topics peripheral to the primary interests of the survey owner will be covered cursorily or without a good knowledge of the relevant research. While it is true that a good third of the more than 3,000 questions in each wave are unlikely to interest anyone but a social security administrator, the sequences on other topics show evidence of having been designed with independent researchers in mind, and have clearly been developed on the basis of an expert familiarity with the literature on the respective topic. Thus the skill-related variables, though relatively few in number, all owe their wording to well-regarded specialist predecessor instruments in Australia or overseas, and though there is nothing in the documentation to indicate that they were consciously intended or tested to make up composite scales, in practice the scales which fall most intuitively out of each set of variables show better reliability statistics than many purpose-designed instruments in the literature.

For all their excellent quality, the HILDA data have two major limitations. One is that precisely because HILDA is the first survey of its type in Australia to provide data of this quality on skill use in the workplace, there are no other reliable sources with which they can be compared to set them in the context of what happened before 2001. The other, which affects only this specific research task, is that none of the original set of questions precisely covers any aspect of Spenner's substantive complexity dimension. Thus it has been possible so far to test only a partial version of the metric which cannot be compared directly against Spenner's original, and by this definition equally partial, metric. This limitation must be kept in mind when comparing the results that follow with those of earlier research.

The analysis is based on a set of variables in the HILDA self-completion questionnaire, forming part of a longer sequence in which respondents are asked to rate their agreement with statements about aspects of their main job on a

7-point Likert response scale. The core set, asked over all eight waves, consists of six variables. The variable names listed below are convenience titles adopted for the purpose of these analyses:

- COMPLEX — My job is complex and difficult
- NUSKILLS — My job often requires me to learn new skills
- USESKILL — I use many of my skills and abilities in my current job
- OWNTASK — I have a lot of freedom to decide how I do my job
- HAVESAY — I have a lot of say about what happens in my job
- WORKFLOW — I have a lot of freedom to decide when I do my work

It would seem intuitively obvious to treat the first three of these variables as indicators or elements of skill-intensity, and the last three as indicators of autonomy/control. Principal component analysis conducted on the full sequence of twelve job-quality questions for Waves 1–6 confirmed that the two sets make up reasonably valid factors, each with an eigenvalue exceeding 1 except for skill-intensity in Waves 3 and 4 (Fraser 2009: 148).⁴ The three items in each set were combined into simple additive 21-point scales, which achieved Cronbach's alpha values over the first six waves of between 0.701 and 0.727 for skill-intensity and 0.807 to 0.827 for autonomy/control (Fraser 2009: 151). These figures are not especially strong, but can be considered acceptable for 3-point scales, and compare more than reasonably with comparable scales used in well-regarded overseas studies.⁵

From Wave 5 onwards the sequence has included a supplementary set of nine questions providing specific detail on Spenner's three sub-dimensions of autonomy/control, together with new data on task variety and repetitiveness (which are relevant to substantive complexity), and on work intensification. This extra detail has made it possible to construct extended scales for both primary dimensions, along with new sub-scales covering autonomy, control over job content and control over the organisation of working time. These new scales all show much greater reliability, with Cronbach's alpha running around 0.9 even on the 3-item scales, but are not tracked in this article because so far they cover only four waves of data.

Besides tracking movements in the mean score for all respondents on each scale over successive waves, the analysis involved tracking the mean scores for each occupation and industry, disaggregated at both 1- and 2-digit levels, using the 2006 ANZSCO and ANZSIC classifications. This provided a basis for determining whether the relative skilfulness of each occupation or industry had changed over the period. In addition, the percentage of the employed sample represented by each industry and occupation was calculated at the beginning and end of the period and at an approximately halfway point (Wave 4 for oc-

cupation and Wave 5 for industry). Together these measurements provide the means of establishing whether any change in the all-respondents mean was the result of substantial shifts of employment within the sample towards either areas or types of work demanding lower or higher skill.

To test the polarisation hypothesis, the sample was divided into respondents working in occupations and industries that recorded high, low and medium mean scores on each scale in Wave 1, and the analysis was repeated for Wave 8. As a safeguard to determine whether these end readings represented a consistent trend, a third analysis was once again undertaken at the same halfway points. The analysis was then repeated for individuals over the full sample for each of the eight waves. The high, low and intermediate categories were defined with reference to the standard deviation for each year to ensure that they reflected the actual range and pattern of variation for the year in question. Two alternative cut-off points were set, one at half a standard deviation either side of the mean, and the other at a quarter. The assumption is that any sustained trend towards polarisation over this period will be reflected in a progressive shrinking of the central band (within a quarter or half, respectively, of the standard deviation on either side of the mean) over the eight waves.

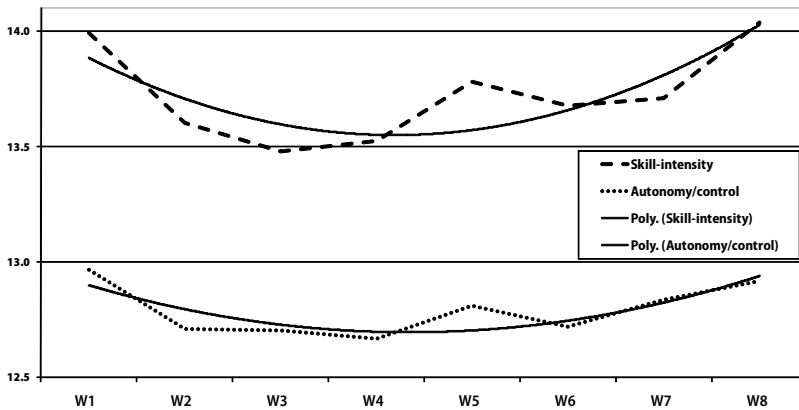
Findings

Before investigating the findings that relate specifically to the four test hypotheses, it should be noted that the component variables in each scale show different patterns across the eight waves. The highest-scoring variable is USEKILL, with a median score of 6 (i.e. clear agreement) in all waves, suggesting that the great majority of Australian workers do not consider themselves to be overskilled for their current jobs. NUSKILLS and OWNTASK recorded median scores of 5 (weak agreement) across all waves, as did HAVESAY except in Wave 4, when it dropped to 4 (neutral). COMPLEX was the lowest-scoring of the skill-intensity variables, recording a neutral median score over the whole period, while OWNTASK was the only one with a median indicating weak disagreement (3) over the first four waves, albeit rising to 4 over Waves 5–8. These last two variables were the only ones whose mean score had moved notably above the Wave 1 level by the end of the period.

Hypothesis 1: Did the Average Level of Skill Rise or Fall?

The trend on this measure is ambiguous. Results for the first eight waves (2001–2008) indicate that across the full employed sample, the mean composite scores for both skill-intensity and autonomy/control declined over the first half of the period, but had recovered to around 2001 levels by 2008. Figure 2 includes polynomial trendlines which show remarkably similar shallow U-shaped patterns for both scales, albeit that for autonomy/control is considerably less pronounced.

Figure 2: Means, composite scales
All employed respondents, Waves 1–8



These results must be treated with some caution for at least two reasons:

- i) The actual range of year-on-year movement is quite small, especially for autonomy/control, making it difficult to separate real movement in the underlying constructs from chance variation;
- ii) The actual curves (as distinct from the trendlines) in Figure 3 show that the movement is uneven, being dominated on both scales by a relatively steep drop between the first and second waves and an almost equally sharp spike in Wave 5, together with another rise between Waves 7 and 8 for skill-intensity. It is only prudent to suspect that one or more of these could have reflected some kind of response effect, affecting respondents' scores across the full relevant set of questions or even the whole questionnaire, rather than an actual change in perceptions.

The first issue is the more easily resolved by statistical testing. Since the major scales were designed to be treated as interval data, it was possible to apply paired-sample *t*-tests between each pair of waves to determine where the change was statistically significant. Significant change at the 0.01 level was found on the skill-intensity scale between Waves 1 and 2, 1 and 4, 4 and 5, and 7 and 8, while the autonomy/control scale showed significant change between Waves 1 and 2, and between 4 and 8. Together these figures at least encourage confidence that the dip in both trendlines in the middle years is more than just the result of chance variation.⁶

Issue (ii) is harder to resolve. The dip between Waves 1 and 2 could well be at least partly evidence of panel conditioning — i.e. the higher scores in Wave 1 might have occurred simply because respondents were seeing the questions for the first time. On the other hand, the new questions added to the same sequence in Wave 5 showed no such dip over the three waves subsequent to their introduction. The behaviour of the core variables also suggests something more than just a one-off artefact affecting responses across the board, as three out of the six continued to record declining mean scores beyond Wave 2. The Wave

5 spike is if anything harder to explain as a response effect. Since sample loss was especially marked in these two waves (Wooden and Watson 2007: 214), it is tempting to attribute the movement to this cause.

To test whether these movements were in fact generic artefacts, independent of the content of each question, five control variables present in all waves were subjected to the same tests as the main variables of interest (Fraser 2009: 175-176). They related to overall job satisfaction, perceived job security (two questions), intention to leave current job and fairness of remuneration; three were drawn from the same sequence in the self-completion questionnaire as the main variables of interest, and two from the interview questionnaire. All five showed noticeably different patterns over Waves 1-6 from the variables of primary interest, with none recording statistically significant variation between Waves 1 and 2 or between 4 and 5 on a repeated measures ANOVA.

None of this evidence conclusively rules out the presence of an artefact, especially after Wave 1. But taken together, it makes a persuasive case that such an artefact was not generic (i.e. if it existed, it specifically affected this set of six questions) and cannot account for all the movement. In particular, it weakens the case for attributing the movement in Waves 2 and 5 to sample loss alone. Once again, therefore, it justifies reasonable confidence that the shape of the trendlines in Figure 3 is real, and hence that for whatever reason, the representative worker's experience in both dimensions deteriorated after 2001, but returned to earlier levels by 2008.

Hypothesis 2: Was the Movement in Mean Scores Due to Compositional Change?

The high level of job mobility over the seven years suggests that most or all of this change might simply be the result of individual members of the sample moving from lower- to higher-skilled industries or occupations, or vice-versa. Population-weighted estimates prepared for Waves 1-7 by MIAESR indicate that sample members changed jobs, on average, every six years, with around 60 per cent of job changes involving a change of occupation and almost two thirds a change of industry (Wilkins, Warren, Hahn and Hough 2010: 61). Furthermore, the movement in mean scores within individual industries or occupations at the 2-digit level appears too small to explain the full change in the overall average for either dimension, since significant shifts between Waves 1 and 8 are confined to industries and occupations with quite small cell sizes. The polynomial trendlines for all the major (1-digit) occupational categories over the full eight waves are remarkably parallel; the only notable convergence appears between the two lowest-scoring categories, labourers and sales workers.

Yet the observed change in the composition of employment equally fails to explain all the movement in the aggregate means. The twelve 2-digit occupations which recorded the highest mean skill-intensity scores in Wave 8 increased their combined representation over the eight waves from 34.8 per cent of the sample to 37.4 per cent, while the proportion represented by the ten lowest-scoring declined marginally from 22.8 per cent to 21.1 per cent. This shift might have been expected by itself to result in at least a small increase in the aggregate mean

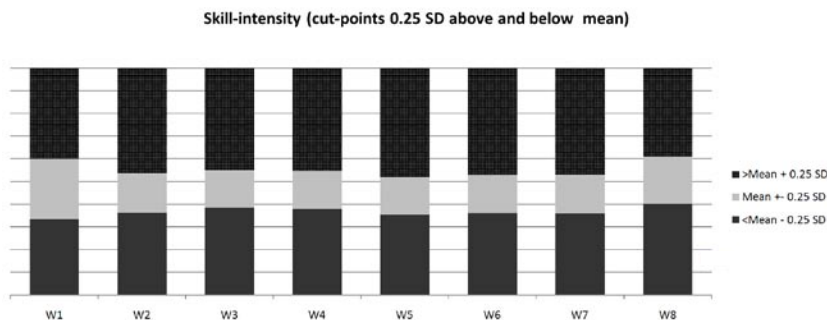
scores by Wave 8, but the actual means recorded were not significantly different from those for Wave 1.

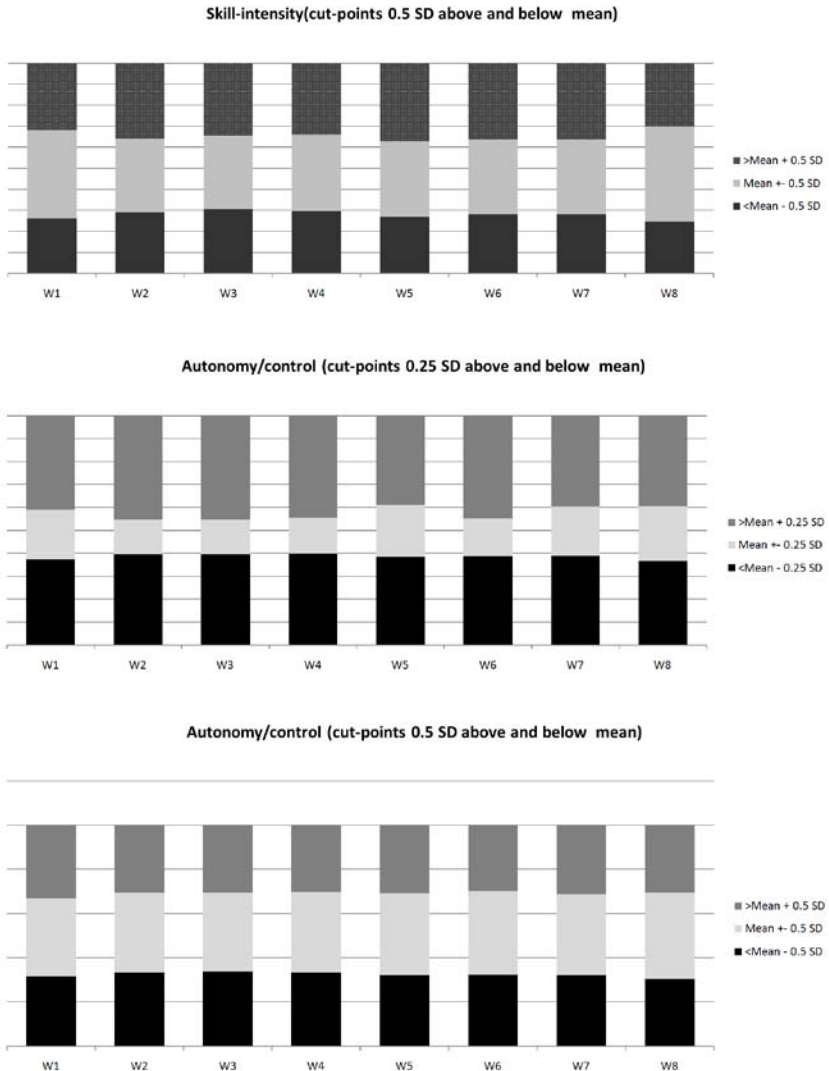
Further evidence against a strong compositional influence on individual outcomes over the first six waves can be found in two alternative regression analyses carried out by Fraser (2009: 183, 190–193). The first pair of models, hierarchical linear regressions using as their dependent variables the change in individual score between Waves 1 and 6 on the skill-intensity and autonomy/control scales respectively, showed that the combined impact of changes in the individual respondent's industry of employment, occupation, sector of employment, contract of employment and score on the ANU4 Occupational Status Scale, at any time over the five years, contributed only 3 per cent of the total variance explained by the model for skill-intensity, and only 0.7 per cent for autonomy/control. In the second pair of models, logistic regressions using the same predictor variables with direction of change on each scale as the dependent variable, only changes in employment sector and occupational status made a statistically significant contribution to direction of change for skill-intensity, and only change in occupational status for autonomy/control. In both sets of models the effect sizes for the compositional variables were too small to be regarded as ecologically significant.

Hypothesis 3: Did Any Skill Polarisation Occur?

Once again, no conclusive evidence was found to support this hypothesis over this admittedly short period. The size of the central band for skill-intensity diminished notably after Wave 1, but returned to something like its original size by Wave 8. Figure 3 shows an inconsistent pattern across waves, which is broadly the same regardless of which cutoff point is used. In years when the balance of the sample shifts towards the ends of the scale, the gain appears to be shared almost evenly between the high and low bands, and the converse applies when the distribution moves back towards the middle. For autonomy/control the overall pattern is similar but shows much less movement, especially using the 0.5 Standard Deviation cutoff.

Figure 3: Distribution of scores, all employed respondents, Waves 1–8





So far as anything can be deduced from these patterns, the thing that stands out is the parallelism between movements in the distribution of scores across the central and outer bands and movements in the overall mean. Years with high aggregate mean scores are characterised by a more centralised distribution, especially in the case of autonomy/control where the spike in mean scores in Wave 5 is accompanied by a one-off increase in the size of the central band.

Clearer shifts emerge between Waves 1 and 8 when the analysis is disaggregated by industry and occupation, but the trends are not consistent. The analysis by occupation showed both growing polarisation and an upward shift in the overall distribution for skill-intensity. The same shift towards the upper band was evident for autonomy/control, but the evidence on polarisation runs in opposite

directions depending on the cutoff point chosen. The only significant change to emerge from the analysis by industry was a strong growth in polarisation for autonomy/control. In all cases, however, the distribution was more polarised at the halfway point, suggesting that these changes should not necessarily be interpreted as sustained trends without further analysis.

Hypothesis 4: Do the Results Suggest that Skill-Intensity and Autonomy/Control are Linked?

On the surface, the broadly parallel trendlines for both scales seem to indicate that the two constructs move in tandem: when the mean for skill-intensity moves up, so does that for autonomy/control, and vice-versa. The finding, noted earlier, that the five control variables (all of which involved different aspects of the respondents' perception of job quality) did not share the same pattern suggests that this is a characteristic specific to these constructs and not simply a reflection of some generic change in the response from year to year (e.g. a higher or lower average level of respondent optimism). Together these results add credence to the hypothesis that the two are linked aspects of a single broader construct, and hence that it might be possible, with considerable further work, to combine them (along with substantive complexity) into a single informative scale or index.

On the other hand, the correlations between the two scales for individual respondents in individual waves are relatively weak at a little over 0.2, though this rises to around 0.3 when the more sensitive scales available from Wave 5 onwards are used. While this seems unimpressive compared with the figures of 0.5–0.7 quoted by Spenner (1990: 403) for the observed correlations between substantive complexity and autonomy/control in earlier studies, it should be remembered that nobody has previously investigated the correlation for skill-intensity, and there is no reason to believe that it will relate to autonomy/control in the same way, or with the same strength. (It should also be noted that the studies Spenner had in mind, though not referenced in his article and no longer identifiable, are unlikely to have had access to data of anything like the same quality or reliability as HILDA.)

More interesting results emerge when the statistics are disaggregated. The mean correlation is around twice as strong for respondents who work in private business as for those whose employer is a government agency (an average ratio of 2.12:1 across the first six waves, though fluctuating between extremes of 3.18:1 and 1.58:1 in individual waves). Of 80 2-digit industries in Wave 8, only 28 have mean skill-intensity scores which rank within ten places of their ranking on autonomy/control, while for 15 industries the ranks differ by forty or more places. For occupations in the same wave, the rankings are understandably somewhat closer, with 22 out of 42 having mean skill-intensity scores within five places in the rank order either side of their mean autonomy/control score. Even here, however, some striking discrepancies appear, with eight occupations showing gaps of fifteen places or more between their ranks on the two scales. By far the strongest negative discrepancies in rank between autonomy/control and skill-intensity are found in occupations dominated by the public sector: education professionals, health professionals and protective service workers.

Similarly, two characteristically public-sector industries, preschool and school education and hospitals, rank over 60 places on autonomy/control below their ranking on skill-intensity.

Discussion

The main findings show that the combined metric, even without data on substantive complexity, is capable of picking up statistically significant variation in the aggregate means for the two main constructs of interest at reasonably frequent intervals — roughly every 4–5 years, though the movement is not uniform. None of the earlier studies summarised by Spenner in his three review articles was able to identify aggregate mean change over a period anywhere near as short as this, though the UK surveys have exposed a broadly similar pace of change at least up to 2001 (Felstead, Gallie and Green 2002: 42–44).

It must however be borne in mind that statistical significance indicates only that the variation is not the result of chance. What remains unanswered, and what the data cannot yet tell us by themselves, is whether this source of variation is important enough, in the context of all the other things that influence such outcomes, to be worth worrying about: i.e. whether these results are *ecologically* significant.

In some ways it would have been easier to argue an ecological significance for the headline trends when there were only four or five waves of data and the trend appeared to be both unidirectional and counter-intuitive in the context of a growing industry and political concern over skill shortages. Now that the trend for both scales has shown an equally puzzling upturn and returned (in the aggregate, though not in the finer detail) to somewhere like where it was in 2001, it makes more sense to view what we are seeing as a fluctuation or a cyclical phenomenon. However, it is hard to read much into even that interpretation as long as the run of data remains so relatively short and we have no useful information on the longer-term trend to provide a context for it. It seems reasonable to assume that even within a long-term upward or downward trend in the overall skilfulness of work, there would be a normal range of variation leading to a certain amount of fluctuation or counter-trends from year to year without affecting the overall direction of history. But while we remain ignorant about what that normal range is, there is no way of telling whether the movement described here lies within or outside it, and hence whether it is worth taking notice of.

If it is hard to explain the significance of the movement, it is even harder to explain the reasons behind it. Certainly it is difficult to identify any major event over this period which would logically have triggered either the initial decline in average skilfulness or the subsequent recovery. Most of the factors one would normally associate with either deskilling or upskilling — structural change in response to global market exposure, the declining role of high-value-added manufacturing, the move to short-term labour contracts, growing education levels, private-sector underinvestment in training, the increasing pervasiveness of ICT in all forms of work, the relentless rollback of the rights of both organised labour and the individual employee — had been in progress for at least a decade, and in some cases nearly a quarter of a century, by the time HILDA first went

into the field, and went through no obvious major acceleration or check over these seven years. The most probable explanation is that the observed movement was in some way connected to the business cycle. That hypothesis at least will be testable in the near future, as the next two waves of data to be released will reveal the immediate impact of the global financial crisis.

There is no real alternative but to remain agnostic for a few more years on this question of significance, and hence of whether a long-term trend exists. As more waves of data accumulate, the point will eventually be reached where data are available for a complete business cycle and it becomes clearer what the normal range of variation is. The UK surveys have provided enough information in twenty years to reach reasonable certainty about which part of the observed movement is long-term trend and which is chance or cyclical fluctuation. Provided either HILDA can be continued that long (which will require funding for a further eight waves, in a fiscal climate unsympathetic to such investments in evidence), or substitute series can emerge to pick up the baton, many of the uncertainties that attend the current findings are likely to be resolved conclusively. There might also be potential to address some of the same issues through the analysis of data from existing surveys in other countries — in particular, general social panel surveys comparable to HILDA — which have been running longer and thus provide a longer-term historical perspective as well as an international comparative one.

In the meantime, it needs to be remembered that simply because a movement is small, within the normal range of variation or even contrary to the historical trend, it does not necessarily lack historical significance. From a systems perspective, which underlies today's mainstream innovation theory (Edquist 2005), most epoch-making social changes originate in first moves or initial conditions which are inconsiderable and easily overlooked in their own right but for some reason turn into cascades, and phenomena like asynchronous feedback cycles can easily disguise emerging trends in a confusion of out-of-synch or counteracting short-term outcomes. For reasons already mentioned, statistical research may do better than qualitative research in revealing such emerging trends at the point where they are still only small, indistinct but at least non-random movements. Whilst their historical significance may only become apparent in retrospect, being able to trace them back to their early stages can lead to a better understanding of how such emergent processes occur in other areas. It is in this context that the more technical concern of statistical significance assumes its real importance.

Pending better resolution of these headline issues, the main surprises to emerge from the research so far involve what might have been regarded as side-issues. The discovery that polarisation is not just a long-term cumulative phenomenon, but something which grows and contracts over cycles of only a couple of years, suggests that further analysis is needed to establish the mechanisms behind it, its apparent link to the overall average skilfulness of jobs, and the extent to which individuals may remain stuck in jobs at either end of the skill spectrum even when the overall trend goes into reverse. In the light of such analyses its significance could well turn out to be quite different from what was previously assumed.

Similarly the lack of strong evidence that compositional change, as conventionally understood, is influencing the overall trend in skilfulness, even in times of high job turnover, raises the question of whether we should be looking for different kinds of compositional effect. One obvious possibility is that job change across industries or occupations, or the return to employment of workers who have spent some time out of work, raises the overall level of learning simply because the first 3–12 months in an unfamiliar job necessarily involve more learning even if the job itself requires few special skills and the skill requirement does not subsequently evolve. Thus years of high job growth or intense economic restructuring may show higher averages on this incomplete metric simply because of the higher proportion of employed respondents who are new to their jobs. Fortunately the use of panel data makes it relatively straightforward to test such hypotheses, and their clarification is high on the agenda for the next phase of the research.

But perhaps the most interesting evidence concerns the exact nature of the relationship between skill-intensity and autonomy/control. The observed correlations between the two scales, though not compelling, are strong enough (particularly for the more sensitive extended scales) to suggest the existence of some logical link, even if that link is little evident in some occupational categories. That said, it must be recalled that Spenner himself never suggested his two dimensions would coincide in every case; in fact he found evidence that autonomy/control had been declining in the aggregate over the same period when substantive complexity was on the increase (Spenner 1983: 834), a finding replicated with greater empirical force by the UK surveys up to 2001 (Felstead, Gallie and Green 2002: 68). So the anomalies in the results for Australia have good precedents.

One way of looking at the discrepancies is that there is a 'natural' or optimal relationship between the two, a hypothesis which has received ample though uneven empirical support from several decades of research in the wake of Karasek (1979) who used broadly similar constructs,⁷ but that the two fall out of mesh in some parts of the labour market. This then raises the question whether such cases are simply contingent — the result of specific characteristics of the work and the circumstances in which it needs to be carried out — or whether they represent system failures implying a productivity problem which can and should be addressed. Given the evidence that such anomalies — specifically, those where autonomy/control falls well below skill-intensity — are most common in highly skilled professions which are over-represented in the public sector, there would appear to be a good case for investigating the latter hypothesis.

Conclusions

As was noted at the beginning, the headline findings from the research so far are neither especially exciting nor especially informative in their own right. The project is still at that point where it opens up more questions than it answers, and indeed where most of its value lies in defining new questions which warrant further investigation. As will be clear from the preceding discussion, there are several more years of analysis ahead to make full sense even of the data

already available, and each new wave of data that is released changes the story significantly, while adding new clarity to results which have so far been tentative or speculative. Without going any further into detail of the analyses still to be undertaken, three main directions can be signalled for future research:

- The full information potential of panel data needs to be exploited by moving beyond cross-sectional analyses into genuine longitudinal research tracking individuals and cohorts across waves. This will require a shift from the standard inferential methods used so far into newer and more sophisticated modelling techniques specifically designed to identify longitudinal and system effects.
- The metric cannot be fully validated until some data on substantive complexity are incorporated. Several of the new questions added in Wave 5 cast some light on this dimension, and once Wave 9 is released there should be enough data to justify developing, if not a scale, then at least a basis for ranking industries and occupations on this criterion. In the long run, however, it is unlikely that HILDA by itself will provide enough information to implement the metric to its best advantage.
- Finally, there is a limit to how long the project can continue relying solely on a found object — even a find as felicitous as HILDA — for its data source. Eventually, even if the funding for HILDA is further extended, it will become necessary to supplement or replace it with a more specialised instrument which can not only fill the data gap on substantive complexity but link the findings on individual employees' experience of work to such matters as workplace culture, work and management practices, and the employer's product or commercial strategy. Ideally this would involve matched or at least coordinated employer surveys to complement the employee data, an option which has been developed with some success as part of the British research program (Green, Mayhew and Molloy 2003).

What does seem clear at this stage is that the addition of skill-intensity to the metric gives it the potential to pick up statistically significant change, some of it probably symptomatic or predictive of larger historical trends, over unprecedentedly short intervals of time. It also creates a basis for eventually formulating unambiguous answers to many of the key questions that eluded the original deskilling debate, even while the metric remains in its present incomplete form. And even at this early stage it has led to the identification of a possible productivity problem affecting highly skilled labour in industries that are critical to Australia's welfare and economic future. Together these findings suggest that the expansion of the Spenner model into a third dimension is a worthwhile step forward in the quantitative analysis of skill change.

Notes

1. Many of the analyses using the first six waves of data are set out in Fraser (2009). Details of the other analyses reported here can be obtained from the author on request.
2. It is true that some cases of unused skill, also known as overskilling, will be frictional (e.g. where an organisation is still adapting its processes to accommodate the skills brought by a new recruit) or transitory (e.g. where a student in the later years of a degree is stacking shelves in a supermarket as a part-time job), while other skills may remain unused simply because they are obsolete or no longer relevant to current practice. However, the use of longitudinal panel data makes it possible to isolate at least the first two of these components from the overall economic impact of longer-term overskilling.
3. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) survey. The HILDA project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in this paper are, however, those of the author and should not be attributed to either FaHCSIA or the MIAESR.
4. COMPLEX is the weakest item in the skill-intensity component, loading more strongly on job stress than on skill-intensity in Wave 5 and moderately on both components in the first two waves. This by itself explains the lesser reliability of the skill-intensity scale, and could be read as evidence that this question straddles the skill-intensity and substantive complexity constructs.
5. Karasek (1979: 291), using data from the US Quality of Employment Survey, recorded an alpha of 0.64 for a seven-item 'job demands' scale. Jonsson (1998: 631) recorded the same figure for a 'human capital growth' scale corresponding broadly to the skill-intensity scale, and 0.63 for an 'autonomy' scale, both derived from the Swedish Level of Living Survey.
6. A non-parametric (and hence marginally less sensitive) test, the Wilcoxon signed-rank test, was used to assess the statistical significance of between-waves change in the component variables, which need to be analysed as ordinal data. NUSKILLS showed the highest variation, achieving statistical significance at the 0.01 level between each of the first four pairs of waves and between 7 and 8. By contrast, OWNTASK did not vary significantly between any two adjacent waves, though the variation over the full eight waves was statistically significant. All the remaining core variables except COMPLEX varied significantly between Waves 1 and 2, while COMPLEX was the only other one to show significant variation between 4 and 5.
7. Among the more recent reviews of this research are de Lange et al. (2003) and Taris and Kompier (2005).

References

- Attewell, P. (1987) 'The deskilling controversy', *Work and Occupations* 14(3), pp. 323–346.
- Attewell, P. (1990) 'What is skill?', *Work and Occupations* 17(4), pp. 422–448.
- Braverman, H. (1974) *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*, Monthly Review Press, New York.
- Buchanan, J. (2006) *From 'Skill Shortages' to Decent Work: The Role of Better Skill Ecosystems*, NSW Board of Vocational Education and Training, Sydney.
- De Lange, A., Taris, T., Kompier, M., Houtman, L. and Bongers, P. (2003), '“The very best of the millennium”: Longitudinal research and the demand-control-(support) model', *Journal of Occupational Health Psychology*, 8(4), pp. 282–305.
- Edquist, C. (2005) 'Systems of innovation: Perspectives and challenges' in J. Fagerberg, D. Mowery, D. and R. Nelson (eds) (2005) *The Oxford Handbook of Innovation*, Oxford University Press, Oxford, pp. 181–208.
- Felstead, A., Gallie, D. and Green, F. (2002), *Work Skills in Britain 1986–2001*, Department for Education and Skills, London.
- Felstead, A., Gallie, D., Green, F. and Zhou, Y. (2007) *Skills at Work, 1986 to 2006*, SKOPE, University of Oxford, Oxford.
- Fraser, D. (2009) Australia's national skilling system: A model and analysis for the period 2001–2006, doctoral thesis, University of Tasmania, available: <http://www.melbourneinstitute.com/hilda/biblio/hbiblio-stu.html> [accessed 17 November 2010].
- Form, W. (1987) 'On the degradation of skills', *Annual Review of Sociology*, 13, pp. 29–47.
- Green, F., Molloy, E., and Mayhew, K. (2003) *Employers' Perspectives Survey*, SKOPE, Oxford: Oxford University, SKOPE, available: <http://www.skope.ox.ac.uk/> [accessed 17 November 2010].
- Jonsson, J. (1998) 'Class and the changing nature of work: Testing hypotheses of deskilling and convergence among Swedish employees', *Work, Employment and Society*, 15(3), pp. 539–563.
- Karasek, R. (1979) 'Job demands, job decision latitude, and mental strain: Implications for job redesign', *Administrative Science Quarterly*, 24, pp. 285–308.
- Maglen, L. and Hopkins, S. (1998) Linking VET to productivity differences: An evaluation of the Prais program, and its implications for Australia, Working Paper 18, Monash University Centre for the Economics of Education and Training.
- Mavromaras, K., McGuinness, S. and Y. K. Fok (2007) Assessing the incidence and wage effects of overskilling in the Australian labour market, IZA Discussion Paper 2837, Bonn, Forschungsinstitut zur Zukunft der Arbeit.
- Payne, J. (2007) 'Skills in context: What can the UK learn from Australia's skill ecosystem project?', *Policy and Politics*, 36(3), pp. 307–323.
- Spenner, K. (1983) 'Deciphering Prometheus: Temporal change in the skill level of work', *American Sociological Review*, 48(6), pp. 824–837.

- Spenner, K. (1985) 'The upgrading and downgrading of occupations: Issues, evidence and implications for education', *Review of Educational Research*, 55(2), pp. 125–154.
- Spenner, K. (1988) 'Social stratification and personality', *Annual Review of Sociology*, 14, pp. 69–97.
- Spenner, K. (1990) 'Skill: Meanings, methods and measures', *Work and Occupations*, 17(4), pp. 399–421.
- Taris, T. and Kompier, M. (2005) 'Job characteristics and learning behavior: Review and psychological mechanisms', *Research in Occupational Stress and Well Being*, 4, pp. 127–166.
- Vallas, S. P. (1990) 'The concept of skill: A critical review', *Work and Occupations*, 17(4), pp. 379–398.
- Wilkins, R., Warren, D., Hahn, M. and Houg, B. (2010) A statistical report on Waves 1 to 7 of the Household, Income and Labour Dynamics in Australia Survey, *Families, Incomes and Jobs*, Volume 5, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
- Windsor, K. (2008) *Skill Ecosystem National Project: Mid-term Evaluation Report*, NSW Department of Education and Training, Darlinghurst.
- Wooden, M. and Watson, N. (2007) 'The HILDA survey and its contribution to economic and social research (so far)', *The Economic Record*, 83(261), pp. 208–231.

Author Details

» **Dr Doug Fraser** is a Visiting Research Fellow at the Australian Defence Force Academy and a Research Associate of the Industrial Relations Research Centre, University of New South Wales. He can be contacted at doug.fraser@tassie.net.au.